Special Event Time Predication for Mine Belt Conveyor Based on Hidden Markov Model

Xiaobin Li[1,2]
1.School of Computer Science and Technology, China University of Mining and Technology, Xuzhou Jiangsu,China
2.School of Computer Science and Technology, Jiangsu Normal University, Xuzhou Jiangsu,China
Email:Wb0817002@163.com

Jiansheng Qian
School of Computer Science and Technology, China University of Mining and Technology, Xuzhou Jiangsu,China
Email: qianjsh@cumt.edu.cn

Zhikai Zhao
School of Computer Science and Technology, China University of Mining and Technology, Xuzhou Jiangsu,China
Email: zhaozhikai2001@163.com

Abstract—**Coal mine belt conveyor can guarantee the coal mine production stable and efficient. On how to effectively predict abnormal accident occurrence time, this paper puts forward a method to predict the abnormal accident occurrence time based on Hidden Markov Model and Hidden Semi-Markov Model. Large amount of time series is collected through belt conveyor protection sensors. The corresponding HMM or HSMM model could be built after feature extraction. At last the accident occurrence time is able to be predicted based on the HMM model or the HSMM model. Experiments carried on the actual production data set illustrate that HMM and HSMM model can effectively predict specific event occurrence time.**

Index Terms—**Hidden Markov Model, Hidden Semi-Markov Model, Belt Conveyor, Prediction**

I. INTRODUCTION

Coal mine belt conveyor plays a key role in coal mine industrial production. It has several advantages such as large transportation, easy to use, well economic effects, easy to maintenance and high reliability. So it has been widely applied into transportation area especially in coal mine industrial production. Not only on the ground but underground, belt conveyor already becomes the major transportation equipment. For example in a coal mine which has one million tons coal production each year, if the belt conveyor break down for one hour because of abnormal accident there will be hundreds of coal lost. Obviously it will bring a significant bad effect for the profit of the coal mine. How to guarantee the belt conveyor running stable is of great importance in coal mine production[1].

Automatic system is now been widely configured in many coal mine production enterprise. These enterprises install kinds of monitor sensors for belt conveyor, so the running status parameters of belt conveyor could be gathered by sensors and transfer to monitor workstation. Inspection of these gathering information people could ascertain the running status of belt conveyor. During the running procedure of belt conveyor, kinds of protection signal such as slider signal, pile signal, slope signal, tear signal and smog signal are installed to monitor the abnormal accident. Once there is an abnormal accident happened, the corresponding protection equipment would suspend the belt conveyor running. If the accident can be predicated through relative technology, people would take measures to remove correspond hidden damages and make the belt conveyor running smoothly.

Fig. 1 illustrates a waveform example gathered from a belt conveyor of a coal mine enterprise. The top graph shows the running power of the belt conveyor (the power data has been preprocessed already). The bottom graph shows the coal piled information. There are three coal piled states: none, light-pile and weight-pile. The belt conveyor is light-pile when the value of pile is 0.5. When

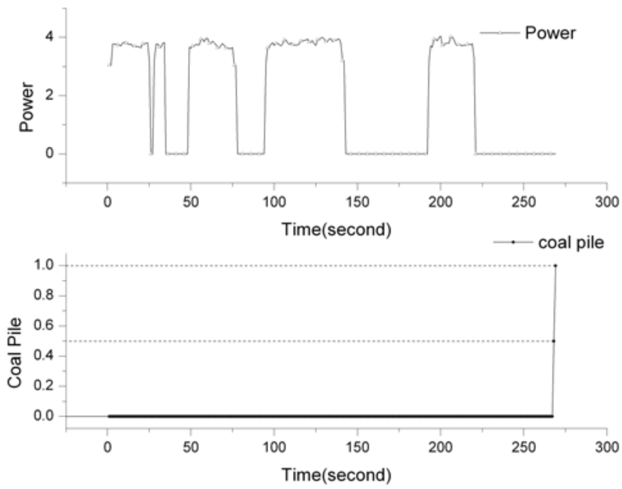

Figure1. A time series of power and coal pile

© 2013 ACADEMY PUBLISHER
doi:10.4304/jsw.8.1.142-150

the value of pile is 1 the belt conveyor is weight-pile and the belt conveyor will shut down spontaneously.

Our work is almost built based on HMM and HSMM. HMM has exhibited its excellent ability in many areas such as speech recognition[2], face recognition[3] and sport[4]. The articles[5] [6] [7] introduce how to apply HMM in the prognostics and health management. Some researchers also apply HMM in Condition Based Maintenance[8]. The article[9] build a continuous HMM model on radar emission and measure the health status of radar by KL distance. Some research [10] draw a conclusion that Markov process can be applied in reliability estimation in multi-state system. Meanwhile, many researchers also raised up some extension structure of HMM. These extensions HMM exhibit stronger performance than HMM in some application area. Because there slacks duration description ability for general HMM some researchers propose HSMM. The article[11] systematically introduces the model structure and inference algorithm of HSMM.

Abnormal accident occurrence time predication is a critical for coal mine production. There is some similarity between the problem and CBM. However, there is particularity in the problem. The reason is accident occurrence could be caused not only from equipment fault but from the outside event.

