Tool Sequence Analysis and Performance Prediction in the Wafer Fabrication Process

Kittisak Kerdprasop and Nittaya Kerdprasop

***Abstract*—**Many modern manufacturing plants are dealing with large scale multi-dimensional data that are daily and automatically collected from hundreds of operational units in a production line. Maintaining high yield through the statistical process control as a sole monitoring method is obviously inefficient in such highly complicated operations. Recent trend in intelligent manufacturing is to apply data mining techniques to automatically identify patterns and causal relationships leading to poor yield. We thus present in this paper a sequence analysis method, which is one of the advanced data mining techniques, to identify tool patterns from the wafer fabrication processing data. The extracted patterns can reveal and differentiate low performance processes from the normal ones. We also present in this paper a feature selection technique to speed up the data mining steps and show the comparative results of performance prediction.

***Keywords*—** Intelligent manufacturing, Performance prediction, Sequence analysis, Pattern identification, Tool fault detection.

I. Introduction

One of the most complex processes in modern industries is semiconductor manufacturing. Among hundreds of steps, the major processes are production of silicon wafers from pure silicon material, fabrication of integrated circuits onto the raw silicon wafers, assembly by putting the integrated circuit inside a package to form a ready-to-use product, and testing of the finished products [15]. A constant advancement in this industry is due mainly to persistent improvement of the wafer fabrication process. The fabrication process consists of a series of steps to cover special material layers over the wafer surface. Wafers re-enter the same processing machines as each layer is successively covered. Some defects in this complicated process can make the final products fail the test. Early fault detection during this critical manufacturing process can obviously improve product quality and reliability.

Recent trend in intelligent manufacturing is to apply the data mining techniques to automatically identify patterns and causal relationships leading to poor yield [2], [3], [4], [13]. In this paper, we expand the frontier of data mining application to the manufacturing process area by proposing sequence data mining technique to analyze patterns of tool sequences that lead to a poor yield.

Sequence is an ordered set of elements in which each element can be numerical, categorical, or a mixture of different kinds of attributes. The order of elements could be determined by their occurring time or positions. If the order is by time and the elements of a sequence are real values, it is a time series. When the sequence elements are discrete, it is a categorical sequence [20]. Sequence mining is a recently active field of research in knowledge discovery and data mining. The applications of the available techniques are mostly in the areas of bioinformatics and financial analysis. In this paper, we demonstrate the potential application of sequence data mining to discover the operational sequences of tools causing low performances in the semiconductor manufacturing process.

Data mining has been proven to be an efficient tool to automatically discover patterns in manufacturing processes. But for a high dimensional data, the data mining process may significantly slow down due to the excessive memory usage problem. Applying data mining technique to such high dimensional data is not a straightforward task because the induced patterns are normally low accurate in their predictive performances.

In this paper, we thus also present a feature selection technique to remedy the high dimensional problem and also to help improving the pattern accuracy. Our main purpose is the discovery of patterns to detect tool fault that leads to low performance of a wafer lot from the semiconductor fabrication process.

The organization of this paper is as follows. Section 2 is the discussion of several work related to the application of intelligent and data mining techniques to help monitoring and improving yield in the manufacturing process. Section 3 is our proposed method for tool sequence analysis, including the experimental results. Section 4 is the presentation of our feature selection technique, especially designed for the case of high dimensional data. The proposed selection technique is based on the Bayesian method. The experimental results are also presented to confirm the time reduction and accuracy improvement of the induced predictive models. Section 5 is the conclusion of this paper.

This work was supported by grants from the National Research Council of Thailand (NRCT) and Suranaree University of Technology through the funding of Data Engineering Research Unit.

K. Kerdprasop is with the School of Computer Engineering and Data Engineering Research Unit, Suranaree University of Technology, Nakhon Ratchasima, Thailand (e-mail: KittisakThailand@gmail.com).

N. Kerdprasop is an associate professor and the director of Data Engineering Research Unit, School of Computer Engineering, Suranaree University of Technology, 111 University Avenue, Muang District, Nakhon Ratchasima 30000, Thailand (phone: +66-44-224-432; fax: +66-44-224-602; e-mail: nittaya@sut.ac.th).

II. Related Work

In most manufacturing processes, cost, quality, and delivery time are key factors for enterprises to attain long-term competition. During the manufacturing processes, process engineers have to monitor and identify the specific characteristics of abnormal products as soon as possible [5], [7], [14], [22], [26], [27]. Process control is crucially important to the semiconductor industries that operate the multistage manufacturing systems on the product scale of lesser 300 nanometers [17].

