Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 693450, 13 pages
http://dx.doi.org/10.1155/2015/693450

Research Article

Modeling the Process of Event Sequence Data Generated for Working Condition Diagnosis

Jianwei Ding,[1,2,3] Yingbo Liu,[2,3] Li Zhang,[2,3] and Jianmin Wang[2,3]

[1]*Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China*
[2]*Institute of Information System & Engineering, School of Software, Tsinghua University, Beijing 100084, China*
[3]*School of Software, Tsinghua University, East Main Building, Beijing 100084, China*

Correspondence should be addressed to Jianwei Ding; dingjw09@mails.tsinghua.edu.cn

Received 1 June 2015; Accepted 5 July 2015

Academic Editor: Xiaoyu Song

Condition monitoring systems are widely used to monitor the working condition of equipment, generating a vast amount and variety of telemetry data in the process. The main task of surveillance focuses on analyzing these routinely collected telemetry data to help analyze the working condition in the equipment. However, with the rapid increase in the volume of telemetry data, it is a nontrivial task to analyze all the telemetry data to understand the working condition of the equipment without any a priori knowledge. In this paper, we proposed a probabilistic generative model called working condition model (WCM), which is capable of simulating the process of event sequence data generated and depicting the working condition of equipment at runtime. With the help of WCM, we are able to analyze how the event sequence data behave in different working modes and meanwhile to detect the working mode of an event sequence (working condition diagnosis). Furthermore, we have applied WCM to illustrative applications like automated detection of an anomalous event sequence for the runtime of equipment. Our experimental results on the real data sets demonstrate the effectiveness of the model.

1. Introduction

Currently, with the rapid development of technology for *the Internet of Things* [1, 2], condition monitoring systems (CMSs) [3, 4] are widely used to monitor the working condition of equipment. *KOMTRAX* (*KOMTRAX*: http://www.komatsuamerica.com/komtrax) from Komatsu and *IEM* (*IEM*: http://www.sanygroup.com/group/en-us/) from SANY are well-known CMSs that generate a large amount of telemetry data while monitoring the working condition of equipment at runtime. Event sequence data especially are one main type of telemetry data, which record a sequence of the operations on the equipment. If we can analyze the working mode of equipment according to these event sequence data, it will help us better understand the working condition of equipment at runtime.

Diagnosis of working condition of equipment mainly depends on analyzing the telemetry data collected at the runtime of the equipment. In most CMSs, data analysis is the only way for engineers to diagnose the working condition of equipment. Telemetry data mainly contain operation events, performance counters, alert events, and others in the real CMSs. Most telemetry data can be classified into two categories: continuous time series data and temporal event data. Time series data is a sequence of real-valued data points, captured and sampled typically at successive time points equally spaced with a uniform time interval. For example, the engine temperature of equipment is a typical example of time series in the CMSs. An event sequence in CMSs is used to record the occurrences of a specific message indicating that something such as an operation has happened in the equipment. For example, an event sequence of "pumping concrete" in a concrete pump truck contains an event *pumping concrete*, which represents the idea that the concrete pump truck starts pumping the concrete.

As illustrated in Figure 1, an event sequence records all the specific operations on the equipment at runtime, which usually provides enough information to engineers for working condition diagnosis of equipment. The previous studies on analysis of event sequence data for working

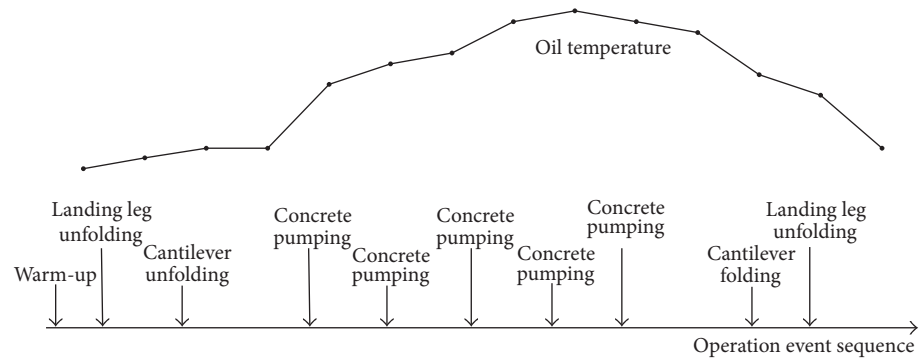

Figure 1: Actual example of telemetry data of a concrete pump truck at one runtime. Oil temperature is a typical continuous time series, which records the variation of oil temperature at one runtime of a concrete pump truck. Operation event sequence is a typical event sequence, which records the occurrences of six operation events (*warm-up, landing leg unfolding, cantilever unfolding, concrete pumping, landing leg folding,* and *cantilever folding*) on the concrete pump truck at one runtime.

condition diagnosis are mainly grouped into two categories: the correlation analysis between distinct event sequences [6–8] and events based process mining algorithms [9, 10]. The correlation analysis between distinct event sequences usually provides useful hints for causality analysis. Although correlated metrics may not exactly be the root causes of events, they could also provide intermediate useful information that pinpoints the root causes of events. The process mining algorithms focus on the occurrence order of distinct events with the help of process models [11, 12], which mainly indicate the working process of equipment.

However, the occurrence frequency of events also provides us with important information for working condition diagnosis, which is ignored by the previous studies. We take a concrete pump truck as an example. As illustrated in Figure 1, at a normal runtime of concrete pump truck, a concrete pump truck needs five *concrete pumping* events to finish pumping a hopper of the concrete. With the wear and tear of concrete pump truck, the concrete pump truck needs seven or more *concrete pumping* events to finish pumping a hopper of the concrete. Although the occurrence order of the six events (*warm-up* → *landing leg unfolding* → *cantilever unfolding* → *concrete pumping* → *landing leg folding* → *cantilever folding*) is the same, yet the working condition of the concrete pump truck has changed a lot. If the occurrence frequencies of events are taken into consideration for working condition diagnosis, it will enhance the ability of working condition diagnosis for engineers, which will help us better understand the working condition of equipment.

In this paper, we proposed a probabilistic generative model called working condition model (WCM), which is capable of depicting the working condition of equipment at runtime. According to event sequence data in different working condition of equipment, we simulated the process of event sequence data generated in order to get the WCM of equipment. Furthermore, with the help of WCM, we extended the application of event sequence data to more domains such as anomaly detection and the variation trend analysis of working condition. Motivated by the real requirement of working condition diagnosis, our working condition model tries to answer the following three questions:

(a) How many types of working modes (details in Section 3) does the equipment have at runtime?

(b) In each type of working mode, how does an event sequence behave?

(c) For a new event sequence at runtime, which type of working mode does it belong to?

Our evaluation consists of multiple phases. First, we model the WCM of the real event sequence data sets collected from 279 concrete pump trucks over a period of 6 months. We analyze the performance of the WCM of the concrete pump truck. Then, we apply the WCM of the concrete pump truck for more applications including anomaly detection and the variation trend analysis of working condition.

Our work presents a probabilistic generative model named WCM to simulate the process of event sequence data generated and to depict the working condition of equipment at runtime. The contributions of this paper are as follows:

(i) Motivated by real applications, we propose the WCM to depict the working condition of equipment. To the best of our knowledge, this is the first attempt to simulate the process of event sequence data generated for working condition diagnosis.

(ii) We illustrate two useful applications based on WCM: automated detection for a new work cycle and automated detection for anomalous work cycles.

(iii) The experiments on real data from a well-known Chinese construction machinery manufacturer show the effectiveness of our model.

The rest of the paper is organized as follows: In Section 2, we introduce some related works. The problem statement and formulation are introduced in Section 3. We introduce our approach in Sections 4 and 5. The empirical evaluation is shown in Section 6. Finally, we conclude our work in Section 7.

2. Related Work

2.1. Analysis of Event Sequence Data. An event is a happening of interest [13, 14]. In the surveillance of equipment, the interest in events comes mostly from the state of equipment changes that are produced by equipment manipulation operations [15]. Example events in the actual surveillance of the concrete pump truck include *warm-up, landing leg unfolding, cantilever unfolding, concrete pumping, landing leg folding,* and *cantilever folding* as shown in Figure 1. When a sequence of events takes place, we refer to these occurrences to get the event sequence data. The main idea of analysis of event sequence data is to process events to gather meaningful or valuable information and then to derive actions from them.