In this paper we will make feature extraction on the time series data gathered from sensors equipped on conveyor. Then Hidden Markov Model or Hidden Semi Markov Model can be constructed on these data. Based on these models the abnormal accident occurrence time can be predicated. This paper is organized as follows: Sect.2 gives a detail related work in the remaining useful life application with HMM model. Sect.3 provides a general theoretical background for HMM. Sect.4 gives a more detail description for HSMM. Sec.5 provides an algorithm to calculate the duration of state and the occurrence time of special state. Sec.6 gives one actual example. Sec.7 draws conclusions.

II. Related Work

Remaining useful life estimation could judge the remaining available time before a specific operating point in system [1]. This problem is the core issue of condition-based maintenance, fault prognostics and equipment health management system. Remaining useful life of system or components is a random variable. It can be estimated based on life cycle model, expert system or statistical data model. The statistical data model has many advantages such as wide adaptability, low cost, no need of experts' prior knowledge, so it is widely concerned.

Hidden Markov model (HMM) have demonstrated its superior performance in many fields such as speech recognition, it is suitable for none stationary stochastic process signal modeling. The paper [2] first put forward HMM model for monitoring Westland helicopter gears, in which HMM model can estimate component failure and the remaining useful life. The paper [3, 4] apply HMM model in metal cutting tool health assessment, where the model can also be used for fault diagnosis and fault prognostics. In the previous application area HMM model is constructed based on data collected from sensors installed on equipment. Paper [5] makes a model for the health of system by a hierarchical HMM, in which health condition can be represented by the top-level node and remaining useful life can be obtained through state transition probability from a hierarchical HMM. Paper [6] decomposed bearing vibration signals by using wavelet packet, then the decomposed parameters are fed into HMM model. Paper [7] applied HMM model into nozzle working condition monitoring. Paper [8] represent the actual dynamic system based on combination of HMM model and belief rule base; eventually make fault diagnosis and prognostics for complex system.

To overcome the expression ability limitation of HMM for state duration, some scholars put forward HSMM model. Paper [9] introduced HSMM model structure and reasoning algorithm. Paper [10] apply HSMM model into state modeling for UH-60A Black Hawk helicopter main transmission device and hydraulic pump, then make equipment health assessment and fault prognostics. Paper [11] set Gauss distribution as the state duration probability distribution in HSMM model. Multiple sensor information can be combined together according to their weight factor to make fault diagnosis or fault prognostics. Paper [12] put forward an unified framework for fault diagnosis and fault prognostics, where each equipment health state is considered as a HSMM model and another HSMM model is built based on the life cycle model separately. Based on the previous two kinds of HSMM remaining useful life of the system could be estimated. As the HSMM model and the HMM model assume that observations are independent from state, while it is not true in practical application. Paper [13] proposed a regression HSMM model, in which experimental results demonstrate the method is effective in some application. Furthermore paper [14] proposed a non-stationary segmented HSMM model where a time decay factor is introduced to represent the system status gradually degraded. Also paper [15] reach the same purpose based a mixture model combined HSMM model with grey model.

Belt conveyer abnormality occurrence time prediction is an actual demand in industrial field. The problem is similar to fault prognostics while it is particular for exceptions on belt conveyor can be caused with equipment fault or external events. As we know there is little research in this area. On the basic work in paper [16] multi-state system could be considered generally as a Markov process, which may be suit for state modeling for belt conveyor. Here the belt conveyor running states are divided into three main operating parts. We can built HMM and HSMM model based on corresponding extraction features for time series collected from coal mine automation systems. Once these models are built belt conveyor running states could be judged and belt conveyor abnormal occurrence time could be predicted. The on-the-spot user could take effective measures to eliminate corresponding hidden trouble based on prediction results to ensure there is no interruption to coal

© 2013 ACADEMY PUBLISHER

mine production. In this paper coal pile abnormal events, often appeared during belt conveyor running, will be studied as a specific example.

III. HMM MODEL

HMM is a doubly stochastic process [17]. A HMM model usually consists of the following elements as defined below:

(1) $S = \{s_1, s_2, ..., s_N\}$ is a finite set of states where each element means a distinct state.

(2) $V = \{v_1, v_2, ..., v_M\}$ is a set of output symbols.

(3) A is a state transition probability matrix where each

$$a_{ij} = P[q_{t+1} = s_j \mid q_t = s_i], 1 \le i, j \le N \quad (1)$$

means a probability of the transition at q_i to q_{i+1}.

(4) $B = \{b_j(k)\}$ is an observation value probability distribution where

$$b_j(k) = P[v_k \mid q_t = s_j], 1 \le j \le N, 1 \le k \le M \quad (2)$$

(5) $\pi = \{\pi_i\}$ is an initial state probability distribution where each element means a probability of the initial state.

$$\pi_i = P(q_1 = s_i), 1 \le i \le N \quad (3)$$

Usually a more compact model

$$\lambda = (A, B, \pi) \quad (4)$$

is used to represent a HMM model for there is an implication definition N and M in A and B.

