

Applying data mining to data analysis in manufacturing

*H. Maki, A. Maeda, T. Morita, H. Akimori¹⁾
Systems Development Laboratory, Hitachi, Ltd.
1099 Ohzenji, Asao, Kawasaki, Kanagawa, 215-0013 JAPAN,
Tel: +81-44-966-9111
Fax: +81-44-959-0850(FAX)
e-mail: maki@sdl.hitachi.co.jp, maeda@sdl.hitachi.co.jp,
morita@sdl.hitachi.co.jp
¹⁾Device Development Center, Hitachi, Ltd.
6-16-3 Shinmachi, Ome, Tokyo 198-8512 JAPAN*

Abstract

In analyzing data from advanced manufacturing processes, it is important to integrate various types of data such as numerical, symbolic, and time series data, however, so large is the volume of data created by integration that engineers are not able to examine all of it. Data mining is a method for extracting information from large databases that can help to analyze the integrated data obtained from advanced manufacturing processes. We have developed a data mining method for analyzing manufacturing data that consists of three steps – feature extraction, combinatorial search, and presentation. We applied the method to LSI fault analysis and found that data mining is useful for indicating to engineers where to focus their attention when looking for faults.

Keywords

data mining, rule induction, semiconductor, feature extraction

1 INTRODUCTION

Recent progress in computer technology has made it possible to collect and store large amounts of data that can be measured in terabytes. Data mining is the term for a class of methods used to extract information from large databases. Association rule extraction is one well-known data mining method (Agrawal, 1994). We have developed another data mining method called CHRIS (characteristic rule induction by subspace search), which is more appropriate than association rule extraction for describing the dependence between the occurrence

of events (Maeda, 1998, Maki, 1998). A comparison of the two methods is discussed in a previous paper (Maki, 1998).

In analyzing data in advanced manufacturing processes, it is important to integrate various types of data, such as numerical, symbolic, and time series data. Significant information is often contained in the integrated data, however, such a huge volume of data is created by integration that engineers are not able to examine it all. Although data mining is suitable for this kind of problem, our data mining method is designed to analyze simple types of data and it had to be modified to handle different data types. In this paper we propose a data mining method for use in analyzing manufacturing processes which can accept various

2 DATA MINING PROCESS

2.1 Overview

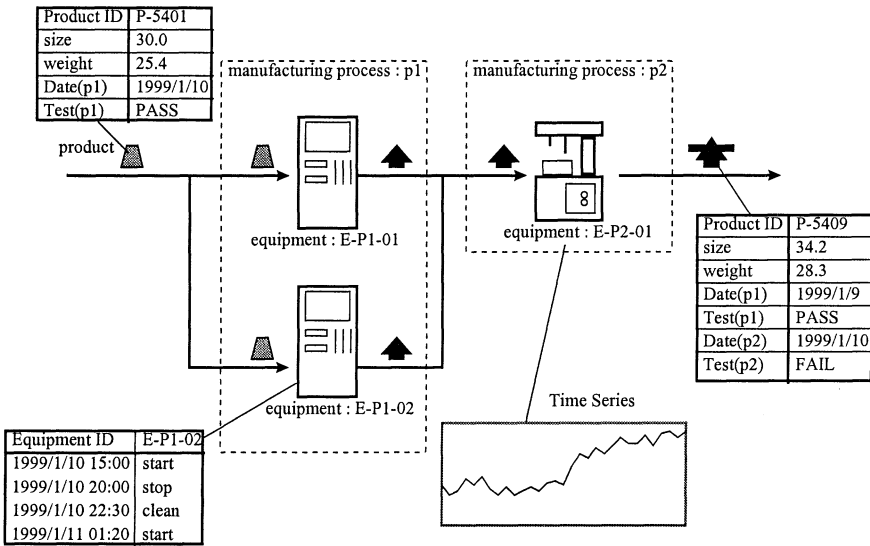


Figure 1 Aspects of and data found in a manufacturing process.

In the manufacturing process there are two basic aspects – the products to be manufactured and the equipment used for manufacturing them. There are also various kinds of data used to monitor the process, such as numerical values, symbolic values, and time series. Examples of the aspects and data are shown in Figure 1 .

To accept these kinds of data, the data mining method that we propose consists of the three steps shown in Figure 2 – feature extraction, combinatorial search, and presentation. First, the data is converted and integrated in the feature extraction

step. Then it is analyzed in the combinatorial search step. Finally, the analysis results are obtained in the presentation step.