Recent manufacturing tools are equipped with sensors to facilitate real-time monitoring of the production process. These tool-state sensor data provide an opportunity for efficient control and optimization. Unfortunately, such data are so overwhelming that timely detection of any fault during the production process is difficult. The fault detection model can however be automatically built from the existing sensor data by means of data mining.

Data mining is about the application of learning algorithm to build model that can best characterize underlying data and accurately predict the class of unknown data. Engineers may potentially use various available data mining techniques to identify specific hidden patterns such as the process fault model to assist the timely investigation of the root causes of the defects.

Ison and colleagues [11], [12] proposed a decision tree classification model to detect fault of plasma etch equipment. The model was built from the five sensor signal data. Many researchers also studied the fault detection problem during the etch process. Goodlin *et al* [9] proposed to build a specific control chart for detecting specific type of faults. They collected tool-state data directly from the etcher. These data consist of 19 variables. The work of Spitzlsperger and colleagues [21] was also based on the statistical method. They adopted the multivariate control chart method to maintain changes in the mean and standard deviation coefficients by remodeling technique.

Recent interest in fault detection has been shifted toward the non-parametric approaches. He and Wang [10] proposed to use the k-nearest neighbor rule for fault detection. Verdier and Ferreira [24], [25] also applied the k-nearest neighbor method, but they proposed to use the adaptive Mahalanobis distance instead of the traditional Euclidean distance. Tafazzoli and Saif [23] proposed a combined support vector machine methodology for process fault diagnosis. Ge and Song [8] applied support vector data to the principal component analysis method to detect process abnormalities.

Most work on fault detection methods has studied the process control problem with a few features of tool-state and process-state measurement data. McCann and his team [16] proposed a rather different setting in which the measurement data from the wafer fabrication process contain as much as 590 features. They applied feature selection technique to select only 40 features for further analysis.

In this work, we apply a data mining technique that can handle 150 features of sequential data, rather than independent and discrete data as proposed in all the previous work. Sequence data mining of manufacturing process appeared in the literature just a few years ago [18], [19].

Our work presented in this paper is different from others in that we apply sequence analysis as an exploratory tool, instead of the classification tool. Moreover, we present a novel feature selection technique based on Bayesian analysis to reduce dimensions of data. A model induced from the reduced feature data set shows not only the increase in prediction accuracy, but also the decrease in model building time.

III. The Tool Sequence Analysis Method

A. Manufacturing Process Data

SETFI (SEmiconductor Tool level Fault Isolation) is a simulated dataset [1] that closely mimics the actual high complexity of semiconductor manufacturing process. The dataset contains 4000 records of the wafer fabrication process. During the process each, a wafer goes through sequence of operations in batch, which is called lot in this dataset. The sequences of hundreds of operations might be different from lot to lot, but these operations involve only twenty tools, number 1 to 20. At each operation unit, only a single tool is in operation.

At the end of the fabrication process, a number of inspection steps are carried out to measure the product performance. Wafer lots that fail the inspection tests need re-processing. Low performance metric is often caused by a small subset of tools. Identifying such problematic tools at an early stage can obviously improve yield performance of the semiconductor manufacturing. Missing values in this dataset are around 25%.

The original SETFI dataset has 300 operational units together with the timestamps of each operation. In this sequence analysis study, we remove the first column (lot#) because it plays no role to the discovered sequence patterns. We also ignore the timestamps as our main objective is categorical sequence analysis, not a time series analysis. We perform sequence analysis over the first 150 operational units because of limitation in our main memory capacity.

From the manufacturing process dataset that contains information of 150 operational units of 4000 wafer lots, we designed the performance sequence analysis framework as illustrated in Fig. 1. The first step of data preparation for our analysis method is to extract features (or variables) containing the tools used in the first 150 operational units together with the performance metric, which is the last column in the SETFI dataset.