A. *Three Basic Problems*

There are three basic problems of interest that must be solved for the model to be useful in real-world applications.

Problem 1: Given the observation sequence $O = O_1O_2...O_T$, and a model $\lambda = (A, B, \pi)$, how do we efficiently compute $P(O \mid \lambda)$, the probability of the observation sequence, given the model?

Problem 2: Given the observation sequence $O = O_1O_2...O_T$, and a model $\lambda = (A, B, \pi)$, how do we choose a corresponding state sequence $Q = q_1q_2...q_T$ which can efficiently explain the observations obey some standards?

Problem 3: How do we adjust the model parameters $\lambda = (A, B, \pi)$ to maximize $P(O \mid \lambda)$?

B. *Solutions to the Previous Problems*

(1) Solution to Problem 1

Firstly we define the forward variable

$$\alpha_t(i) = P(O_1O_2...O_t, q_t = S_i \mid \lambda) \quad (5)$$

i.e., the probability of the partial observation sequence, $O_1O_2..., O_t$, and state S_i at time *t*, given the model λ. We can solve for $\alpha_t(i)$ inductively, as follows:

1) Initialization:

$$\alpha_1(i) = \pi_i b_i(O_1), 1 \le i \le N \quad (6)$$

2) Induction:

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1}), \begin{matrix} 1 \le t \le T-1 \\ 1 \le j \le N \end{matrix} \quad (7)$$

3) Termination:

$$P(O \mid \lambda) = \sum_{i=1}^{N} \alpha_T(i) \quad (8)$$

(2)Solution to Problem 2

Viterbi algorithm can be used to find the single best state sequence, $Q = \{q_1, q_2...q_T\}$, for the given observation sequence $O = \{O_1O_2...O_T\}$, we need to define the quantity

$$\delta_t(i) = \max_{q_1q_2...,q_{t-1}} P[q_1q_2...q_t = i, O_1O_2...O_t \mid \lambda] \quad (9)$$

i.e., $\delta_t(i)$ is the best score (highest probability) along a single path, at time *t*, which accounts for the first *t* observations and ends in state S_i. By induction we have

$$\delta_{t+1}(j) = [\max_i \delta_t(i) a_{ij}] \cdot b_j(O_{t+1}) \quad (10)$$

The complete procedure for finding the best state sequence can now be stated as follows:

1) Initialization:

$$\delta_1(i) = \pi_i b_i(O_1), 1 \le i \le N \quad (11)$$

$$\psi_1(i) = 0 \quad (12)$$

2) Recursion:

$$\delta_t(j) = \max_{1 \le i \le N}[\delta_{t-1}(i) a_{ij}] b_j(O_t), \begin{matrix} 2 \le t \le T \\ 1 \le j \le N \end{matrix} \quad (13)$$

$$\psi_t(j) = \arg\max_{1 \le i \le N}[\delta_{t-1}(i) a_{ij}], \begin{matrix} 2 \le t \le T \\ 1 \le j \le N \end{matrix} \quad (14)$$

3) Termination:

$$P^* = \max_{1 \le i \le N}[\delta_T(i)] \quad (15)$$

$$q_T^* = \arg\max_{1 \le i \le N}[\delta_T(i)] \quad (16)$$

4) Path backtracking:

$$q_t^* = \psi_{t+1}(q_{t+1}^*), t = T-1, T-2, ..., 1 \quad (17)$$

(3)Solution to Problem 3

To solve the problem 3, the backward variable $\beta_t(i) = P(O_{t+1}O_{t+2}...O_T \mid q_t = S_i, \lambda)$. should be defined.

i.e., the probability of the partial observation sequence from *t*+1 to the end, given state S_i at time *t* and the model λ. Also we can solve for $\beta_t(i)$ inductively.

© 2013 ACADEMY PUBLISHER

1) Initialization:

$$\beta_T(i) = 1, 1 \le i \le N \quad (18)$$

2) Induction:

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(O_{t+1}) \beta_{t+1}(j), \begin{array}{l} t = T-1, T-2, ..., 1 \\ 1 \le i \le N \end{array} \quad (19)$$

Once we get forward variable $\alpha_t(i)$ and backward variable $\beta_t(i)$, we can define $\xi_t(i,j)$, the probability of being in state S_i at time *t*, and state S_j at time *t*+1, given the model and observation sequence, i.e.

$$\begin{aligned} \xi_t(i,j) &= P(q_t = S_i, q_{t+1} = S_j \mid O, \lambda) \\ &= \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{P(O \mid \lambda)} \\ &= \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)} \end{aligned} \quad (20)$$

Then we can define

$$\gamma_t(i) = \sum_{j=1}^{N} \xi_t(i,j) \quad (21)$$

as the probability of begin in state S_i at time *t*, given the observation sequence and the model.

For a given observation sequence *O*, we can obtain the expected number of times that state S_i is visited and the expected number of transitions from state S_i to state S_j, i.e,

$$\gamma = \sum_{t=1}^{T-1} \gamma_t(i) \quad (22)$$

is used as the expected number of transitions from S_i

$$\xi = \sum_{t=1}^{T-1} \xi_t(i,j) \quad (23)$$

is used as the expected number of transitions from S_i to S_j.