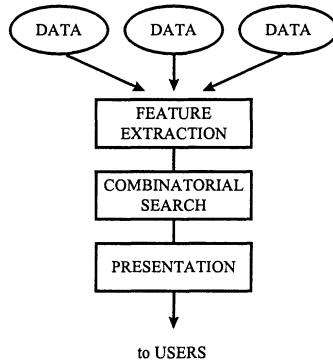


Figure 2 Three steps of the data mining method.

2.2 Feature extraction

In the feature extraction step, the symbolic values which represent the features of the data to be analyzed are extracted. There are several methods for feature extraction, such as clustering, classification, and fluctuation detection of a time series.

For example, numerical data obtained from product inspections, such as size, weight, and voltage, are often classified into three categories – ‘SMALL’, ‘MEDIUM’, and ‘LARGE’. The boundaries of the categories are specified by users or determined automatically. In most cases they are determined by dividing the range of the data into regular intervals.

Figure 3 shows an example of time series data. Using time series analysis methods, fluctuations in the time series are detected and the data is converted into symbolic values such as ‘increase’, ‘decrease’, ‘suddenly rise’ and ‘suddenly fall’.

The symbolic values obtained in this step are regarded as attributes of the two aspects of the manufacturing process, and concatenation of the attributes forms the feature vectors of these aspects (Figure 4). Several methods can be applied to data simultaneously, therefore, a lot of attributes are obtained, which means that the feature vector tends to be multidimensional.

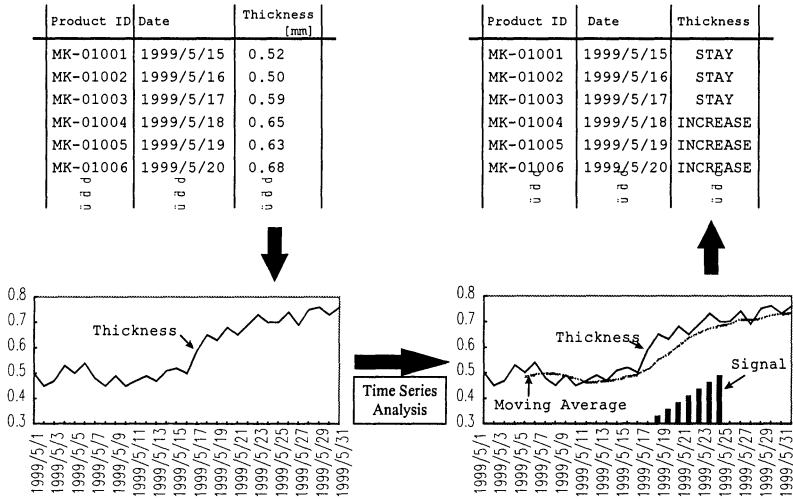


Figure 3 An example of time series feature extraction.

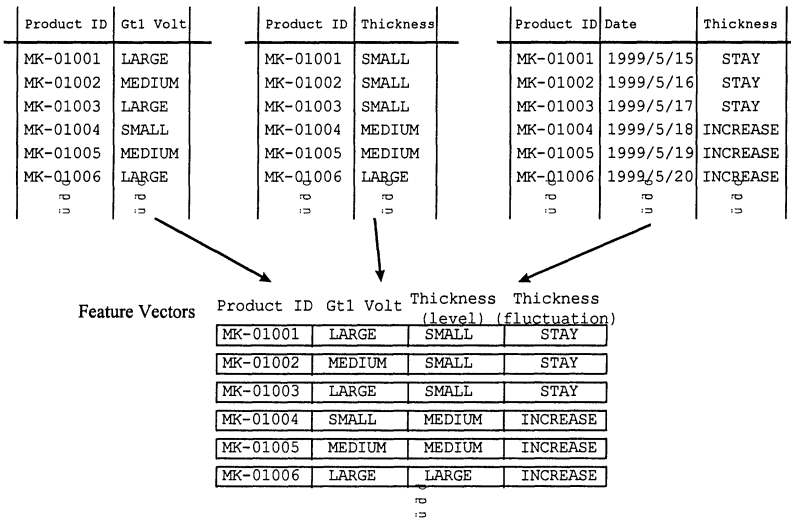


Figure 4 Forming feature vectors.