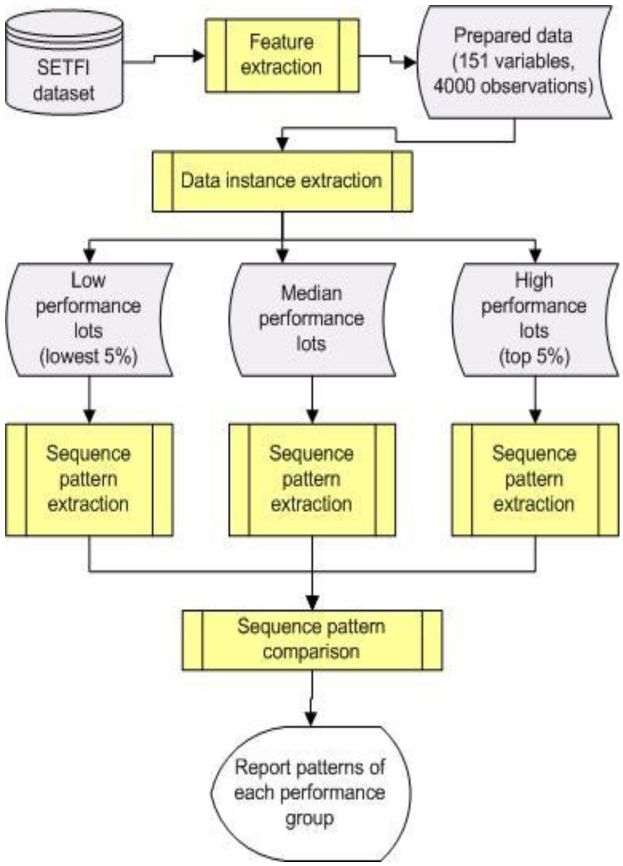


Fig. 1 A framework of the manufacturing sequence analysis

Table 1. Performance of each wafer subgroup

Wafer lots	Performance			
	Max.	Min.	Average	**S.D.**
Low (200 lots)	2574.012	2177.438	2503.816	**66.95**
Median (201 lots)	2790.671	2778.334	2784.345	**3.70**
High (201 lots)	3293.183	2992.259	3062.469	**63.95**
All (4000 lots)	**3293.183**	**2177.438**	**2787.924**	**125.84**

We then divided the dataset into three subgroups: low, median, and high performance lots. Each subgroup contains approximately 200 to 201 data instances. Performance statistics (maximum, minimum, average values, and standard deviation within the subgroup) of the three subgroups are summarized in Table 1.

B. Performance Pattern Mining Technique

Data division into three subgroups and sequence extraction are performed by a computer program, implemented with the R language. Program coding is provided in Fig. 2.

```
library(TraMineR)
mainT <- function(from=1,to=150,lot='low',per=0.05,min=0.3)
{ dat  <- read.csv('C:/sequence-mining/SETFI.csv')
  dat <- dat[-1]       # remove first column
  dat2 <- dat          # make a copy of dataset
  dat2all <- dat2[with(dat2,order(res)),]
  dat2low <- dat2all[1:(per*nrow(dat2)),]
  dat2mid <- dat2all[(0.5*nrow(dat2)-(per*nrow(dat2))/2):
                (0.5*nrow(dat2)+(per*nrow(dat2))/2),]
  dat2top <- dat2all[((1-per)*nrow(dat2)):4000,]
  if (lot=='low') dat2 <- dat2low
  if (lot=='high') dat2 <-dat2top
  if (lot=='all') dat2 <-dat2all
  if (lot=='mid') dat2 <-dat2mid
  mvad.seq <- seqdef(dat2,var=from:to,missing=NA)
   #  Event sequence analysis
  mvad.seqe <- seqecreate(mvad.seq)
  fsubseq <- seqefsub(mvad.seqe, pMinSupport = min )
  print(fsubseq[1:50])
   # plot the 15 most frequent sequences
  plot(fsubseq[1:15],
      main=paste(nrow(dat3),'records at',lot,'lot ,Columns :Col',
        from,'-',to))
}
```

Fig. 2 R program coding for manufacturing sequence analysis

The program calls seqdef, seqecreate, and seqefsub functions from the library TraMineR [6]. There is only one function in the program, which is mainT. The SETFI dataset is in the comma-separated-value (csv) format and stored in the folder sequence-mining. The first command in the program is to read the data and store in the variable 'dat'. The first column, which is the lot number, is then removed. Then the dataset has been sorted in descending order according to the performance value. The ordered data of 4000 wafer lots are called 'dat2all'. This dataset is divided into three subsets: 'data2low', 'data2mid', and 'data2top'.

If the function mainT() has been called with no parameter, the 'data2low' will be processed to search for sequences. To extract sequence patterns from other subsets, the parameter has to be specified. For example, mainT(high) is a command to extract patterns from a group of wafer lots with high performance. The parameters 'from' and 'to' are for identifying data columns to be analyzed. Parameter 'per' is a percentage to split data into high, low, and median subgroups. The last parameter is minimum support value, 'min', in which the value 0.3 has been set as default. To analyze wafer lot patterns, we have to run this program three times, i.e., each execution for each data subset.