Events in an event sequence are often interrelated and form complex relationships. The correlation analysis of event sequence data [7, 8, 16, 17] focuses on detecting these relationships and is extended to other related applications such as anomaly detection [18–20]. A temporal, spatial, or causal relationship of events can determine the partial order between events [16]. Hence, event sequence data based process mining algorithms focus on the causal relationship of events by analyzing the occurrence order of distinct events [21, 22]. There have been many existing process mining algorithms [23, 24] and tools [25, 26] to mine the causal relationship of events, which is capable of instructing engineers to better understand the operation procedure of equipment [27].

2.2. Working Condition Diagnosis. Working condition of equipment [28] is the condition in which the equipment works, including but not limited to such things as amenities, physical environment, stress and noise levels, degree of safety or danger, and the like. The working condition diagnosis usually uses specific models or variables for different applications. In the correlation analysis of event sequence data for working condition diagnosis, the correlation coefficients of event sequence data, for example, the Pearson correlation [6, 7, 29, 30] and the Rank correlation [31], are used to depict the working conditions of equipment. In the process mining algorithms of event sequence data, varieties of process models, for example, Petri net [32] and business process modeling notation [33, 34], are specific to depicting the working condition of equipment for different process mining algorithms. In this paper, we take the occurrences frequencies of events into consideration and simulate the process of event sequences generated in different working modes of equipment. Hence, we use the occurrence probability of events in the event sequence to depict the working condition of equipment at runtime.

2.3. Probabilistic Generative Model. In probability and statistics, a generative model [35] is a model for randomly generating observable data, typically given some hidden parameters. It specifies a joint probability distribution over observable data. Generative models are used in machine learning [36] either for modeling data directly (i.e., modeling observations drawn from a probability density function), or as an intermediate model to forming a conditional probability density function. A conditional distribution can be formed from a generative model through the Bayesian rule [36].

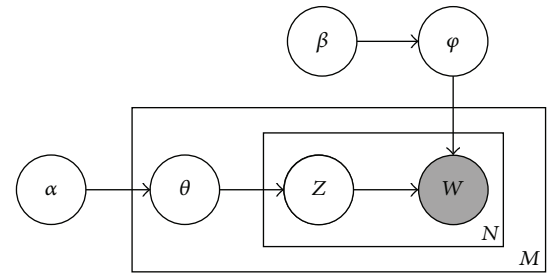

Figure 2: Graphic model of a typical generative model LDA. The boxes are plates representing replicates. The outer plate represents documents, while the inner plate represents the repeated choice of topics and words within a document. M denotes the number of documents and N the number of words in a document. $\alpha, \beta, \theta, \varphi, Z$ are hidden parameters and W is observations. Details about LDA refer to [5].

For example, latent Dirichlet allocation (LDA) [5, 37, 38] is a typical generative model, which is widely used in many domains. In natural language processing, LDA is capable of simulating the process of documents generated well, where observations are words collected into documents, and it posits that each document is a mixture of a small number of topics and that each word's creation is attributable to one of the document's topics. Figure 2 illustrates the graphic model [39] of LDA [5]. With plate notation, the dependencies among variables can be captured concisely.

3. Terminology and Notation

As the working condition of equipment is always corresponding to a period, for example, one day or one week, we first determine the basic unit of observation for working condition diagnosis, intended to ease the working condition diagnosis according to the event sequence data.

Definition 1 (work cycle). A *work cycle* of a piece of equipment, denoted by s, is a complete work period, that is, a complete usage period of the equipment from the time the equipment starts working until it shuts down.

We define the idea that s consists of elements that are integers from $1, \ldots, S$, where S is the number of work cycles. There is one important advantage in adopting the work cycle as the basic unit of observation in terms of event sequence data analysis for working condition diagnosis. In our opinion, no matter in what kind of circumstances (e.g., different places and different climates) the equipment works, the working condition of equipment in one work cycle mainly behaves similarly. For example, in a work cycle of the concrete pump truck illustrated in Figure 1, different concrete pump trucks usually have a similar working process: *warm-up* $\rightarrow$ *landing leg unfolding* $\rightarrow$ *cantilever unfolding* $\rightarrow$ *concrete pumping* $\rightarrow$ *landing leg folding* $\rightarrow$ *cantilever folding*, even though the concrete pump truck works in different working circumstances.

Definition 2 (event). An *event* of the equipment, denoted by e, is to record an occurrence of a specific message indicating that something such as an operation has happened in the equipment.

For example, in Figure 1, there are six events (*warm-up, landing leg unfolding, cantilever unfolding, concrete pumping, landing leg folding,* and *cantilever folding*), which reflect the working condition of some component in the concrete pump truck, respectively. We will use integers to denote the entries in the event set, with each event e taking a value from $1, \ldots, E$, where E is the number of unique events in the event set denoted by $\mathscr{E}$.

Definition 3 (event sequence). An *event sequence*, denoted by $\mathbf{e}_s$, of the equipment consists of a sequence of events that occur in work cycle s.

An event sequence is represented as a vector of events, $\mathbf{e}_s$, with N_s entries. For example, in Figure 1, the event set of the concrete pump truck contains six ($E = 6$) events, denoted by $\mathscr{E}_{\text{pump}} = (1, \ldots, 6)$, where the integers represent the entry of the events *warm-up, landing leg unfolding, cantilever unfolding, concrete pumping, landing leg folding,* and *cantilever folding*, respectively. Hence, the event sequence is equal to a vector with the length $N_s = 10$, denoted by $\mathbf{e}_s = (1, 2, 3, 4, 4, 4, 4, 4, 5, 6)$.

Suppose that the data set has S work cycles of the equipment, corresponding to S event sequences. The data set with S event sequences is represented as a concatenation of the event sequence vectors, which we will denote by $\mathbf{e}$, having $N = \sum_{s=1}^{S} N_s$.

In a work cycle, an event sequence provides us a main working process of the equipment. However, an occurrence of the event is also related with the working place and working date of the equipment. For example, the concrete pump truck will add an operation event *concrete mixing* in order to prevent the concrete setting if the working temperature is low. The working temperature is directly related with the working place (e.g., north or south of China) and working date (e.g., winter or summer).

In addition to these events, we have the information about the characteristics of each event sequence (work cycle): working place, working date, and equipment pieces number of the work cycle. We define $\mathbf{p}_s$ to be the set of working places of work cycle s. $\mathbf{p}_s$ consists of elements that are integers from $1, \ldots, P$, where P is the number of working places which generated the event sequences in the data set. P_s will be used to denote the number of working places of work cycle s. We define τ to be the set of working dates of work cycle s. τ_s consists of elements that are integers from $1, \ldots, T$, where T is the number of working dates. (In order to ease the notation, in the working date, we just record the working month of the work cycle, which means $T = 12$.) T_s will be used to denote the number of working dates of work cycle s. We define ω_s to be the set of equipment number of work cycle s. ω_s consists of elements that are integers from $1, \ldots, \Omega$, where Ω is the number of the equipment pieces.

Definition 4 (work cycle characteristic). A *work cycle characteristic* (WCC) is five-tuple set, denoted by $\mathscr{W}_s = \{\mathscr{E}, \mathbf{e}_s, \mathbf{p}_s, \tau_s, \omega_s\}$, which record all the information about the work cycle s.

A WCC is corresponding to a work cycle, so the original data set is redefined as a group of WCCs, denoted by $\mathscr{D} = \{\mathscr{W}_1, \ldots, \mathscr{W}_S\}$. The WCCs of two work cycles are likely to be different, though they have the same working place and working date. The main differences between the work cycles center on the occurrence of the events. However, the occurrence disciplines of the events are akin to each other for the work cycles in the same working mode. For example, the concrete pump truck has two main working modes: pumping mode and traveling model. For the work cycle in the pumping mode of the concrete pump truck, the occurrence of the event *concrete pumping* is frequent, as shown in Figure 1. However, for the work cycle in the traveling mode, the occurrence of the event *concrete pumping* is none, since the concrete pump truck can not pump concrete in the traveling mode.

Definition 5 (working mode). A *working mode*, denoted by π, is on behalf of a kind of work cycles that is about a specific subject, has an identifiable purpose, and can stand alone.

For event set $\mathscr{E}$, we define working mode vector (WMV) $\mathscr{G}(\pi) = \{(1, c_1), \ldots, (E, c_E)\}$ to be the set of events e associated with its occurrence frequency c_e, where $\sum_{e=1}^{E} c_e = 1$. The WMV $\mathscr{G}(\pi)$ is able to depict the occurrence disciplines of events according to the occurrence frequency of events. Therefore, if we can get a group of WMVs for a group of work cycles, it will help us better understand the occurrence disciplines of events.