Then the three model parameters of HMM could be reasoned from the following three formulas.

$$\overline{\pi_i} = \gamma_1(i) \quad (24)$$

$$\overline{a_{ij}} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \quad (25)$$

$$\overline{b_j}(k) = \frac{\sum_{\substack{t=1 \\ s.t.O_t = v_k}}^{T} \gamma_t(j)}{\sum_{t=1}^{T} \gamma_t(j)} \quad (26)$$

For some applications when the observations are continuous it is possible to quantize these signals via codebooks, etc. Also continuous observation density could be introduced to model an HMM[17].

IV. HSMM Model

Hidden Markov Model has been successfully applied in many areas. In the model any hidden state has its correspond state transition probability a_{ii}, general $a_{ii} \ne 0$. Meanwhile any state extension cycle probability density function $p_i(d)$ can be calculated:

$$p_i(d) = (a_{ii})^{d-1}(1 - a_{ii}) \quad (27)$$

d means the adjacent conditions are maintained for *d* time unit with a fixed state S_i.

Apparently, *d* obeys geometric distribution which implies the default duration in HMM follows geometric distribution, which is obviously limited in practical field.

In HSMM [9, 17] a specific probability distribution can freely be given for each state, so that the hidden state for several continuous observed symbols is constant. In some researcher's opinion, HSMM can also be called clear cycle HMM, variable cycle HMM, generalization HMM, segmented HMM, etc.

A. HSMM Definition

The following is related HSMM definitions [11]. Suppose that a discrete markov chain is the state set $Q = \{1, ..., M\}$. Then the state sequence is $S_{1:T} \equiv S_1, ..., S_T$,In which $S_t \in Q$. Then some expressions are defined as follows:

(1) $S_{t_1:t_2} = i$ means that the system state from t_1 to t_2 always remain in state *i*, that implies $S_{t_1} = i$, $S_{t_1+1} = i$,..., $S_{t_2} = i$,meanwhile the previous state S_{t_1-1} and the latter state S_{t_2+1} may not necessarily be *i*.

(2) $S_{[t_1:t_2]} = i$ means that the system state from t_1 to t_2 always remain in the state *i*, that implies $S_{t_1} = i$, $S_{t_1+1} = i$,..., $S_{t_2} = i$,meanwhile the previous state S_{t_1-1} and the latter state S_{t_2+1} must not be *i*.

(3) $S_{[t_1:t_2} = i$ means that the system state from t_1 to t_2 always remain in state *i*, that implies $S_{t_1} = i$, $S_{t_1+1} = i$,..., $S_{t_2} = i$, meanwhile the previous state S_{t_1-1} must not be *i*, and the latter state S_{t_2+1} may not necessarily be *i*.

(4) $S_{t_1:t_2]} = i$ means that the system state from t_1 to t_2 always remain in state *i*, that implies

© 2013 ACADEMY PUBLISHER

$S_{t_1}=i$, $S_{t_1+1}=i$,..., $S_{t_2}=i$, meanwhile the previous state S_{t_1-1} may not necessarily be *i*, and the latter state S_{t_2+1} must not be *i*.

$O_{1:T}=O_1,...,O_T$ is an observation sequence where $O_t\in V$ is the observation symbol at time *t*. $V=\{v_1,v_2,...,v_k\}$ is observation symbols set. The former observation sequence *O* is corresponding to hidden states sequence $S_{1:T}$, in which $S_{1:d_1]}=i_1$, $S_{[d_1+1:d_1+d_2]}=i_2$,..., $S_{[d_1+...+d_{1n-1}+1:d_1+...+d_n]}=i_n$,for $m=1,...,n-1$ .There is a state transition $(i_m,d_m)\rightarrow(i_{m+1},d_{m+1})$ in which a formula

$$\sum\nolimits_{m=1}^{n} d_m=T \ . \tag{28}$$

must be satisfied. Besides, $i_1,...,i_n\in Q$, $d_1,...,d_n\in D$ where *D* is a positive integer which means each state duration.

Here a compact form of HSMM $\lambda=(A,B,\pi)$ is given:

(1) *A* is state transition probability matrix where each element in the matrix means a state transition probability

$$a_{(i,d')(j,d)} \overset{def}{=} P[S_{[t+1:t+d]}=j \mid S_{[t-d'+1:t]}=i] \tag{29}$$

(2) *B* is observation probability matrix where each element means a probability of observation symbol stayed in a specific state

$$b_{j,d}(o_{t+1:t+d}) \overset{def}{=} P[o_{t+1:t+d} \mid S_{[t+1:t+d]}=j] \tag{30}$$

(3) π is the initial state probability distribution where

$$\pi_{j,d} \overset{def}{=} P[S_{[t-d+1:t]}=j], t\le 0, d\in D \tag{31}$$

B. *HSMM Parameter Estimation*

If HSMM parameter is unknown, some existing relevant prior knowledge of sequence O is needed to estimate HSMM parameters.

As is similar to HMM structure, the forward and backward variables are defined.