C. Sequence Analysis Results and Discussion

The patterns that are returned as a result from running the program can be shown as in Figs 3-5. To save space, we show only the top 25 frequent patterns obtained from each run. The subsequence such as (3>2) means the operation with tool number 3 often followed by the operation with tool number 2. Support value is between 0 and 1. The higher the value, the most frequent this pattern has occurred.

	Subsequence	Support	Count
1	(3>2)	0.960	192
2	(3>4)	0.960	192
3	(1>3)	0.950	190
4	(4>2)	0.945	189
5	(2>4)	0.940	188
6	(4>3)	0.940	188
7	(1>2)	0.935	187
8	(4>1)	0.935	187
9	(2>1)	0.930	186
10	(3>1)	0.925	185
11	(5>2)	0.925	185
12	(5>4)	0.925	185
13	(2>3)	0.920	184
14	(3>5)	0.920	184
15	(4>5)	0.920	184
16	(5>3)	0.920	184
17	(2>5)	0.915	183
18	(5>1)	0.910	182
19	(1>5)	0.900	180
20	(1>4)	0.865	173
21	(3>4)-(3>2)	0.840	168
22	(3>2)-(3>2)	0.830	166
23	(4>2)-(3>2)	0.830	166
24	(5>3)-(3>2)	0.830	166
25	(3>4)-(1>3)	0.825	165

Fig. 3 Sequence patterns of the low performance wafer lots

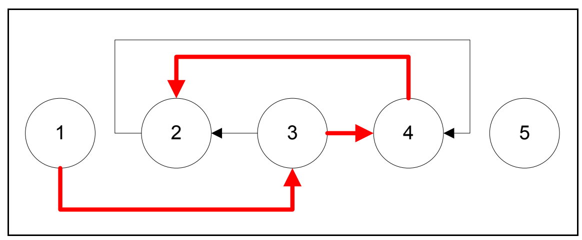
	Subsequence	Support	Count
1	(3>2)	0.9651741	194
2	(2>3)	0.9601990	193
3	(2>5)	0.9552239	192
4	(4>3)	0.9552239	192
5	(2>1)	0.9502488	191
6	(3>4)	0.9452736	190
7	(4>2)	0.9452736	190
8	(5>3)	0.9402985	189
9	(3>5)	0.9253731	186
10	(4>1)	0.9253731	186
11	(3>1)	0.9203980	185
12	(1>3)	0.9104478	183
13	(2>4)	0.9104478	183
14	(5>1)	0.9054726	182
15	(5>2)	0.9054726	182
16	(5>4)	0.9054726	182
17	(1>2)	0.9004975	181
18	(1>5)	0.8905473	179
19	(1>4)	0.8855721	178
20	(3>2)-(2>3)	0.8606965	173
21	(4>5)	0.8557214	172
22	(2>3)-(4>3)	0.8507463	171
23	(3>2)-(3>2)	0.8507463	171
24	(3>2)-(2>1)	0.8457711	170
25	(2>3)-(3>2)	0.8407960	169

Fig. 4 Sequence patterns of the median performance wafer lots

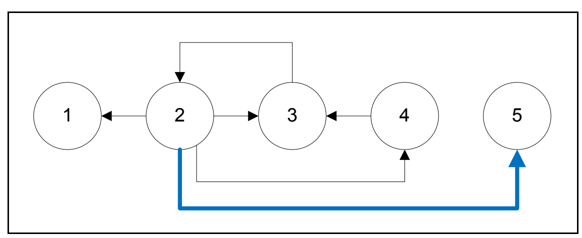
	Subsequence	Support	Count
1	(3>2)	0.9800995	197
2	(2>4)	0.9601990	193
3	(3>1)	0.9502488	191
4	(2>3)	0.9452736	190
5	(1>2)	0.9353234	188
6	(2>1)	0.9353234	188
7	(4>2)	0.9353234	188
8	(4>3)	0.9353234	188
9	(5>1)	0.9303483	187
10	(5>3)	0.9303483	187
11	(1>4)	0.9203980	185
12	(2>5)	0.9203980	185
13	(3>4)	0.9104478	183
14	(5>4)	0.9104478	183
15	(1>3)	0.9054726	182
16	(3>5)	0.9054726	182
17	(1>5)	0.8855721	178
18	(5>2)	0.8855721	178
19	(4>5)	0.8805970	177
20	(4>1)	0.8756219	176
21	(3>2)-(2>3)	0.8606965	173
22	(2>3)-(3>2)	0.8507463	171
23	(4>2)-(2>3)	0.8457711	170
24	(4>2)-(3>2)	0.8457711	170
25	(3>2)-(3>2)	0.8308458	167