Definition 6 (working mode space (WMS)). A *working mode space* (WMS), denoted by $\mathbb{G} = \{\mathscr{G}(1), \ldots, \mathscr{G}(\Pi)\}$, is a set of WMVs for a group of given work cycles of equipment.

Actually, the WMS is akin to a group of cluster centers, each of which depicts the working condition of equipment in different working modes.

4. The Inference of WMS

In this section, we develop effective algorithms for the inference of the WMS for a group of given work cycles of equipment. Before proceeding, we formulate our problem as follows.

WMS Inference Problem. Given a group of work cycles associated with the corresponding WCCs $\mathscr{D} = \{\mathscr{W}_1, \ldots, \mathscr{W}_S\}$, the inference problem is to infer the WMS model $\mathbb{G} = \{\mathscr{G}(1), \ldots, \mathscr{G}(\Pi)\}$, where Π represents the number of working modes.

With the help of WMS, we can find that, in different working places and different working dates, the work cycles of equipment have different working modes. Meanwhile, there are several working modes in the same working place and the same working date. The WMV of working mode reflects

the working condition of its corresponding work cycle, especially the occurrence disciplines of events.

In the remainder of this section, we first introduce the WCM for learning the WMS for a group of given work cycles, and then introduce the inference framework of the WCM.

4.1. The WCM. The WCM is a hierarchical generative model in which each event e in a work cycle is associated with three latent variables: a working place, **x**, a working date, **y**, and a working mode, **z**. These latent variables augment the E-dimensional vector **e** (indicating the values of all events in the event set $\mathscr{E}$) with three additional E-dimensional vectors, **x**, **y**, and **z**, indicating working place, working date, and working mode assignments for the E events.

As we observed, the sets of working places and the sets of working dates for each work cycle are observed. This leaves the unresolved issue of having unobserved working places and working dates and avoids the need to define a prior on working places and working dates, which is outside of the scope of our model. Each working place is associated with a multinomial distribution over working mode, and each working date is also associated with a multinomial distribution over working mode. Conditioned on the set of working places and the set of working dates, associated with their distributions over working modes, the process by which the corresponding event sequence for a work cycle is simulated can be summarized as follows: first, a working place and a working date are, respectively, chosen uniformly at random for each event that will appear in the work cycle; next, a working mode is sampled for each event both from the distribution over working mode associated with the working place of that event and from the distribution over working mode associated with the working date of that event; finally, the events themselves are sampled from the distribution over events associated with each working mode.

This simulating process can be expressed more formally by defining some of the other variables in the WCM. Assume we have Π working modes. We can parameterize the multinomial distribution over working modes for each working place using matrix Θ of size $\Pi \times P$, with elements $\theta_{\pi p}$ that stand for the probability of assigning working mode π to an event occurring in working place p. Thus $\sum_{\pi=1}^{\Pi} \theta_{\pi p} = 1$, and for simplicity of notation we will drop the index π when convenient and use θ_p to stand for the pth column of the matrix Θ. Similarly, we use matrix Δ of size $\Pi \times T$ to parameterize the multinomial distribution over working modes for each working date, where elements $\delta_{\pi\tau}$ stand for the probability of assigning working mode π to an event occurring in the working date τ. Thus, $\sum_{\tau=1}^{T} \delta_{\pi\tau} = 1$, and we will also drop the index π when convenient and use δ_τ to stand for the τth column of the matrix Δ, intended to simplify the notation. The multinomial distributions over events associated with each working mode are parameterized by matrix Φ of size $E \times \Pi$, with elements $\phi_{e\pi}$ that stand for the probability of simulating to make event e occur in the working mode π. Again, $\sum_{e=1}^{E} \phi_{e\pi} = 1$, and ϕ_e stands for the eth column of the matrix Φ. These three multinomial distributions are assumed to be generated from symmetric Dirichlet priors with hyperparameters α, β, and γ, respectively. In the results of this paper, we assume that these hyperparameters are fixed. This notation is summarized in Notations.

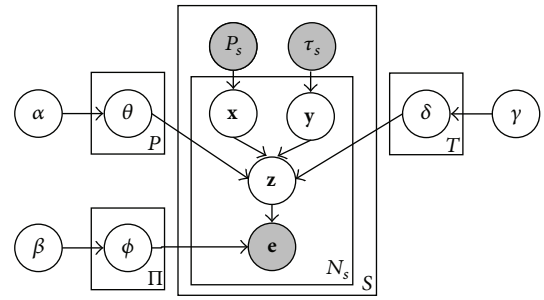

Figure 3: The graphic representation of WCM.

The sequential simulating procedure of first picking a working place and a working date, respectively, followed by picking a working mode, and then simulating an event to occur in this working mode according to the probability distributions, leads to the following generative process:

(1) For each working place $p = 1, \ldots, P$ choose $\theta_p \sim$ Dirichlet(α);

for each working date $\tau = 1, \ldots, T$ choose $\delta_\tau \sim$ Dirichlet(γ);

for each working mode $\pi = 1, \ldots, \Pi$ choose $\phi_\pi \sim$ Dirichlet(β).

(2) For each work cycle $s = 1, \ldots, S$,

given the vector of working places $\mathbf{p}_s$,
given the vector of working dates τ_s,
for each event $i = 1, \ldots, N_s$,

conditioned on $\mathbf{p}_s$ choose working place $x_{si} \sim$ Uniform($\mathbf{p}_s$),
conditioned on τ_s choose working date $y_{si} \sim$ Uniform(τ_s),
conditioned on x_{si} and y_{si} choose working mode $z_{si} \sim$ Discrete($\theta_{x_{si}}, \delta_{y_{si}}$),
conditioned on z_{si} choose event $e_{si} \sim$ Discrete($\phi_{z_{si}}$).

The graphical model corresponding to this process is shown in Figure 3. Under this simulating process, the working mode is drawn independently when conditioned on Φ, and each working mode is drawn independently when conditioned on Θ, Δ, and Π. The probability of the event sequence **e**, conditioned on Θ, Δ, and Φ (and implicitly on a fixed number of working modes Π), is

$$P(\mathbf{e} \mid \Phi, \Delta, \Theta, \mathscr{P}, \mathscr{T}) = \sum_{s=1}^{S} P(\mathbf{e}_s \mid \Phi, \Delta, \Theta, \mathbf{p}_s, \tau_s). \quad (1)$$

With the help of (1), we can first obtain the probability of the event sequence in each work cycle, $\mathbf{e}_s$, by summing over the latent variables $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$, to get what is shown in (3). Consider

$$P(\mathbf{e}_s \mid \Phi, \Delta, \Theta, \mathscr{P}, \mathscr{T}) = \prod_{i=1}^{N_s} P(e_{si} \mid \Phi, \Delta, \Theta, \mathbf{p}_s, \tau_s) = \prod_{i=1}^{N_s} \sum_{\tau=1}^{T} \sum_{p=1}^{P} \sum_{\pi=1}^{\Pi} P(e_{si}, z_{si} = \pi, x_{si} = p, y_{si} = \tau \mid \Phi, \Delta, \Theta, \mathbf{p}_s, \tau_s)$$
$$= \prod_{i=1}^{N_s} \sum_{\tau=1}^{T} \sum_{p=1}^{P} \sum_{\pi=1}^{\Pi} P(e_{si} \mid z_{si} = \pi, \Phi) P(z_{si} = \pi \mid x_{si} = p, \Theta) P(z_{si} = \pi \mid y_{si} = \tau, \Delta) P(x_{si} = p \mid \mathbf{p}_s) P(y_{si} = \tau \mid \tau_s), \tag{2}$$

$$P(\mathbf{e}_s \mid \Phi, \Delta, \Theta, \mathscr{P}, \mathscr{T}) = \prod_{i=1}^{N_s} \frac{1}{P_s} \frac{1}{T_s} \sum_{p \in \mathbf{p}_s} \sum_{\tau \in \tau_s} \sum_{\pi=1}^{\Pi} \phi_{e_{si}\pi} \theta_{\pi p} \delta_{\pi\tau}, \tag{3}$$

$$P(\mathbf{e} \mid \alpha, \beta, \gamma, \mathscr{P}, \mathscr{T}) = \int_\Theta \int_\Delta \int_\Phi P(\mathbf{e} \mid \Theta, \Delta, \Phi, \mathscr{P}, \mathscr{T}) P(\Theta, \Delta, \Phi \mid \alpha, \gamma, \beta) \, d\Theta \, d\Delta \, d\Phi \tag{4}$$

$$= \int_\Theta \int_\Delta \int_\Phi \left[\prod_{i=1}^{N_s} \frac{1}{P_s} \frac{1}{T_s} \sum_{p \in \mathbf{p}_s} \sum_{\tau \in \tau_s} \sum_{\pi=1}^{\Pi} \phi_{e_{si}\pi} \theta_{\pi p} \delta_{\pi\tau} \right] P(\Theta, \Delta, \Phi \mid \alpha, \gamma, \beta) \, d\Theta \, d\Delta \, d\Phi. \tag{5}$$