Forward variable:

$$\alpha_t(j,d) \overset{def}{=} P[S_{[t-d+1:t]}=j, o_{1:t} \mid \lambda] \tag{32}$$

Backward variable:

$$\beta_t(j,d) \overset{def}{=} P[o_{t+1:T} \mid S_{[t-d+1:t]}=j,\lambda] \tag{33}$$

Once forward variable is given the probability of a sequence complied with a specific HSMM model can be determined as

$$P(O\mid\lambda)=\sum_{i=1}^{N}\alpha_T(i) \tag{34}$$

Meanwhile the model parameter can be substituted with the former formula so that

$$\alpha_t(j,d) = \sum_{i\in S\setminus\{j\}}\sum_{d'\in D}\alpha_{t-d}(i,d')\cdot a_{(i,d')(j,d)}\cdot b_{j,d}(o_{t-d+1:t}) \tag{35}$$

and

$$\beta_t(j,d) = \sum_{i\in S\setminus\{j\}}\sum_{d'\in D} a_{(j,d)(i,d')}\cdot b_{i,d'}(o_{t+1:t+d'})\cdot\beta_{t+d'}(i,d') \tag{36}$$

can be deduced.

Based on the above two definition the following expression could be determined.

$$\eta_t(j,d) \overset{def}{=} P[S_{[t-d+1:t]}=j, o_{1:T} \mid \lambda] = \alpha_t(j,d)\beta_t(j,d)) \tag{37}$$

$$\xi_t(i,d':j,d) \overset{def}{=} P[S_{[t-d'+1:t]} = i, S_{[t+1:d]}=j, o_{1:T}\mid\lambda] = \alpha_t(i,d')a_{(i,d')(j,d)}b_{j,d}(o_{t+1:t+d})\beta_{t+d}(j,d) \tag{38}$$

$$\xi_t(i,j) \overset{def}{=} P[S_{t]}=i, S_{[t+1}=j, o_{1:T}\mid\lambda] = \sum_{d'\in D}\sum_{d\in D}\xi_t(i,d':j,d) \tag{39}$$

$$\gamma_t(j) \overset{def}{=} P[S_t=j, o_{1:T}\mid\lambda] = \sum_{\tau\ge t}\sum_{d=\tau=t+1}\eta_\tau(j,d) \tag{40}$$

Furthermore the HSMM parameter weight estimation process is as follow:

(1) An initial HSMM model parameters λ_0 should be assumed.

(2) According to the existing parameters λ_k $\alpha_t(j,d)$, $\beta_t(j,d)$, $\eta_t(j,d)$, $\xi_t(i,d':j,d)$, $\xi_t(i,j)$ and $\gamma_t(j)$ are calculated respectively.

(3) The new HSMM model parameters can be calculated based on result generated from (2). Initial distribution

$$\hat{\pi}_{j,d} = \frac{\eta_t(j,d)}{\sum_{j,d}\eta_t(j,d)} \tag{41}$$

transition probability matrix

$$\hat{a}_{(i,d')(j,d)} = \frac{\sum_t \xi_t(i,d':j,d)}{\sum_{j\ne i}\sum_t \xi_t(i,d':j,d)} \tag{42}$$

© 2013 ACADEMY PUBLISHER

and observation matrix

$$\hat{b}_{j,d}(v_{k1}:v_{kd}) = \frac{\sum_t [\eta_t(j,d)\cdot I(o_{t+1:t+d} = v_{k1:kd})]}{\sum_t \eta_t(j,d)} \quad (43)$$

can be calculated. Meanwhile when $o_{t+1} = v_{k1},..., o_{t+d} = v_{kd}$, $I(o_{t+1:t+d} = v_{k1:kd}) = 1$, otherwise $I(o_{t+1:t+d} = v_{k1:kd}) = 0$.

(4) Repeat step (2) (3) until $L(\lambda_k) = P[o_{1:T} \mid \lambda_k]$ convergence to a fixed point.

V. State Occurrence Time Prediction Algorithm

A. State Occurrence Time Calculation

Inspiration by state duration and remain useful life predication algorithm proposed in literature [11, 18] a state occurrence time calculation algorithm is proposed as below.

The current system state is *i* and the current time is T_i^t . Assume the state firstly appeared time is T_i^0 . Based on the duration time probability distribution expected duration of current state can be calculated as T_i^u . Then the duration under current state *i* can be adjusted as $T_i^u = T_i^u - (T_i^t - T_i^0 + 1)$.

Assume predicted state *j*=*i*+*k*,*k*=1,2,...,remain useful life to state j ,i.e. L_j ,could be calculated based on the following two step recursive algorithm:

(1) When state is *j*-1

$L_{j-1} = a_{j-1,j-1}(T_{j-1}^u + T_j^u) + a_{j-1,j}T_j^u$ could be obtained.

(2) When state is *j*-*k*(*k*>1)

$L_{j-k} = a_{j-k,j-k}(T_{j-k}^u + L_{j-k+1}) + a_{j-k,j-k+1}L_{j-k+1}$ could be calculated. Obviously j appeared moment is $T_j^t = T_i^t + L_i$.

B. Time Prediction Algorithm

(1) Feature extraction on collected data.