Fig. 5 Sequence patterns of the high performance wafer lots

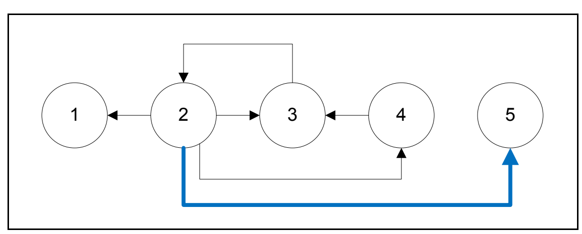
For an easy comparative analysis, we draw the top-5 subsequence diagrams of frequently occurred patterns. The top-5 frequent subsequence patterns of low, median, and high performance groups are shown in Fig.6.



(a) low performance subgroup



(b) median performance subgroup



(c) high performance subgroup

Fig. 6 Top-5 sequence patterns of wafer lots in a low, median, and high performance subgroups

The diagram in Fig.6 shows subsequence states of operations that involve tools number 1-5. There are twenty tools applied in the wafer fabrication line of processes. Tools number 1 through 5 are the set of tools mostly applied (more than 50% of the operational units) in the wafer fabrication process.

On comparing the top-5 subsequence patterns, we notice some unique patterns in each subgroup (denoted as a thick line in the diagram). The unique pattern in a high performance subgroup is the sequence of operation tool3 → tool1, that is, the tool3 has been applied in the process just before the tool1. This pattern does not exist in other two subgroups.

The median performance subgroup has a unique sequence tool2 → tool5. Only this subgroup involves tool5 in the frequent subsequence patterns.

The wafer lots that are in the lowest performance subgroup illustrate three unique patterns: tool1 → tool3, tool3 → tool4, and tool4 → tool2.

From the subsequence patterns, we found that the wafer lots with low performance show high involvement with a tool number 3. We thus further analyze the patterns of frequently occurred operational subsequences. At this stage, we compare only the subsequences with support value greater than or equal to 0.95 (from the full scale of 1.0). The comparative results are illustrated in Fig.7.

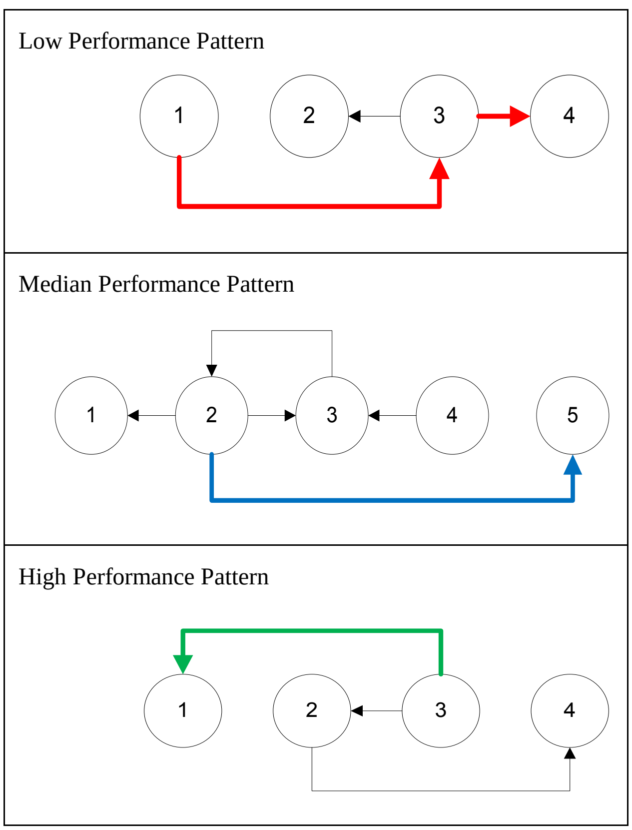


Fig. 7 Comparison of sequence patterns occurring at a high frequency rate ≥ 0.95

From the frequently occurred subsequence pattern analysis, we can thus summarize that high yield wafer fabrication lots involve the tool subsequence tool3 → tool1, and the lots at moderate yield are those involving the subsequence tool2 → tool5. The subgroup of wafer lots with lowest performance shows two distinct subsequence patterns: tool1 → tool3 and tool3 → tool4. Note that these two subsequence patterns are not necessarily continuous.