In (3), the factorization makes use of the conditional independence assumptions of model. Meanwhile, the variables $\mathbf{x}$ and $\mathbf{y}$ are mutually stochastically independent. Equation (3) represents the probability of the events $\mathbf{e}$ in terms of the entries of the parameter matrices Θ, Φ, and Δ as introduced above. The probability distribution over working place assignments, $P(x_{si} = p \mid \mathbf{p}_s)$, is assumed to be uniform over the elements of $\mathbf{p}_s$ and deterministic if $P_s = 1$. Similarly, the probability distribution over working date assignments, $P(y_{si} = \tau \mid \tau_s)$, is assumed to be uniform over the elements of τ_s and deterministic if $T_s = 1$. The probability distribution over working mode assignments, both $P(z_{si} = \pi \mid x_{si} = p, \Theta)$ and $P(z_{si} = \pi \mid y_{si} = \tau, \Delta)$, is the multinomial distributions θ_p and δ_τ in Θ and Δ, respectively, that corresponds to working place p and working date τ, respectively. The probability of an event given a working mode assignment, $P(e_{si} \mid z_{si} = \pi, \Phi)$, is the multinomial distribution ϕ_π in Φ that corresponds to working mode π.

In (4) and (5), we treat Θ, Φ, and Δ as random variables and compute the marginal probability of a corpus by integrating them out. $P(\Theta, \Delta, \Phi \mid \alpha, \gamma, \beta) = P(\Theta \mid \alpha) P(\Delta \mid \gamma) P(\Phi \mid \beta)$ are the Dirichlet priors on Θ, Δ, and Φ, respectively, as we defined before.

5. Inference of WCM from Data

The WCM contains three continuous random variables: Θ, Δ, and Φ. Various approximate inference approaches have recently been proposed for estimating the posterior distribution for continuous random variables in hierarchical Bayesian models. In this paper our inference method is Gibbs sampling [40], which is a special form of Markov chain Monte Carlo.

Our target of estimation is to compute the posterior distribution $P(\Theta, \Delta, \Phi \mid \alpha, \gamma, \beta)$. In order to sample the values of the distribution, we have to use the latent variables $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ to estimate the posterior distribution:

$$P(\Theta, \Delta, \Phi \mid \alpha, \gamma, \beta) = \sum_{\mathbf{x},\mathbf{y},\mathbf{z}} P(\Theta, \Delta, \Phi \mid \mathbf{x}, \mathbf{y}, \mathbf{z}, \alpha, \gamma, \beta) P(\mathbf{x}, \mathbf{y}, \mathbf{z} \mid \alpha, \gamma, \beta). \tag{6}$$

The estimation process mainly involves two steps: first, we use Gibbs sampling to get approximate posterior $P(\mathbf{x}, \mathbf{y}, \mathbf{z} \mid \alpha, \gamma, \beta)$; second, $P(\Theta, \Delta, \Phi \mid \mathbf{x}, \mathbf{y}, \mathbf{z}, \alpha, \gamma, \beta)$ can be computed directly for each sample, by exploiting the fact that the Dirichlet distribution is conjugate to the multinomial.

5.1. Gibbs Sampling. Using Gibbs sampling we can generate a sample from the joint distribution $P(\mathbf{z}, \mathbf{y}, \mathbf{z} \mid D_{\text{train}}, \alpha, \beta)$ by two steps: first, sampling working place assignment x_{si}, working date assignment y_{si}, and working mode assignment z_{si} for individual event e_{si}, conditioned on fixed assignments of working places, working date,s and working modes for all other events in the data set; second, repeating this process for each event. A single Gibbs sampling iteration consists of sequentially performing this sampling of working place, working date, and working mode assignments for each individual event in the data set:

$$P(x_{si} = p, y_{si} = \tau, z_{si} = \pi \mid e_{si} = e, \mathbf{x}_{-si}, \mathbf{y}_{-si}, \mathbf{z}_{-si}, \mathbf{e}_{-si}, \mathscr{P}, \mathscr{T}, \alpha, \beta)$$
$$\propto \frac{C^{E\Pi}_{e\pi,-si} + \beta}{\sum_{e'} C^{E\Pi}_{e'\pi,-si} + E\beta} \frac{C^{\Pi P}_{\pi p,-si} + \beta}{\sum_{p'} C^{\Pi P}_{\pi p',-si} + P\alpha} \cdot \frac{C^{\Pi T}_{\pi\tau,-si} + \beta}{\sum_{\tau'} C^{\Pi T}_{\pi\tau',-si} + T\gamma}. \tag{7}$$

According to (1)~(5), we can derive a basic equation needed for the Gibbs sampler as shown in (7). In (7),

$C^{\Pi P}$ means working mode assigned to working place count matrix, where $C^{\Pi P}_{\pi p,-si}$ means the number of events assigned to working mode π in the working place p excluding the working mode assignment to event e_{si}. Similarly $C^{\Pi T}$ means working mode assigned to working date count matrix, where $C^{\Pi T}_{\pi\tau,-si}$ means the number of events assigned to working mode π in the working date τ excluding the working mode assignment to event e_{si}. Similarly, $C^{E\Pi}$ represents event assigned to working mode count matrix, where $C^{E\Pi}_{e\pi,-si}$ represents the number of events from the eth entry in the event set assigned to working mode π excluding the topic assignment to event e_{si}. Meanwhile, $\mathbf{x}_{-si}, \mathbf{y}_{-si}, \mathbf{z}_{-si}, \mathbf{e}_{-si}$ represents the vector of working place assignment, vector of working date assignment, vector of working mode assignments, and vector of event observations in the data set except for the ith event in the sth work cycle, respectively.

The main sampling steps are as follows: we first initialize the working place, working date, and working mode assignments, **x**, **y**, and **z**, randomly. In each Gibbs sampling iteration, we sequentially draw the working mode, working place, and working date assignment of the ith event from the joint conditional distribution in (7). With the increasing of iterations, the Gibbs sampler will approach its stationary distribution—the posterior distribution $P(\mathbf{z}, \mathbf{y}, \mathbf{z} \mid D_{\text{train}}, \alpha, \beta)$.

5.2. The Posterior Probability. Given $\mathbf{z}, \mathbf{y}, \mathbf{z}, D_{\text{train}}, \alpha, \beta$, and γ, computing posterior distributions on Θ, Δ, and Φ is straightforward. Based on the fact that the Dirichlet distribution is conjugate to the multinomial distribution, then we can get

$$\begin{aligned}
\phi_\pi \mid \mathbf{z}, \beta, D_{\text{train}} &\sim \text{Dilichlet}\left(C^{E\Pi}_{.\pi} + \beta\right),\\
\theta_p \mid \mathbf{x}, \mathbf{z}, \alpha, D_{\text{train}} &\sim \text{Dilichlet}\left(C^{\Pi P}_{.p} + \alpha\right),\\
\delta_\tau \mid \mathbf{y}, \mathbf{z}, \gamma, D_{\text{train}} &\sim \text{Dilichlet}\left(C^{\Pi T}_{.\tau} + \gamma\right),
\end{aligned} \tag{8}$$

where $C^{E\Pi}_{.\pi}$ represents the vector of counts of the number of times each event has been assigned to working mode π. $C^{\Pi P}_{.p}$ and $C^{\Pi T}_{.\tau}$ are similar to $C^{E\Pi}_{.\pi}$. Then we can evaluate the posterior probability of each element of Θ, Δ, and Φ as follows:

$$\begin{aligned}
E\left[\phi_\pi \mid \mathbf{z}, \beta, D_{\text{train}}\right] &= \frac{\left(C^{E\Pi}\right)^k + \beta}{\sum_{e'}\left(C^{E\Pi}_{e'\pi}\right)^k + E\beta},\\
E\left[\theta_p \mid \mathbf{x}, \mathbf{z}, \alpha, D_{\text{train}}\right] &= \frac{\left(C^{\Pi P}\right)^k + \alpha}{\sum_{\tau'}\left(C^{\Pi P}_{\pi' p}\right)^k + P\alpha},\\
E\left[\delta_\tau \mid \mathbf{y}, \mathbf{z}, \gamma, D_{\text{train}}\right] &= \frac{\left(C^{\Pi T}\right)^k + \gamma}{\sum\left(C^{\Pi T}_{\pi' t}\right)^k + T\gamma},
\end{aligned} \tag{9}$$

where $(C^{E\Pi})^k$ is the matrix of working mode assigned to event counts exhibited in $(\mathbf{z})^k$ and k refers to sample k from the Gibbs sampler. These posterior probabilities also provide point estimates for Φ, Θ, and Δ and correspond to the posterior predictive distribution for the next event from a working mode, the next event from a working date, and the next working mode in a work cycle, respectively.