(2) Parameter estimation of HSMM and HMM model training.

(3) Duration probability distribution on each state can be deduced from HSMM or HMM model.

(4) Classify a new sample based on the former models and obtain the current state of this sample.

(5) Based on the current state, duration probability distribution and expected time estimation method suggested in the previous section, future state occurrence time could be obtained.

VI. Experiments

In order to evaluate the prediction algorithm suggested in this paper some software such as R, mhsmm software package [19] and rhmm are used here. The goal is to predict actual protection occurrence time based on the collected data of belt conveyor from Pingdingshan 8th Coal Mine in China. Experimental computer is equipped with CPU Intel Core (TM) 2 Quad 2.66GHz, 1.85G memory.

A. Data Set

We collected sensor data which is range from October 1, 2009 00: 01: 00.000 to 2009 October 31st 23: 59: 00.000 of D-two belt in Pingdingshan 8th Coal Mine. The total records number is 401760. There are nine sensor points equipped in belt: D-two 1 # motor power, D-two skid output, D-two motor feedback, D-two piles of coal feedback, D-two airframe weight deviation feedback, D-two pull rope feedback, D-two tear feedback, D-two material feedback (control sprinkler), and D-two smoke feedback. Each sensor points collect 44640 records and the interval among these records is 1 minute.

There are some periods in which belt did not run. It is necessary to preprocess these samples. Firstly a definition is given that useful sample is referred as that motor power sensor value is not zero. Meanwhile coal pile feedback record value of 0.5 is mild coal pile and record value of 1 is severe coal pile. When the belt conveyor is in mild coal pile, belt conveyor generally did not stop. While belt conveyor is in severe coal pile it is necessary to shut down machine immediately. For artificial processing duration after shutdown cannot be estimated, we define valid sample data as the sequence between power value not zero and coal pile. After these preprocess we obtained 16 valid data sample point.

Afterwards a log transformation is applied into power data as to meet the requirements of statistical. The conversion rule is $PW_n = \log(PW + 1)$ where PW is the original power data and PW_n is the data after log transformation. Fig. 1 illustrates a sample sequence transformed.

B. State and Initial Parameters

According to practical coal pile of belt conveyor during its' running, belt can be divided into three states: normal operation, mild coal pile, severe coal pile (later normal, mild and severe is used to represent these three states respectively). Practical state transition relationship among these three states can be illustrated as Fig. 2.

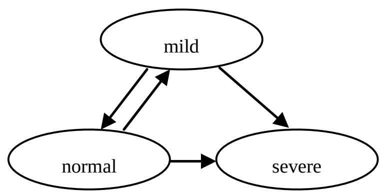

Figure2. Coal pile state transition of belt conveyor

In HMM model initial state probability distribution is set as $\pi_0 = (1,0,0)$,observation probability distribution parameters B_0 is set as a three-dimensional Gaussian distribution where mean vector $u_0 = (0,0,0)$,variance

© 2013 ACADEMY PUBLISHER

vector $\delta_0 = (1,1,1)$ distribution. State transition probability matrix is list in Table 1.

TABLE I.

INITIAL STATE TRANSFER PROBABILITY OF HMM

state	normal	mild	severe
normal	0.00000	0.80000	0.20000
mild	0.50000	0.00000	0.50000
severe	0.00000	0.00000	1.00000

TABLE II.

INITIAL STATE TRANSFER PROBABILITY OF HSMM

state	normal	mild	severe
normal	0.00000	0.80000	0.20000
mild	0.50000	0.00000	0.50000
severe	0.10000	0.90000	0.00000

For a state can not turn to itself in a HSMM model, its initial transition probability matrix is set in Table 2. Other parameters such as initial probability distribution, observed symbol probability distribution are set as the same as the above HMM model. Lastly state duration probability distribution is set as no argument type (non-parametric).

C. Model Parameters Trained

Model parameters could be obtained after model training on samples. State transition probability of HMM model during belt running is show in Table 3. While Table 4 shows transition probability of HSMM model. Table 5 shows other parameters such as initial probability distribution π , mean vector of observation probability distribution μ and covariance vector δ .

TABLE III.

STATE TRANSITION PROBABILITY OF HMM AFTER LEARNING

state	normal	mild	severe
normal	0.96502	0.01107	0.02391
mild	0.00705	0.98072	0.01223
severe	0.29542	0.30043	0.40415

TABLE IV.

STATE TRANSITION PROBABILITY OF HSMM AFTER LEARNING

state	normal	mild	severe
normal	0.00000	0.99460	0.00540
mild	0.05200	0.00000	0.94800
severe	0.00059	0.99941	0.00000

TABLE V.

PARAMETERS OF MODEL

	HSMM	*HMM*
π	(0.8125,0.1819,0.0056)	(1.0000,0.0000,0.0000)
μ	(0.0028,3.7518,3.7685)	(0.0019,3.7727,3.0694)
δ	(0.0002,0.3753,0.0101)	(0.0000,0.0120,1.1323)

D. State Estimation

Before predicting belt conveyor abnormal occurrence time, it is necessary to estimate the current running state of belt conveyor. Precision and recall are two indicators used here to make state assessment. Table 6 lists precision of sequence data classification using HMM and HSMM model. Table 7 lists recall rate of sequence data classification using the two models.