The sequence patterns associating more than two tools appear at much lower support values. We list here some patterns of the three performance subgroups.

Low:	*Subsequence*	*Support*	*Count*
	(3>4)-(3>2)	0.840	168
	(3>2)-(3>2)	0.830	166
	(4>2)-(3>2)	0.830	166
	(5>3)-(3>2)	0.830	166
	(3>4)-(1>3)	0.825	165

Median:	*Subsequence*	*Support*	*Count*
	(3>2)-(2>3)	0.8606965	173
	(2>3)-(4>3)	0.8507463	171
	(3>2)-(3>2)	0.8507463	171
	(3>2)-(2>1)	0.8457711	170
	(2>3)-(3>2)	0.8407960	169

High:	*Subsequence*	*Support*	*Count*
	(3>2)-(2>3)	0.8606965	173
	(2>3)-(3>2)	0.8507463	171
	(4>2)-(2>3)	0.8457711	170
	(4>2)-(3>2)	0.8457711	170
	(3>2)-(3>2)	0.8308458	167

IV. Wafer Performance Prediction Method

A. Performance Prediction Methodology

The feature selection technique and the discovery of tool patterns leading to a wafer lot showing low/high performances can be explained as follows:

Step 1: Data preparation

1.1 Remove irrelevant features, that are, lot # and timestamp ($T_1 - T_{300}$)

1.2 Replace missing value with a symbol '?' (There are around 25% of missing values in several operational units.)

1.3 Sort data in ascending order according to the performance value (these data instances will be referred to by their numbers ranging from 1-4000)

Step 2: Feature selection

2.1 Set the threshold value T as 0.60

2.2 Add a new feature, called class, with two distinct values: c0 (low performance) and c1 (high performance). The first half of instances (number 1-2000) are assigned class c0 and the rest (instance number 2001-4000) are class c1.

2.3 Prepare train and test data sets. Test set contains 200 data instances that are the first 100 instances and the last 100 instances (instances number 1-100 and 3901-4000). The train data set contains 1000 data instances that are instances number 101-600 and 3401-3900.

2.4 For all the available 300 features, do the analysis to justify appropriate training size by

2.4.1 Dividing the train data into 20 subsets; each subset is an increment of 50 data instances (that is, 50, 100, 150, 200, 250, …, 950, 1000 instances)

2.4.2 Then test predictive accuracy of each data subset using Naïve Bayes learning algorithm

2.5 Select a set of features that yield predictive accuracy greater than the threshold T

Step 3: Perform data mining

3.1 Use selected features that have been analyzed by step 2 to extract features in both the train data and test data sets

3.2 Run selected data mining algorithm on the train data set and then test the accuracy of predictive model using the test data set

B. Feature Selection Results

The feature selection (step 2) in the previous methodologies is our main contribution. From the step 2.4 that is the analysis for a proper training size, we obtain the analysis result as shown in Fig.8. In the figure, Y-axis is predictive accuracy, whereas X-axis is the size of training data. Each scale along the X-axis is the power of 50. The best accuracy is at X=7, which is the data subset of size 7×50 = 350 data instances.

From the step 2.5, which is the selection of features giving accuracy higher than 0.6, we obtain the selection result as shown in Fig. 9. With the accuracy threshold of 0.6, there are 7 features selected by the proposed method. These features are columns number 10, 85, 133, 141, 147, 198, and 261. Fig. 10 shows the results of both proper training size and the best discriminative features.

