Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 33 (2007) 192–198
Expert Systems with Applications
www.elsevier.com/locate/eswa

Data mining for yield enhancement in semiconductor manufacturing and an empirical study

Chen-Fu Chien *, Wen-Chih Wang, Jen-Chieh Cheng

Department of Industrial Engineering and Engineering Management, National Tsing Hua University, 101 Section 2 Kuang Fu Road, Hsinchu 30033, Taiwan

Abstract

During wafer fabrication, process data, equipment data, and lot history will be automatically or semi-automatically recorded and accumulated in database for monitoring the process, diagnosing faults, and managing manufacturing. However, in high-tech industry such as semiconductor manufacturing, many factors that are interrelated affect the yield of fabricated wafers. Engineers who rely on personal domain knowledge cannot find possible root causes of defects rapidly and effectively. This study aims to develop a framework for data mining and knowledge discovery from database that consists of a Kruskal–Wallis test, *K*-means clustering, and the variance reduction splitting criterion to investigate the huge amount of semiconductor manufacturing data and infer possible causes of faults and manufacturing process variations. The extracted information and knowledge is helpful to engineers as a basis for trouble shooting and defect diagnosis. We validated this approach with an empirical study in a semiconductor foundry company in Taiwan and the results demonstrated the practical viability of this approach.

Keywords: Data mining; Decision tree; Clustering; Defect diagnosis; Yield enhancement; Semiconductor manufacturing

1. Introduction

The competition of the semiconductor wafer fabrication depends on cost, quality, and the delivery time, especially quality that is a key factor for enterprises to attain long-term competition. In the age of digital information, owing the rise of e-commerce and information technology, a large amount of data has been automatically or semi-automatically collected in modern industry. Decision makers may potentially use the information buried in the raw data to assist their decisions through data mining for possibly identifying the specific patterns of the data.

The semiconductor fabrication processes are complex and lengthy. During the fabrication processes, voluminous data were generated and collected. As the process issue happens, engineers have to identify the root causes of problem as soon as possible to reduce the lost caused by excursion. Most of the engineers rely on their own domain knowledge and experience to identify the specific characteristics of abnormal products. However, such judgments are ineffective and limited by their own domain knowledge. Therefore, it has become an important topic to effectively transfer plethora and complex engineering data into valuable information and knowledge for process improvements and yield enhancement in semiconductor manufacturing. The extracted information and knowledge can assist the engineers as their reference and basis for advanced investigation of the root causes of the defects.

This research aims to propose a framework for mining production data to extract knowledge for manufacturing process monitoring and defect diagnosis in order to remove assignable causes and thus improve the yield. In particular, Kruskal–Wallis test (i.e., K–W test) (Kruskal & Wallis, 1952) and decision tree methodology are applied to analyze and classify abnormal process stages in semiconductor manufacturing. An empirical study was conducted by using

* Corresponding author. Tel.: +886 3 5742648; fax: +886 3 5722685.
E-mail address: cfchien@mx.nthu.edu.tw (C.-F. Chien).

0957-4174/$ - see front matter
doi:10.1016/j.eswa.2006.04.014

real data from a fab to validate the proposed approach. The results showed practical viability of this approach that can efficiently limit the scope for defect diagnosis and derive specific decision rules effectively.

This paper is organized as follows. Section 1 describes research background, significance, and research aims of this study. Section 2 describes the fundamental of this research and reviews related literatures. Section 3 proposes a research framework with detailed procedures for semiconductor data mining and knowledge discovery from database. Section 4 validates the framework with an empirical study. Section 5 concludes with discussion and further research directions.

2. Fundamental

2.1. Data mining

Data mining is defined to be the exploration and analysis, by automatic or semiautomatic means, of large quantities of data stored either in databases, data warehouses, or other information repositories to discover interesting knowledge including meaningful patterns and rules (Berry & Linoff, 1997; Han & Kamber, 2001). Therefore, data mining is discovery-driven not assumption-driven. Owing to e-manufacturing and advanced automation in modern wafer fabs, a huge amount of data is automatically or semi-automatically recorded in data warehouse or engineering data analysis system (EDAS). Indeed, such data can provide a rich resource for knowledge discovery and decision support. To understand, analyze and eventually use this data, a multi-disciplinary approach called data mining has been proposed (Fu, 1997).

Data mining has been applied in many fields to deal with massive data including marketing and sales (Brachman, Khabaza, Kloesgen, Piatetsky-Shapiro, & Simoudis, 1996), finance investment (Deboeck & Kohonen, 1998), astronomy (Fayyad, 1997), health and medical treatment (Evans, Leman, Deters, Fusaro, & Lynch, 1997). In particular, Bose and Mahapatra (2001) conducted a series of data mining applications in many fields including finance, telecommunication, marketing, web analysis, and compared the strengths and weaknesses of various data mining techniques. Only a few number of studies applied data mining techniques in the semiconductor manufacturing to solve the manufacturing issues and optimize the process targets (Braha & Shmilovici, 2002; Fan, Guo, Chen, Hsu, & Wei, 2001; Peng & Chien, 2003; Tsuda, Shiri, Takagi, & Take, 2000).

The data mining models can generally be categorized as four types, i.e., association rules, clustering, classification, and prediction (Fayyad, Piatetsky-Shapiro, & Smyth, 1996; Fu, 1997; Han & Kamber, 2001). Association is the discovery of association rules showing attribute-value conditions that occur frequently together in a given dataset. A popular application of association is market basket analysis, which finds the buying habits of customers by searching for sets of items that are frequently purchased together. Clustering is the process of dividing a dataset into several different groups, those groups also called clusters. The clustering objects are to minimize the inter-class similarity and maximize the intra-class similarity. Classification derives a function or model that identifies the categorical class of an object based on its attributes. A classification model usually is constructed by analyzing the relationship between the attributes and the objects classes in the training dataset. Prediction is a model that predicts a continuous value or future data trends. Other regression model such as polynomial regression, logistic regression and Poisson regression, are used to forecast other particular type of continuous value data.

2.2. Clustering

Clustering is an unsupervised classification of patterns into groups (clusters). Clustering problems arise in many different applications (Kanungo et al., 2002), such as data mining and knowledge discovery, data compression, vector quantization, pattern recognition and pattern classification. A number of algorithms for clustering have been proposed that can be divided into two categories: hierarchal and non-hierarchal approaches.

Among the clustering algorithms that are based on minimizing a formal objective function, the most widely used and studies is *K*-means clustering. The most intuitive and frequently used criterion function in non-hierarchal clustering technique is the squared error criterion, which tends to work well with isolated and compact clusters (Jain, Murty, & Flynn, 1999). However, *K*-means is the simplest and most commonly used algorithm employing a squared error criterion (McQueen, 1967).

2.3. Kruskal–Wallis test

The K–W test is a method for equality of treatment means and is a non-parametric test alternative to the usual ANOVA. The application includes (Daniel, 1990):

(1) The variable of interest is continuous.
(2) The observations are independent both within and among samples.
(3) The populations are identical except for possible differences in location for at least one population.
(4) The data for analysis consist of k random samples of sizes.

Because the K–W test is used to test the null hypothesis that the k treatments are identical against the alternative hypothesis that some of the treatments generate observations that are larger than others are. Therefore, the procedure is designed to be sensitive for testing differences in means.

In this study, K–W test is applied to examine whether there are significant differences among the machine at the same process stages since data distribution is either unknown or with non-normal property.