2.3 Combinatorial search

In the combinatorial search step, a set of the feature vectors obtained in the feature extraction step are accepted, and ‘if-then’ rules (characteristic rules) are generated. We have developed a search method which we call CHRIS for characteristic rule induction. A characteristic rule is a combination of the predicates. For example, the characteristic rule shown in Figure 5 is the combination of two predicates, ‘VOLTAGE is LARGE’ and ‘FUNCTION is FAULT’. ‘VOLTAGE’ and ‘FUNCTION’ are two of the attributes which the feature vectors consist of.

The significance of each characteristic rule is evaluated by what we call the ‘u-measure’ (utility measure). This is the product of two factors – the generality and the accuracy of a rule. The u-measure for the rule ‘if A then B’ is defined as follows:

$$u = P(A)^\alpha \{P(B|A)\log[P(B|A)/P(B)]\}. \tag{1}$$

where $P(A)$ is the probability that predicate A is satisfied, $P(B)$ is the probability that predicate B is satisfied, and $P(B|A)$ is the probability for predicate B when predicate A is satisfied. The factor $P(A)$ corresponds to the generality and $P(B|A)\log[P(B|A)/P(B)]$ corresponds to the accuracy. In general, these two factors are mutually exclusive. The symbol α represents the trade-off parameter by which users can control the relative importance between generality and accuracy. We usually use $0 \leq \alpha \leq 1$.

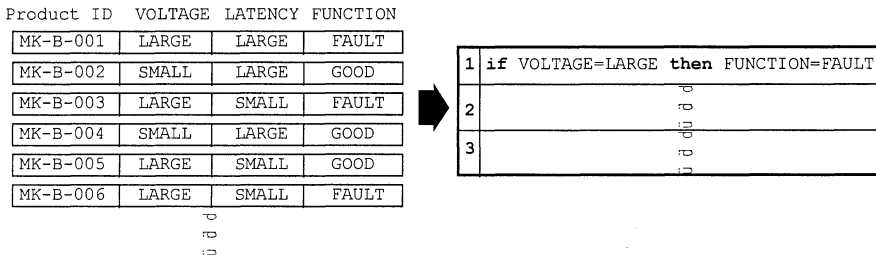


Figure 5 Examples of characteristic rules.

When a user specifies one predicate as the ‘target’, CHRIS searches all possible combinations of the predicates for the characteristic rules whose u-measure values are larger than others. A list of the rules with large u-measure values is obtained as a result of the combinatorial search step.

2.4 Presentation

In the presentation step, the results obtained in the previous step are provided in various ways. One of the simplest ways of presenting the rules given by the results is to write each rule on a display screen, however, data analyzing tools like data

visualization are sometimes necessary to examine the results in detail. Therefore, a data mining system should be able to cooperate with other systems.

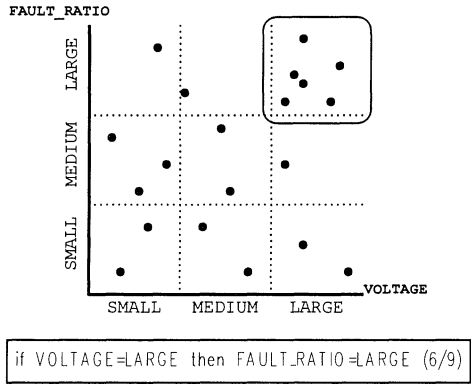


Figure 6 An example of the visualization of data distribution.

Suppose that an engineer wants to use a data visualization tool to see the distribution of the data described by a certain characteristic rule. Figure 6 shows a rule and the data distribution diagram corresponding to it. The attributes, 'VOLTAGE' and 'FAULT_RATIO' are originally numerical, and categorized in the feature extraction step. In order to visualize the distribution of the data, the original numerical data must be retrieved from the database. When the user selects one of the rules in the rule list generated in the combinatorial search step, the data mining system extracts the attribute names which appear in the rule and transfers them to a visualizing tool. Using the attribute names, the visualizing tool retrieves the data from the database and draws the diagram describing the data distribution.

3 SYSTEM ARCHITECTURE

The architecture of the data mining system we propose is shown in Figure 7. Various kinds of data obtained from a manufacturing process are stored in several database servers. There are several feature extraction programs in the system. Each retrieves data from the database servers and performs the feature extraction process individually. The results of the feature extraction process are stored in the feature vector database and are combined with one another to create feature vectors. The combinatorial search program retrieves the feature vectors from the feature vector database and generates characteristic rules. After the combinatorial search, the presentation programs are invoked and the generated rules are passed to them. Some of the presentation programs present the data mining result to users and others transfer it to other information processing systems.