TABLE VI.

ACCURACY OF STATE JUDGMENT

model	normal	mild	severe
HSMM	0.99405	0.00165	0.00000
HMM	0.99451	0.00000	0.00455

TABLE VII.

RECALL OF STATE JUDGMENT

model	normal	mild	severe
HSMM	0.34287	0.41176	0.00000
HMM	0.35119	0.00000	0.12500

TABLE VIII.

EXPECTATION AND DEVIATION OF STATE DURATION PROBABILITY DISTRIBUTION

state	*HSMM*		*HMM*	
	μ	δ	μ	δ
normal	28.3	38.1	28.4	40.5
mild	2.7	0.6	55.9	71.0
severe	1.0	0.0	1.7	1.1

E. State Duration Distribution

By modeling on sequence data, state duration distribution can be obtained. Table 8 lists the expectation and deviation of duration for three states in the two models.

© 2013 ACADEMY PUBLISHER

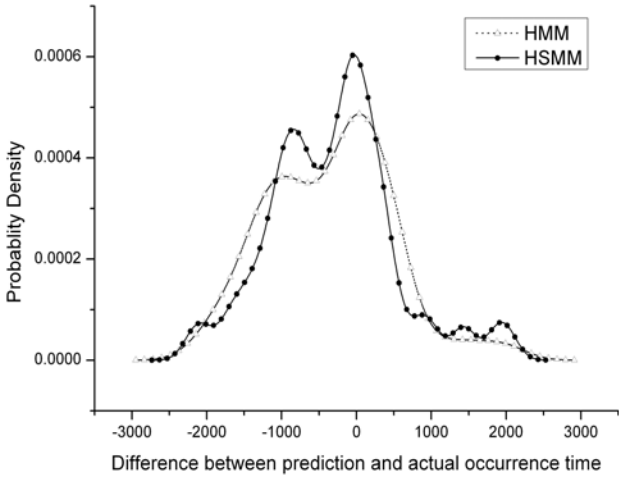

Figure3. The probability density of difference time between the predict occurrence time and actual occurrence time to mild coal pile

F. *Abnormal Occurrence Time Prediction*

In order to compare time prediction effect between HMM and HSMM, Fig. 3 shows the time point prediction for mild and severe coal pile under HMM where the probability density of difference time between the prediction and actual occurrence time is illustrated. While the Fig. 4 shows the probability density of difference time between the prediction and actual occurrence time under HSMM. We can conclude that HSMM model is better than HMM model in predicting occurrence time between mild coal pile and severe coal pile. For mild coal pile time prediction, the difference between prediction and actual time is distributed near zero, that is to say, the predictions are consistent with actual situation. While for severe coal pile time prediction, the sample amount is relatively rare, so the prediction result has large deviation. Besides, occurrence time span of mild coal pile in training samples is relatively large so that the model learned will give a larger span prediction for mild pile.

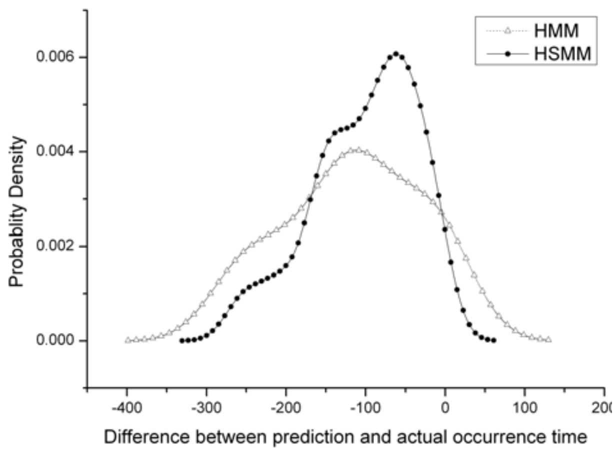

Figure4. The probability density of difference time between the predict occurrence time and actual occurrence time to severe coal pile

VII. CONCLUSION

Coal mine belt conveyor can guarantee coal mine production stable and efficient. On how to effectively predict abnormal accident occurrence time, this paper puts forward a method to predict the abnormal accident occurrence time based on Hidden Markov Model (HMM) and Hidden Semi-Markov Model (HSMM). Large amount of time series is collected through belt conveyor protection sensors firstly. The corresponding HMM or HSMM model could be built after feature extraction. At last accident occurrence time is able to be predicted based on the HMM model or the HSMM model. Experiments carried on the actual production data set illustrate that HSMM model can effectively predict specific event's occurrence time. In terms of the effect of time point prediction, HSMM is better than HMM model.

Nevertheless, in this paper HMM default has geometric distribution for state duration, but HSMM model directly selects probability distribution of non-parametric which resulted in defect of large span in predicting final time points. So how to select an appropriate HSMM model extension time distribution function in specific application is a worthwhile topic. Meanwhile here we assumed that false alarm of normal condition is equivalent with false positive of coal pile but false positive of coal pile is more serious than false alarm of normal operation condition. It is a difficult problem to predict abnormal occurrence time with imbalance cost.