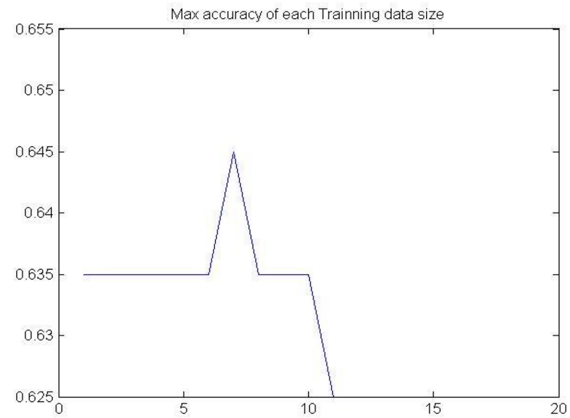


Fig. 8 Justification for selecting appropriate training data size

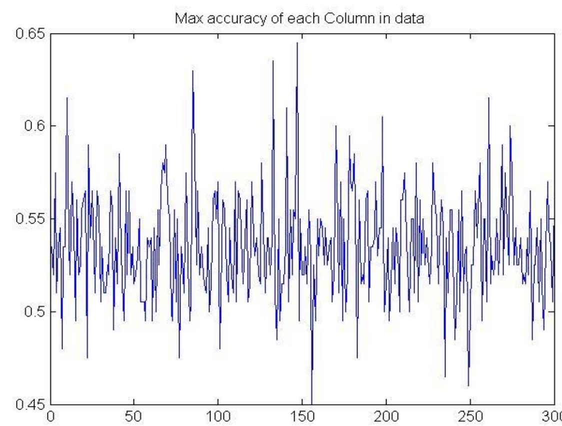


Fig. 9 The analysis result of feature selection

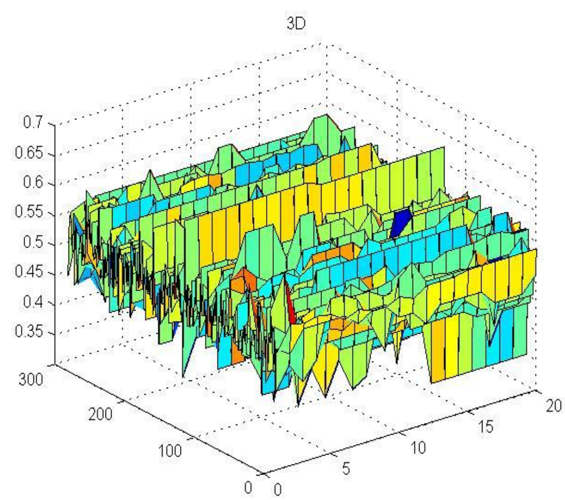


Fig. 10 Analysis results of both proper training data size (Z-axis, each scale is a power of 50) and the best discriminative features (X-axis) according to the predictive accuracy metric (Y-axis)

C. Comparison of Wafer Performance Predictive Models

After selecting proper features and justifying appropriate training size, the prepared data are then used to induce a predictive model. We perform experimentation with seven model induction methods:

- decision tree induction (C4.5 algorithm),
- random forest (RF),
- naïve Bayes (NB),
- k-nearest neighbors (k-NN, with k=10),
- Adaboost (AdaB),
- artificial neural network (ANN, using voted perceptron algorithm), and
- support vector machine (SVM).

All experimentations have been tested with the same data set. The predictive performances of the seven algorithms are summarized in Table 2 and graphically compared in Figs. 11 and 12.

It can be seen that our proposed feature selection technique can improve predictive performance, as well as can reduce training and testing time. The best predictive model in terms of accuracy is the k-NN when ten nearest neighbors (that is, k=10) are taken into account. But most users might consider the performance predictive model represented as a decision tree be more comprehensible as it can convey information in an easy-to-understand format.

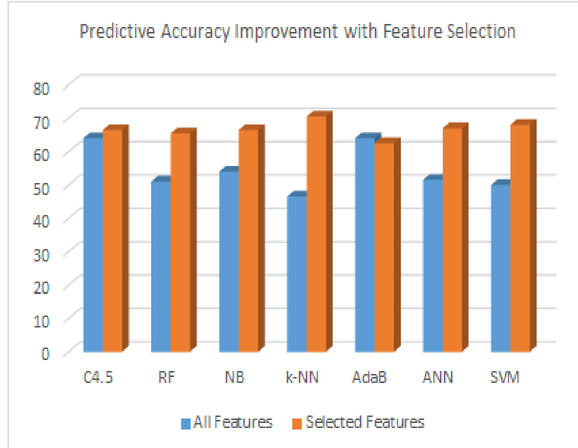


Fig. 11 Accuracy improvement with the proposed feature selection method

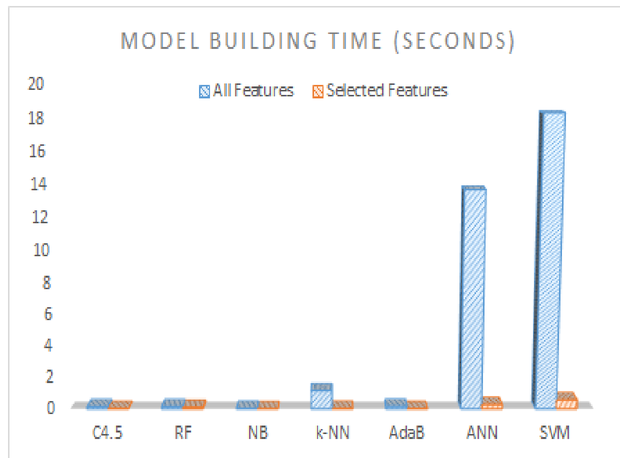


Fig. 12 Time reduction in learning to create predictive models

Table 2. Predictive performance of each learning algorithm.