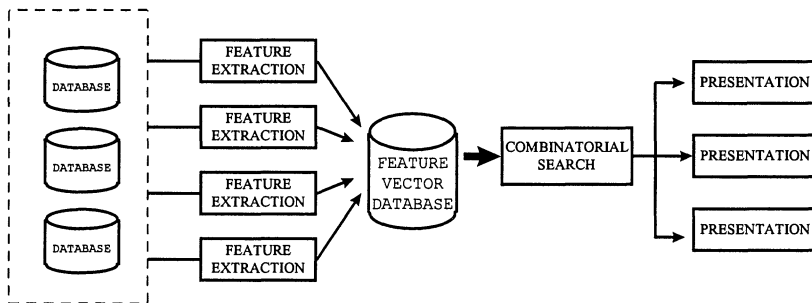


Figure 7 Data mining system architecture.

4 APPLICATION TO LSI MANUFACTURING

As an example of the use of our data mining method in a manufacturing process, we applied it to the analysis of LSI faults.

We analyzed the conditions where faults in the functioning of LSI's occur. The data to be analyzed was a set of values measured in electrical inspections of LSI chips located on a wafer. There were about 100 attributes, and they were all numerical, and placed into three categories – 'LARGE', 'MEDIUM', and 'SMALL'. The predicate 'FUNCTIONAL_FAULT_RATIO is LARGE' was specified as the target, where 'FUNCTIONAL_FAULT_RATIO' was one of the attributes. The result of the combinatorial search was visualized in the distribution diagram. The diagram in Figure 8 indicates the relationship between the fault ratio and the short resistance in layer n2. Engineers examined in detail the relationship between the fault and the condition found by data mining. The examination strongly suggested that something in layer n2 causes the functional fault. Finally, the cause of the fault was identified by a detailed inspection of layer n2 using an electron microscope.

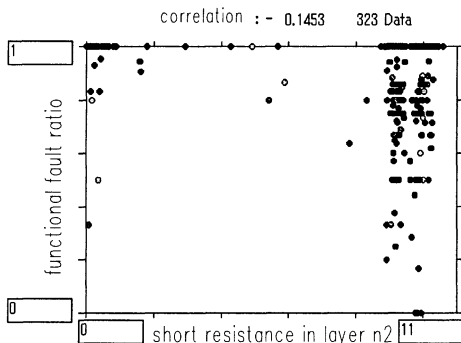


Figure 8 Examination of data mining result.

5 CONCLUSION

We proposed a data mining method for use in manufacturing processes that can analyze various kinds of data. The method consists of three steps – feature extraction, combinatorial search and presentation. We also described the system architecture for using the method.

We applied our data mining method to LSI fault analysis. In the experiment we found that the method was useful for indicating to engineers where they should focus their attention when searching for faults.

6 REFERENCES

- Agrawal, R. and Srikant, R. (1994) Fast Algorithm for Mining Association Rule. *20th VLDB Conference*, Chile
- Maeda, A., Maki, H., and Akimori, H. (1998) Characteristic Rule Induction Algorithm for Data Mining. *PAKDD-98*, Australia
- Maki, H., Maeda, A., and Akimori, H. (1998) Data Mining Application to LSI Fault Analysis. *International Conference on Electrical Engineering (ICEE'98) Vol. 1*, Korea

7 BIOGRAPHY

Hideyuki Maki

He received his BE from Kyoto University, Japan in 1990 and joined Systems Development Laboratory, Hitachi, Ltd. in 1990. His current research interests include data mining and data warehouse systems.

Akira Maeda

He received his MSc from Tokyo University, Japan in 1980 and joined Systems Development Laboratory, Hitachi, Ltd. in 1981. His current research interests include large scale database systems.

Toyohisa Morita

He received his BE in 1987 and his ME in 1989 from Tokyo University, Japan. He has been working at Systems Development Laboratory, Hitachi, Ltd. since 1989. His research interests are data mining and its application to the real world.

Hiroyuki Akimori

He received his BE from Hokkaido University, Japan in 1984. He joined Device Development Center, Hitachi, Ltd. in 1984. His current research interests include improving the wafer fabrication method and yield of CIM systems.