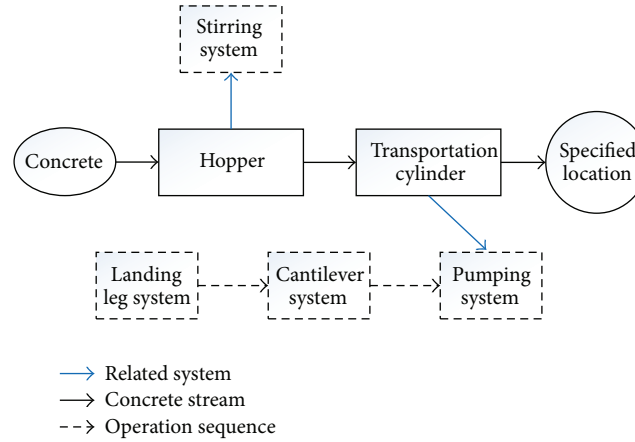

Figure 4: The stream of the concrete in the concrete pump truck and the operation sequence of the concrete pump truck at runtime.

6. Experimental Evaluation

6.1. Data Preparation. We trained the WCM on a real world data set collected from a well-known Chinese construction machinery manufacturer. The data set is a set of event sequence data from the concrete pump truck in 6 months (from June 2012 to November 2012). This data set contains $S = 32,632$ work cycles, $P = 5$ different working places, $T = 6$ different working dates, a total of $N = 22,418,756$ event tokens, and an event set size of $E = 33$ unique events. The working date of each work cycle is according to its real working month, which means the working date set $\mathcal{T} = \{\text{Jun}, \text{Jul}, \text{Aug}, \text{Sep}, \text{Oct}, \text{Nov}\}$. Because the event sequence data are all collected in the Chinese Mainland, we divide the working places into 5 regions according to administrative region of China: Northern China, Northeastern China, Eastern China, Mid-Southern China, and Western China.

The concrete pump truck is a type of construction machinery, which is a truck associated with a concrete pump. It alternates between two working statuses: traveling and pumping. In the pumping status, it will push the concrete to the specified location. In the traveling status, it is just a truck. In the experiment, we mainly focus on events in the pumping status. Figure 4 shows the stream of the concrete in the concrete pump truck at runtime and operation sequence of different systems in the concrete pump truck. The concrete pump truck first switches to pumping status and then unfolds and fixes the landing leg. Next, it unfolds cantilever to the specified location. Afterwards, the concrete is poured to the hopper, and meanwhile the stirring system initiates stirring the concrete. Finally, the pumping system initiates pumping the concrete in the hopper to the specified location. When the pumping ends, the concrete pump truck stops the pumping system and then folds the cantilever and landing leg

Table 1: Event set.

Event	Abbr.	Type	Related system
Stop pumping mandatorily	*SPM*	Alarm event	All
Reminder of concrete import	*RCI*	Alarm event	Hopper
Concrete piston withdrawing	*CPW*	Alarm event	Pumping system
Reminder of concrete cylinder water	*RCSW*	Alarm event	Hopper
Swing cylinder initiate	*SCI*	Operation event	Pumping system
Stalling of engine	*SoE*	Alarm event	All
Alteration of operation mode (*remote or close*)	*AOM*	Operation event	Pumping system
Alteration of pump truck status (*pumping or travelling*)	*APTS*	Operation event	All
Control of pumping displacement	*CPD*	Operation event	Pumping system
Transportation cylinder initiate	*TCI*	Operation event	Pumping system
Manual control of master cylinder	*MCMC*	Operation event	Pumping system
Manual control of swing cylinder	*MCSC*	Operation event	Pumping system
Detection of system pressure	*DSP*	Alarm event	Pumping system
Manual control of engine speed	*MCES*	Operation event	Pumping system
High pressure mode initiate	*HPMI*	Operation event	Pumping system
Warm-up initiate	*WUI*	Operation event	Pumping system
Water pump initiate	*WPI*	Operation event	Hopper
Concrete stirring initiate	*CSI*	Operation event	Stirring system
Cantilever folding initiate	*CFI*	Operation event	Cantilever system
Temperature control initiate	*TCI*	Operation event	Pumping system
Cantilever movement	*CM*	Operation event	Cantilever system
Landing leg movement	*LLM*	Operation event	Landing leg system
Detection of oil pressure	*DOP*	Alarm event	Pumping system
Landing leg folding	*LLF*	Operation event	Landing leg system
Rotary table movement	*RTM*	Operation event	Cantilever system
Oil pump initiate	*OPI*	Operation event	Pumping system
Energy accumulator initiate	*EAI*	Operation event	Pumping system
Bypath valve initiate	*BVI*	Operation event	Pumping system
Concrete pumping initiate	*CPI*	Operation event	Pumping system
Master cylinder initiate	*MCI*	Operation event	Pumping system
Cantilever shock absorbers initiate	*CSAI*	Alarm event	Cantilever system
Initiate of system cooling	*ISC*	Operation event	Pumping system
Hydraulic oil supplement	*HOS*	Operation event	Pumping system

successively. Table 1 shows the relations between systems and events in the concrete pump truck.

Table 1 shows all the events in the event set. There are two types of events: alert event and operation event. The occurrence of an alarm event is to remind the operator that some emergency happens. For example, the occurrence of event RCI means to remind the operator to import concrete into the hopper. The alarm event is not a regular operation. The operation event is the real record of regular operations in the concrete pump truck.

6.2. Analysis for Gibbs Sampling Using Perplexity. As mentioned earlier, in the experiment described in this paper we do not estimate the hyperparameters α, β, and γ. Instead they are fixed at $50/\Pi$, 0.01, and $50/\Pi$, respectively. In this paper we use the perplexity of the model on test work cycles to evaluate when the performance of the model begins to stabilize.

The perplexity of new unobserved work cycle s that contains events $\mathbf{e}_s$ and is conditioned on the working places $\mathbf{p}_s$ and working dates τ_s of the work cycle is defined as

$$\text{Perplexity}\left(\mathbf{e}_s \mid \mathbf{p}_s, \tau_s\right) = \exp\left(-\frac{\log P\left(\mathbf{e}_s \mid \mathbf{p}_s, \tau_s\right)}{N_s}\right), \quad (10)$$

where $P(\mathbf{e}_s \mid \mathbf{p}_s, \tau_s)$ is the probability assigned by the WCM. To simplify notation here, we do not consider the explicit dependency on the hyperparameters. For multiple work cycles, we report the average perplexity over work cycles defined as follows:

$$\overline{\text{Perplexity}} = \sum_{s=1}^{S} \frac{\text{Perplexity}\left(\mathbf{e}_s \mid \mathbf{p}_s, \tau_s\right)}{S}. \quad (11)$$

The lower the perplexity the better the performance of the model. We can obtain an approximate estimate of perplexity

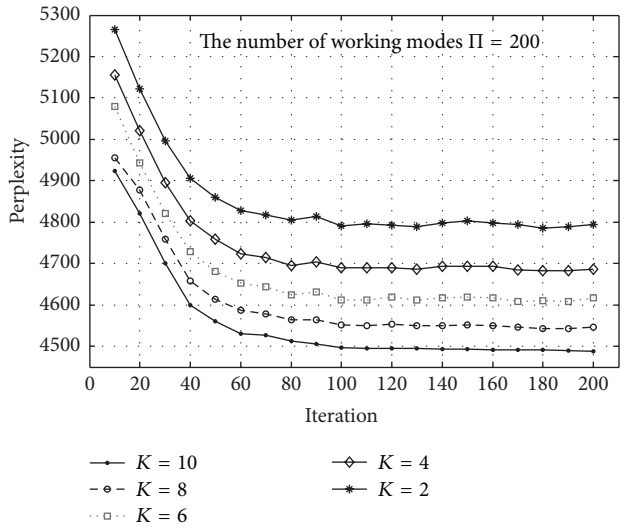

Figure 5: Perplexity as a function of iterations of the Gibbs sampler for a $\Pi = 200$ model, respectively. Each curve shows the perplexity from averaging for different settings of Π, but now over a larger range of sampling iterations.

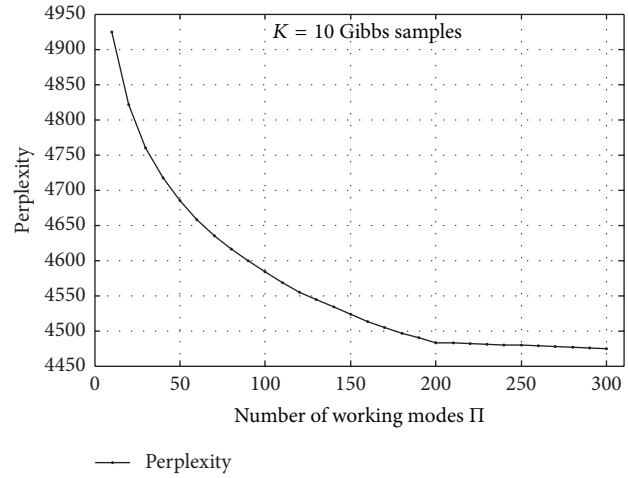

Figure 6: Perplexity as a function of the parameter Π of the Gibbs sampler for $K = 10$ samples.

by averaging over multiple samples according to (9), calculated as follows:

$$P(\mathbf{e}_s \mid \mathbf{p}_s, \tau_s) = \frac{1}{K}\sum_{k=1}^{K}\prod_{i=1}^{N_s}\frac{1}{P_s T_s}\sum_{p\in\mathbf{p}_s,\tau\in\tau_s,\pi} E\left[\theta_{\pi p}\delta_{\pi\tau}\phi_{e_{si}\pi} \mid \mathbf{x}^k, \mathbf{y}^k, \mathbf{z}^k\right]. \quad (12)$$

Experimental results, using different values for K, indicated that $K = 10$ samples is a reasonable choice to get a good approximation of the perplexity. Because of the exchangeability of the working modes, it is possible that quite different solutions of working modes are detected across different samples. In practice, however, we have also found that the solutions of working modes are relatively stable across samples, with only a small subset of unique working modes appearing in any sample. Hence, we use the average perplexity values across samples in the experiment.

Figure 5 illustrates the perplexity as a function of iterations of the Gibbs sampler, for a $\Pi = 200$ model to fit the data set, respectively. It appears from Figure 5 that performance of models (for different settings of parameter K) trained using the Gibbs sampler appears to stabilize rather quickly (after about 100 iterations), at least in terms of perplexity on the data set. This indicates that the perplexity values flatten out after a 100 or so iterations of the Gibbs sampler.

6.3. The Number of Working Modes Π. Although the perplexity computation is able to be averaged over different Gibbs sampler runs, other applications of the model rely on the analysis of each working mode and are based on the analysis of each sample. Meanwhile, the setting of the parameter Π is also determined according to the perplexity. The parameter Π represents the number of working modes.

Figure 6 illustrates the perplexity as a function of the parameter Π in $K = 10$ Gibbs samples. Empirical settings of the parameter Π show that the average perplexity over the data set decreases with the increase of the parameter Π. Experimental results confirm that the average perplexity indeed decreases as we made analysis. In particular, the perplexity values flatten out after the parameter Π is set to 200. This indicates that the parameter $\Pi = 200$ fits the data set in the model.

6.4. Analysis of the WCM Results. About the analysis of the WCM results, we can use the point estimate of the WCM parameters to look at specific Θ, Δ, and Φ distributions and related quantities that can be derived from these parameters (such as the probability of a working place and a working date given a randomly selected event from a working mode). In the following results, we take a specific sample, $\mathbf{x}_k$, $\mathbf{y}_k$, and $\mathbf{z}_k$, after 100 iterations from a single arbitrarily selected Gibbs run and then generate point estimates of Θ, Δ, and Φ using (9).

There are totally 200 working modes (parameter $\Pi = 200$). Each working mode, using a WMV, helps us to better understand the occurrences of events. For the sake of analysis, we list the highest probability working modes for each working place and each working date from the WCM in Table 2. In each working mode, we list the top 10 events most likely to be generated in the most likely working mode conditioned on both the working place and working date. For example, in the working place of Northern China, for the most likely working mode (numbered 101 in the 200 working modes), the top 10 events (OPI, SPM, EAI, HOS, BVI, MCI, AOM, APTS, CPD, and TCI) are most likely to occur in the working date of June.

Experimental results show that different working places have different working modes in spite of the same working date, and the same working place also has different working modes for different working dates. It indicates that the working mode is indeed related with the working place and working date. Events related with the pumping system, such

Table 2: The highest probability working mode for each working place and each working date from the WCM.

Working date	Probability	Working mode	Events
Working place = Northern China			
Jun.	0.0251	101	*OPI, CM, EAI, RTM, BVI, SPM, AOM, APTS, CPD,* and *TCI*
Jul.	0.0341	164	*OPI, CSI, RTM, CM, ISC, APTS, SPM, AOM, CPD,* and *TCI*
Aug.	0.0051	62	*LLF, CFI, APTS, CSI, AOM, ISC, RTM, SPM, CPD,* and *TCI*
Sep.	0.0342	12	*OPI, RTM, CM, CPI, BVI, LLF, SPM, AOM, APTS,* and *CPD*
Oct.	0.0351	49	*RTM, OPI, CM, BVI, MCI, SPM, AOM, APTS, CPD,* and *TCI*
Nov.	0.0353	129	*OPI, ISC, SPM, EAI, CSI, APTS, AOM, CPD, TCI,* and *MCMC*
Working place = Northeastern China			
Jun.	0.0258	176	*OPI, SPM, EAI, HOS, BVI, MCI, AOM, APTS, CPD,* and *TCI*
Jul.	0.0263	29	*OPI, LLF, ISC, SPM, CFI, APTS, CPI, HOS, AOM,* and *CPD*
Aug.	0.0141	71	*OPI, CSI, RTM, CM, ISC, APTS, SPM, AOM, CPD,* and *TCI*
Sep.	0.0114	111	*RTM, BVI, OPI, CM, MCI, HOS, EAI, SPM, AOM,* and *APTS*
Oct.	0.0146	69	*ISC, LLF, CSI, AOM, APTS, OPI, CFI, SPM, CPD,* and *TCI*
Nov.	0.0257	93	*RTM, OPI, BVI, MCI, CM, CPI, SPM, AOM, APTS,* and*CPD*
Working place = Eastern China			
Jun.	0.0279	177	*OPI, HOS, CPI, SPM, LLF, RTM, EAI, BVI, AOM,* and *APTS*
Jul.	0.0201	72	*OPI, EAI, CPI, SPM, MCI, RTM, HOS, AOM, APTS,* and *CPD*
Aug.	0.0277	87	*OPI, BVI, EAI, RTM, AOM, SPM, MCI, APTS, CPD,* and *TCI*
Sep.	0.0274	9	*OPI, EAI, BVI, RTM, HOS, SPM, AOM, APTS, CPD,* and *TCI*
Oct.	0.0214	191	*RTM, MCI, CPI, CM, EAI, OPI, HOS, SPM, AOM,* and *APTS*
Nov.	0.0255	170	*OPI, MCI, BVI, RTM, CPI, HOS, SPM, AOM, APTS,* and *CPD*
Working place = Mid-Southern China			
Jun.	0.0122	74	*OPI, EAI, CSI, CPI, ISC, MCI, SPM, AOM, APTS,* and *CPD*
Jul.	0.0177	33	*OPI, CPI, CM, MCI, HOS, SPM, AOM, APTS, CPD,* and *TCI*
Aug.	0.0262	187	*HOS, MCI, CPI, OPI, EAI, BVI, CSI, SPM, AOM,* and *APTS*
Sep.	0.0205	104	*RTM, EAI, BVI, OPI, SPM, MCI, CFI, APTS, AOM,* and *CPD*
Oct.	0.0193	39	*OPI, HOS, BVI, CM, RTM, SPM, AOM, APTS, CPD,* and *TCI*
Nov.	0.0133	158	*OPI, BVI, RTM, MCI, CM, SPM, AOM, APTS, CPD,* and *TCI*
Working place = Western China			
Jun.	0.0037	4	*OPI, RTM, BVI, CM, EAI, SPM, CPI, MCI, AOM,* and *APTS*
Jul.	0.0134	144	*HOS, MCI, CPI, OPI, CFI, EAI, SPM, AOM, APTS,* and *CPD*
Aug.	0.0126	126	*OPI, SPM, CM, BVI, AOM, LLF, APTS, CSI, CPD,* and *TCI*
Sep.	0.0122	88	*OPI, HOS, CPI, CM, LLF, AOM, CFI, MCI, BVI,* and *SPM*
Oct.	0.0104	37	*OPI, EAI, MCI, HOS, CSI, ISC, CFI, LLF, SPM,* and *AOM*
Nov.	0.0135	78	*OPI, HOS, RTM, BVI, CSI, EAI, MCI, APTS, AOM,* and *SPM*