Learning algorithm	**Use all features (300 operations)**		**Feature selection (7 operations)**	
	Accuracy	**Time (seconds)**	**Accuracy**	**Time (seconds)**
Decision Tree	64.00%	0.14	66.50%	0.02
Random Forest	51.00%	0.14	65.50%	0.10
Naïve Bayes	54.00%	0.03	66.50%	0.00
k-NN (k =10)	46.50%	1.17	70.50%	0.03
AdaBoost	64.00%	0.15	62.50%	0.01
ANN (Voted Perceptron)	51.50%	13.67	67.00%	0.27
Support Vector Machine	50.00%	18.28	68.00%	0.58

V. Conclusion

For the highly complex manufacturing processes such as semiconductor industries, hundreds of metrology data are available for process engineers to analyze for the purpose of maintaining efficient operations and getting optimum yield of high quality products. For such a large volume of measurement data, automatic data analysis technique is essential. We thus investigate the application of advanced mining technique, namely sequence data mining, to help analyzing problematic sequences in the wafer fabrication process of semiconductor industries.

We design a tool sequence analysis framework to group wafer operational data into three categories: processes with low, high, and moderate performance metrics. Data in each category are then analyzed with the sequence mining program written in R language. We find from the experimental results that the frequently occurred subsequence patterns of each category show unique patterns.

From the sequence analysis experiments, we also notice that the mining program has been confronted with the memory space limitation. This is due to the tremendous amount of candidate subsequence patterns to be generated during the search process. We then devise a feature selection technique to remedy this problem.

We propose a novel technique for selecting only discriminative features and also propose a technique to extract appropriate amount of train data, instead of learning from the huge amount of all available data. The experimental results confirm efficiency of our feature selection technique. From the seven learning algorithms, six algorithms show significant improvement in terms of predictive accuracy and model induction time. The k-nearest neighbor model shows the highest improvement; predictive accuracy has been improved from 46.50% to 70.50%. But we found that the decision tree model is good at describing operational unit, which is annotated with specific tool number that has been applied in that unit. Such specific information is helpful for searching the root cause of problematic lots.

Sequence analysis technique presented in this paper is semi-automatic in the sense that unique pattern inspection has to be done by human. We thus plan to further our research towards the design and implementation of an automatic tool to timely detect process trends leading to low performance products.

References

[1] Advanced Analytics – Intel, SETFI: Manufacturing data: Semiconductor tool fault isolation. *Causality Workbench Repository*, http://www.causality.inf.ethz.ch/repository.php, access on June, 2014.

[2] H. Ahmadi, A. Moosavian, and M. Khazaee, "An appropriate approach for misalignment fault diagnosis based on feature selection and least square support vector machine," *International Journal of Mechanics*, vol. 6, issue 2, 2012, pp. 97-104.

[3] T. Chen, H. Tsai, and H. Wu, "Wafer lot output time prediction with a hybrid artificial neural network," *WSEAS Transactions on Computers*, vol. 5, issue 5, 2006, pp. 817-823.

[4] T. Chen, and Y. Wang, "A hybrid intelligent approach for output projection in a semiconductor fabrication plant," *Intelligent Data Analysis*, vol. 12, no. 1, 2008, pp. 129-144.

[5] R. Furferi and L. Governi, "Development of an artificial vision inspection system for real-time defect detection and classification on circular knitting machines," *WSEAS Transactions on Computers*, vol. 5, issue 6, 2006, pp. 1186-1193.

[6] A. Gabadinho, G. Ritschard, M. Studer, and N. Muller, *Mining Sequence Data in R with the TraMineR Package: A User's Guide*. University of Geneva, Switzerland, http://mephisto.unige.ch/traminer/, accessed on December, 2013.

[7] M. Gallo, G. Guizzi, and V. Zoppoli, "An integrated approach to develop a simulation model in manufacturing processes," *International Journal of Systems Applications, Engineering & