Available online at www.sciencedirect.com
ScienceDirect
IFAC-PapersOnLine 48-17 (2015) 028–032

Improvement for Large scale database-based Online Modelling of Melting Furnace

Chunlin Yu *, Harutoshi Ogai*
Norihiro Tanaka**, Hidehiko Furuya**, Yuhei Akaike**

**Graduate School of Information, Production and Systems, Waseda University, Fukuoka, Japan (Tel: 81-93-692-5147;e-mail: ogai@waseda.jp)*

***NIPPN STEEL & SUMIKIN ENGNEERING CO.,LTD. 1-5-1 Osaki, Shinagawa, Tokyo, Japan (Tel: 81-3-6665-2222; e-mail: tanaka.norihiro.m2y@eng.nssmc.com).*

Abstract: Recently, attention has been drawn by the local modeling techniques of an idea called "Just-In-Time (JIT) modeling". To apply "JIT modeling" to a large amount of database online, "Large-scale database-based Online modelling (LOM)" has been proposed. LOM is a technique that makes the retrieval of neighboring data more efficient by using stepwise method and quantization. This paper reports the improvement of LOM, a computerized sequential prediction method which uses a local regression model by PLS, on the shaft-furnace direct melting system (melting furnace). The gas temperature prediction of the furnace is done without delay.

Keywords: LOM, JIT modeling, PLS, sequential prediction, top gas temperature prediction

1. INTRODUCTION

With the development of the computing machine and the database system recently, accumulating and retrieving a large amount of data at high speed has become possible. As a result, local modelling technique of a new idea that is called "Just-In-Time (JIT) modelling", "Lazy Learning" and "Model on Demand" have gotten attention. This method accumulates the measured data in the wide operating range to a data-base. Whenever there is a need for prediction in system, the data which is highly relevant with "demand point (Query)" of the input is retrieved from the database as "neighbour" data. And an output of retrieved data is interpolated to obtain an output of "the demand point" by compositing the local model. The method has a feature that deals with changes easily by discarding an existing local model and recomposing a new local model, whenever further accumulation of observed data occurs.

Moreover, to apply "JIT modelling" to a large amount of database online, "Large-scale database-based Online modelling (LOM)" has been proposed. The LOM is a technique that makes retrieval of "neighbour" data more efficient by using "stepwise method" and quantization. The stepwise method decreases the dimension of multi-dimensional space of actual process. In addition to that, the multi-dimensional space is quantized.

However, ordinarily in the case of the prediction with LOM method, LOM doesn't consider the effect by operating variables of future process. And the process values in the case of operating a general operation are estimated. Hence, for achieving the long-term estimation enough to reflect future operating amount having practical enough prediction accuracy there is an agenda which is achieved a prediction method considering future operating variables.

As the method estimating after a few hours status the sequential prediction method considering process operating variables by repeating prediction with LOM was proposed. And in the event where setting future operating variables is difficult, long-term prediction can be performed. By modelling the general operation based on LOM, a long-term process behaviour can be estimated using a sequential prediction method without future operating information.

2. LOM (Large-scale database-based Online Modelling)

2.1 Just-In-Time Modelling

A target process is a nonlinear dynamic system, and the characteristics of this system are expressed as a regression model, which is shown as following equation (1).

$$y(t+p) = f\{y(t), y(t-1), \ldots, y(t-n_y), u(t-d), u(t-d-1), \ldots, u(t-d-n_u)\} \quad (1)$$

Where,

$u(t)$ is the control input vector of system at time t,

$y(t)$ is the observational output vector of system at time t,

n_u is the order of control input vector,

n_y is the order of observational output vector,

p is the estimate time (or the predict time),

d is the time delay,

f is the unknown nonlinear function.

10.1016/j.ifacol.2015.10.072

Furthermore, the system input vector x^k and the system output vector y^k are redefined as the following equation (2) and (3).

$$y^k = y(k+p) \quad (2)$$

$$x^k = \{y(k), y(k-1), \ldots, y(k-n_y) \\ u(k-d), u(k-d-1), \ldots, u(k-d-n_u)\} \quad (3)$$

As the time progressing, a large number of data, which is composed of the system input vector x^k and the system output vector y^k, for example (x^1, y^1), (x^2, y^2), … , are stored in the system as a data set $\{(x^k, y^k)\}$, $(k = 1,2\ldots)$,where k is the discrete time. Then, JIT modelling is equal to find the nonlinear function f from the stored data set $\{(x^k, y^k)\}$ whenever it required to estimate (or predict, control).