as OPI, MCI, and CPI, are most likely to occur in most working modes, which indicates that the working modes of the concrete pump truck are consistent with the actual situations. Meanwhile events related with the cantilever system and landing leg system, such as LLF and CFI, have less occurrences as compared with events of the pumping system. Moreover, in the working date of summer (working date = June, July, and August), the alert event SPM is more likely to occur, which indicates that the concrete pump truck more likely fails in the hot climate. The operation event AOM is more likely to occur, which indicates that the operators prefer to operate the concrete pump truck in the remote manner.

Because the probability of working mode reflects the probability of its occurrence, we can analyze the work loads of different working places in different working dates. According to the probability of the working mode in Table 2, we can find that the working modes in the working place of Eastern China are more likely to occur than the working modes in the working place of Western China. It indicates that the concrete pump trucks in the working place of Eastern China have more work loads than that in the working place of Western China. Meanwhile, the concrete pump trucks in the working date of June have more work loads than that in the working date of November. Generally, we can analyze different working modes according to the probability.

6.5. Illustrative Applications for the WCM. In this section we provide some illustrative examples of how the WCM can be used to answer different types of questions and prediction problems concerning working modes of the equipment.

6.5.1. Automated Detection for a New Work Cycle. In real cases, we would like to quickly assess working mode assignments for new work cycles not contained in the training data set, especially for the real-time event sequence flow.

Our automated detection strategy is to apply the Gibbs sampling algorithm that runs only on the event tokens in the new work cycle, instead of rerunning the algorithm for every new work cycle again. Afterwards, the event tokens in the new work cycles are quickly assigned to the most likely working places, working dates, and working modes. The main procedure is as follows: first, we start by assigning events randomly to working places, working dates, and working modes; second, we then sample new assignments of events by applying the Gibbs sampler only to the event tokens in the new work cycle, each time temporarily updating the count matrices $C^{E\Pi}$, $C^{\Pi P}$, and $C^{\Pi T}$ shown in (7).

Table 3 shows the occurrences of events for a new work cycle. After the sampling, the WCM has assigned each event to its most likely working mode. Table 3 illustrates the top 3 most likely working modes assigned to each event for the new work cycle. Note that each event is assigned to different working modes according to its occurrence count. According to (7), although events of this new work cycle are assigned to different working modes, they are assigned to the number 107 working mode with the probability 0.0003. The top 10 most likely events in the number 107 working mode are shown as follows:

> *RTM*, *CM*, *OPI*, *BVI*, *SPM*, *CPI*, *MCI*, *SCI*, *ISC*, and *SoE*

The automated detection result for the new work cycle is indeed consistent with the actual situations, in comparison with the real occurrences of events.

6.5.2. Automated Detection of Anomalous Work Cycles. We illustrate in this section how our model could be useful for detecting anomalous work cycles. A work cycle assigned to a working mode with low probability is considered as an anomalous work cycle.

We also take the work cycle as an example for the automated detection of an anomalous work cycle, shown in Table 3. The work cycle is assigned to the number 107 working mode with the probability 0.0003. As compared with most of other working modes, number 107 working mode has lower probability, so this work cycle is detected as an anomalous work cycle. The alert events SPM and SoE have frequent occurrences both in the work cycle and in number 107 working mode, which indicates that this work cycle is an anomalous work cycle. Meanwhile, we analyzed the real failure records and confirmed that the engine indeed failed frequently during this work cycle. Generally, these anomalous work cycles can be automatically detected efficiently, with the help of the WCM.

Table 3: Actual example of automated detection for a new work cycle. Each event is assigned to its most likely working mode according to its corresponding occurrence count. In the table, we list the top 3 most likely working modes for each event for the new work cycle.

Top 3 most likely working modes Working date = Jun.; working place = Eastern China				
Event	Count	First	Second	Third
SPM	72	107	181	112
AOM	33	169	67	183
APTS	23	90	15	76
CPD	42	145	139	59
TCI	2	118	134	112
MCMC	0	Null	Null	Null
MCSC	0	Null	Null	Null
DSP	0	Null	Null	Null
MCES	0	Null	Null	Null
HPMI	0	Null	Null	Null
WUI	2	159	104	77
WPI	23	54	175	71
CSI	55	147	29	61
CFI	25	2	132	100
TCI	23	95	185	53
CM	127	12	49	192
LLM	55	189	114	23
RCI	0	Null	Null	Null
DOP	0	Null	Null	Null
LLF	40	111	10	42
RTM	297	191	104	52
CPW	0	Null	Null	Null
RCSW	0	Null	Null	Null
OPI	95	177	176	101
EAI	56	126	100	170
BVI	77	177	53	146
CPI	60	164	104	149
MCI	60	177	175	162
SCI	66	120	149	73
CSAI	0	Null	Null	Null
ISC	51	68	149	23
HOS	0	Null	Null	Null
SoE	33	119	112	107

7. Conclusions and Future Work

The working condition model proposed in this paper provides a relatively simple probabilistic model for exploring the relationships between working place, working place, working mode, and events in a work cycle. This model provides significantly improved predictive power in terms of the analysis of working condition according to the event sequence data.

Our future works mainly include the optimization of the model, the model training, and the conduction experiments on different data sets. Furthermore, the further analysis of

the anomalous work cycles detected by our model is also an interesting question.

Notations Associated with the WCM, As Used in This Paper

$\mathscr{P}$: Working places of all the work cycles (set)
$\mathscr{T}$: Working dates of all the work cycles (set)
$\mathbf{p}_s$: Working places of the sth work cycle (P_s-dimensional vector)
P_s: Number of working places of the sth work cycle (Scalar)
τ_s: Working dates of the sth work cycle (T_s-dimensional vector)
T_s: Number of working dates of the sth work cycle (Scalar)
P: Number of working places (Scalar)
S: Number of work cycles (Scalar)
T: Number of working dates (Scalar)
N_s: Number of events in the sth work cycle (Scalar)
N: Number of events in all the event sequences (Scalar)
Π: Number of working modes (Scalar)
E: Number of events in the event set (Scalar)
$\mathbf{e}_s$: Event sequence vector for the sth work cycle (N_s-dimensional vector)
e_{si}: ith event in the sth work cycle (ith component of vector $\mathbf{e}_s$)
$\mathbf{x}$: Working place assignments (N-dimensional vector)
x_{si}: Working place assignment for event e_{si} (ith component of vector $\mathbf{x}_s$)
$\mathbf{y}$: Working date assignments (N-dimensional vector)
y_{si}: Working date assignment for event e_{si} (ith component of vector $\mathbf{y}_s$)
$\mathbf{z}$: Working mode assignments (N-dimensional vector)
z_{si}: Working mode assignment for event e_{si} (ith component of vector $\mathbf{z}_s$)
α, β, γ: Dirichlet prior (Scalar)
Φ: Probabilities of events given working modes ($E \times \Pi$ matrix)
ϕ_π: Probabilities of events given working mode π (E-dimensional vector)
Θ: Probabilities of working modes given working places ($\Pi \times P$ matrix)
θ_p: Probabilities of working modes given working place p (Π-dimensional vector)
Δ: Probabilities of working modes given working dates ($\Pi \times T$ matrix)
δ_τ: Probabilities of working modes given working dates τ (Π-dimensional vector).