Finally, due to the constraints of the production environment, the data acquired often contains kinds of noise and missing values. So how to find more effective method to improve prediction accuracy of abnormal events is a future topic.

ACKNOWLEDGEMENT

This work was supported by Chinese National High Technology Research and Development Program 863 under Grant No 2008AA062200, and by Chinese National High Technology Research and Development Program 863 under Grant No 2012AA062103, and by Jiangsu Province Production Research Foundation under Grant No BY2009114 in China, and by Xuzhou Industry Science Program under Grant No XX10A001, and by Jiangsu Normal University Foundation under Grant No 10XLA13. Also the authors wish to thank the reviewers for this paper.

REFERENCES

[1] X. S. Si, W. Wang, C. H. Hu and D. H. Zhou, Remaining useful life estimation-A review on the statistical data driven approaches. European Journal of Operational Research, 2011.

[2] C. Bunks, D. McCarthy and T. Al-Ani, Condition-based maintenance of machines using hidden Markov models. Mechanical Systems and Signal Processing, 2000. 14(4): p. 597-612.

[3] P. Baruah and R. B. Chinnam, HMMs for diagnostics and prognostics in machining processes. International Journal of Production Research, 2005. 43(6): p. 1275-1293.

© 2013 ACADEMY PUBLISHER

[4] R. B. Chinnam and P. Baruah. Autonomous diagnostics and prognostics through competitive learning driven HMM-based clustering. in Neural Networks, 2003. Proceedings of the International Joint Conference on. 2003.
[5] F. Camci and R. B. Chinnam, Health-State Estimation and Prognostics in Machining Processes. Automation Science and Engineering, IEEE Transactions on, 2010. 7(3): p. 581-597.
[6] H. Ocak, K. A. Loparo and F. M. Discenzo, Online tracking of bearing wear using wavelet packet decomposition and probabilistic modeling: A method for bearing prognostics. Journal of Sound and Vibration, 2007. 302(4-5): p. 951-961.
[7] A. H. Tai, W. K. Ching and L.Y. Chan, Detection of machine failure: Hidden Markov Model approach. Computers & Industrial Engineering, 2009. 57(2): p. 608-619.
[8] Z. J. Zhou, C. H. Hu, D. L. Xu, M. Y. Chen and D. H. Zhou, A model for real-time failure prognosis based on hidden Markov model and belief rule base. European Journal of Operational Research, 2010. 207(1): p. 269-283.
[9] S. Z. Yu, Hidden semi-Markov models. Artificial Intelligence, 2010. 174(2): p. 215-243.
[10] M. Dong, D. He, P. Banerjee and J. Keller, Equipment health diagnosis and prognosis using hidden semi-Markov models. The International Journal of Advanced Manufacturing Technology, 2006. 30(7): p. 738-749.
[11] M. Dong and D. He, Hidden semi-Markov model-based methodology for multi-sensor equipment health diagnosis and prognosis. European Journal of Operational Research, 2007. 178(3): p. 858-878.
[12] M. Dong and D. He, A segmental hidden semi-Markov model (HSMM)-based diagnostics and prognostics framework and methodology. Mechanical Systems and Signal Processing, 2007. 21(5): p. 2248-2266.
[13] M. Dong, A novel approach to equipment health management based on auto-regressive hidden semi-Markov model (AR-HSMM). Science in China Series F: Information Sciences, 2008. 51(9): p. 1291-1304.
[14] M. Dong and Y. Peng, Equipment PHM using non-stationary segmental hidden semi-Markov model. Robotics and Computer-Integrated Manufacturing, 2010.
[15] Y. Peng and M. Dong, A hybrid approach of HMM and grey model for age-dependent health prediction of engineering assets. Expert Systems with Applications, 2011. 38(10): p. 12946-12953.
[16] M. Muhammad, A. Majid, and M. Amin, Reliability and Availability Evaluation for a Multi-state System Subject to Minimal Repair. Journal of Applied Sciences, 2011, 2011.
[17] L. R. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 1989. 77(2): p. 257-286.
[18] Z. Y. Xia, B. Yang, C. H. Yang and J. J. Yang, Equipment Fault Diagnosis Based on State Space Model and Improved Hidden Semi-Markov Model.Mathematics in practice and theory,2011,Vol.41,No.3
[19] J. Oi Connell and S. H jsgaard, Hidden Semi Markov Models for Multiple Observation Sequences-The mhsmm Package for R. Journal of Statistical Software, 2009.

Xiaobin Li He received the B.S. and M.S. degrees in computer science from China University of Mining and Technology in 1999 and 2004, respectively. He is currently pursuing the Ph,D. degree in computer science at the China University of Mining and Technology. His research interests include machine learning and data mining.

Jiansheng Qian He is now a professor in China University of Mining and Technology. His research interests include signal processing , computer network and machine learning.

Zhikai Zhao He is now a Ph,D. candidate in China University of Mining and Technology. His research interests include machine learning and data mining

© 2013 ACADEMY PUBLISHER