2.2 LOM

JIT modeling needs a process which measures distance from every conjugated dataset to retrieve neighboring dataset from a database in each estimating. The calculation load becomes enlarged if database grows in stature.

The LOM makes retrieval of neighbour data more efficient by using "stepwise method" and quantization. The stepwise method decreases the dimension of multi-dimensional space of actual process. In addition, the multi-dimensional space is quantized. Step wise method is a technique that adds and deletes input variables by statistical test to decrease input variables within the bound enough for practical use in the regression model. Schematic diagram of the LOM is shown in Fig. 1.

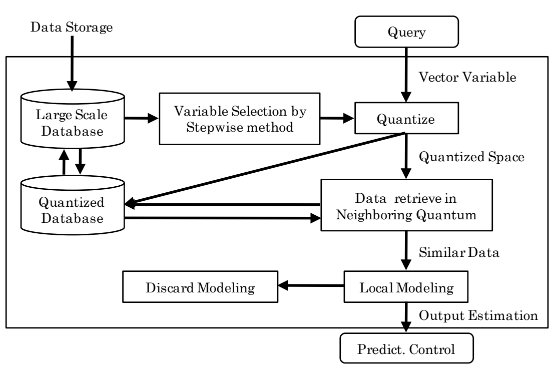

Fig. 1 Schematic diagram of LOM

2.3 Similarity

In this paper, the degree of similarity $s(k_i, k_q)$ between demand data and past data is defined as the infinity norm as follows.

$$s(k_i, k_q) = \left\| x^{k_i} - x^{k_q} \right\|_\infty \quad (4)$$

Where, $\|\cdot\|_\infty$ is infinity norm operator. Datasets of setting number that have small $s(k_i, k_q)$ are adopted as the neighbouring datasets.

3. SEQUENTIAL PREDICTION METHOD by LOM

3.1 Methodology

In order to estimate the long-term plant state, a sequential prediction method is required to repeatedly estimate all control input vectors and observational output vectors. Therefore, the sequential prediction method estimates not only the plant but also the general operation of operator. The outline of the Sequential prediction method is shown in Fig. 2. Moreover, the flow of the method is shown as follows.

1) Demand data $x^{k_q}(t)$ that consists of the present $(t = 0)$ control input vector $u(t)$ and the present observational output vector $y(t)$ is given to LOM. LOM predicts both $u(t+1)$ and $y(t+1)$.

2) Update t to estimate the plant state in next step. Demand data $x^{k_q}(t+1)$ that consists of the control input vector $u(t+1)$ and the observational output vector $y(t+1)$ is given to LOM. LOM predicts both $u(t+2)$ and $y(t+2)$.

3) Repeat processing 2) until the closing time.

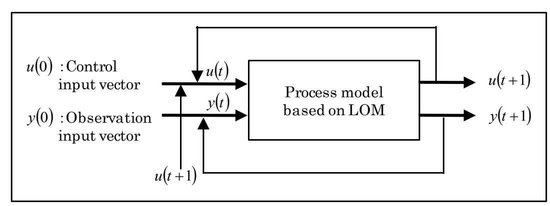

Fig. 2 Outline of Sequential prediction method

3.2 Sequential prediction system

Processing flow for sequential prediction system of LOM is shown in Fig. 3.

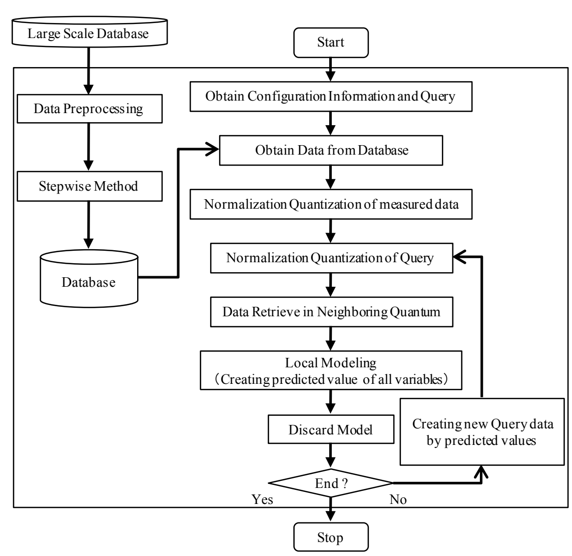

Fig. 3 Processing flow of the Sequential prediction system

3.3 PLS(Partial Least Squares Regression)

PLS is used for improvement of LOM. The procedure is shown as follows.