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] J. Holler, V. Tsiatsis, C. Mulligan, S. Avesand, S. Karnouskos, and D. Boyle, *From Machine-to-Machine to the Internet of Things: Introduction to a New Age of Intelligence*, Academic Press, 2014.

[2] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Sensing as a service model for smart cities supported by Internet of Things," *Transactions on Emerging Telecommunications Technologies*, vol. 25, no. 1, pp. 81–93, 2014.

[3] R. F. Mesquita Brandão and J. A. Beleza Carvalho, "The importance of control monitoring systems in wind parks maintenance," *British Journal of Applied Science & Technology*, vol. 4, no. 10, pp. 1461–1471, 2014.

[4] C. J. Crabtree, D. Zappalá, and P. J. Tavner, "Survey of commercially available condition monitoring systems for wind turbines," Tech. Rep., Durham University, 2014.

[5] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *The Journal of Machine Learning Research*, vol. 3, no. 4-5, pp. 993–1022, 2003.

[6] S. Kandula, R. Mahajan, P. Verkaik, S. Agarwal, J. Padhye, and P. Bahl, "Detailed diagnosis in enterprise networks," in *Proceedings of the ACM SIGCOMM Conference on Data Communication (SIGCOMM '09)*, vol. 39, pp. 243–254, ACM, August 2009.

[7] J.-G. Lou, Q. Fu, Y. Wang, and J. Li, "Mining dependency in distributed systems through unstructured logs analysis," *ACM SIGOPS Operating Systems Review*, vol. 44, no. 1, pp. 91–96, 2010.

[8] C. Luo, J.-G. Lou, Q. Lin et al., "Correlating events with time series for incident diagnosis," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '14)*, pp. 1583–1592, ACM, August 2014.

[9] J. Chen and R. Kumar, "Online failure diagnosis of stochastic discrete event systems," in *Proceedings of the IEEE Conference on Computer Aided Control System Design (CACSD '13)*, pp. 194–199, IEEE, August 2013.

[10] J. Chen and R. Kumar, "Failure diagnosis of discrete-time stochastic systems subject to temporal logic correctness requirements," in *Proceedings of the 11th IEEE International Conference on Networking, Sensing and Control (ICNSC '14)*, pp. 42–47, IEEE, April 2014.

[11] *Business Process Model and Notation (BPMN) Version 2.0*, OMG Specification, Object Management Group, 2011.

[12] F. Leymann, "Bpel vs. bpmn 2.0: should you care?" in *Business Process Modeling Notation*, pp. 8–13, Springer, Berlin, Germany, 2011.

[13] C. C. Aggarwal, *Managing and Mining Sensor Data*, Springer, 2013.

[14] N. H. Gehani, H. V. Jagadish, and O. Shmueli, "Composite event specification in active databases: model and implementation," in *Proceedings of the 18th VLDB Conference Vancouver (VLDB '92)*, vol. 92, pp. 327–338, Citeseer, British Columbia, Canada, 1992.

[15] I. Davidson, S. Gilpin, and P. B. Walker, "Behavioral event data and their analysis," *Data Mining and Knowledge Discovery*, vol. 25, no. 3, pp. 635–653, 2012.

[16] J. Han and M. Kamber, *Data Mining, Southeast Asia Edition: Concepts and Techniques*, Morgan Kaufmann, 2006.

[17] H. R. Motahari-Nezhad, R. Saint-Paul, F. Casati, and B. Benatallah, "Event correlation for process discovery from web service interaction logs," *The VLDB Journal*, vol. 20, no. 3, pp. 417–444, 2011.

[18] F. Skopik and R. Fiedler, "Intrusion detection in distributed systems using fingerprinting and massive event correlation," in *GI-Jahrestagung*, pp. 2240–2254, 2013.

[19] G. A. Wilkin, P. Eugster, and K. R. Jayaram, "Decentralized fault-tolerant event correlation," *ACM Transactions on Internet Technology*, vol. 14, no. 1, article 5, 2014.

[20] H. Wei, "A correlation analysis method for network security events," in *Informatics and Management Science III*, vol. 206 of *Lecture Notes in Electrical Engineering*, pp. 269–277, Springer, London, UK, 2013.

[21] W. Van Der Aalst, A. Adriansyah, A. K. A. de Medeiros et al., "Process mining manifesto," in *Usiness Process Management Workshops*, pp. 169–194, Springer, Berlin, Germany, 2012.

[22] J. C. A. M. Buijs, B. F. van Dongen, and W. M. P. van der Aalst, "Mining configurable process models from collections of event logs," in *Business Process Management*, pp. 33–48, Springer, 2013.

[23] Á. Rebuge and D. R. Ferreira, "Business process analysis in healthcare environments: a methodology based on process mining," *Information Systems*, vol. 37, no. 2, pp. 99–116, 2012.

[24] J. Wang, R. K. Wong, J. Ding, Q. Guo, and L. Wen, "On recommendation of process mining algorithms," in *Proceedings of the IEEE 19th International Conference on Web Services (ICWS '12)*, pp. 311–318, IEEE, Honolulu, Hawaii, USA, June 2012.

[25] R. S. Mans, W. M. P. van der Aalst, and H. M. W. Verbeek, "Supporting process mining workflows with rapidprom," in *Proceedings of the Business Process Management Demo Sessions (BPMD '14)*, vol. 1295, pp. 56–60, Eindhoven, The Netherlands, September 2014.

[26] C. Li, M. Reichert, and A. Wombacher, "Mining business process variants: challenges, scenarios, algorithms," *Data & Knowledge Engineering*, vol. 70, no. 5, pp. 409–434, 2011.

[27] R. Accorsi, T. Stocker, and G. Müller, "On the exploitation of process mining for security audits: the process discovery case," in *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, pp. 1462–1468, ACM, March 2013.

[28] B.-J. Lee, S.-G. Park, K.-B. Min et al., "The relationship between working condition factors and well-being," *Annals of Occupational and Environmental Medicine*, vol. 26, no. 1, article 34, 2014.

[29] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, Routledge Academic, New York, NY, USA, 2013.

[30] P. Bahl, R. Chandra, A. Greenberg, S. Kandula, D. A. Maltz, and M. Zhang, "Towards highly reliable enterprise network services via inference of multi-level dependencies," *ACM SIGCOMM Computer Communication Review*, vol. 37, no. 4, pp. 13–24, 2007.

[31] B. Rosner, *Fundamentals of Biostatistics*, Cengage Learning, 2010.

[32] A. Zimmermann, "Colored petri nets," in *Stochastic Discrete Event Systems: Modeling, Evaluation, Applications*, pp. 99–124, Springer, 2008.

[33] A. Adriansyah, B. F. van Dongen, and W. M. P. van der Aalst, "Towards robust conformance checking," in *Business Process Management Workshops*, vol. 66 of *Lecture Notes in Business Information Processing*, pp. 122–133, Springer, Berlin, Germany, 2011.

[34] M. Weidlich and M. Weske, *Business Process Modeling Notation*, Springer, Berlin, Germany, 2010.

[35] C. M. Bishop and J. Lasserre, "Generative or discriminative? Getting the best of both worlds," in *Bayesian Statistics*, J. M. Bernardo, M. J. Bayarri, J. O. Berger et al., Eds., vol. 8, pp. 3–23, Oxford University, 2007.

[36] C. M. Bishop, *Pattern Recognition and Machine Learning. Volume 1*, Springer, New York, NY, USA, 2006.

[37] D. M. Blei and J. D. Lafferty, "Dynamic topic models," in *Proceedings of the 23rd International Conference on Machine Learning (ICML '06)*, pp. 113–120, ACM, June 2006.

[38] J. Foulds, L. Boyles, C. DuBois, P. Smyth, and M. Welling, "Stochastic collapsed variational Bayesian inference for latent dirichlet allocation," in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 446–454, ACM, 2013.

[39] J. Pearl, *Bayesian Networks*, Department of Statistics, UCLA, 2011.

[40] I. Porteous, D. Newman, A. Ihler, A. Asuncion, P. Smyth, and M. Welling, "Fast collapsed gibbs sampling for latent dirichlet allocation," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '08)*, pp. 569–577, ACM, August 2008.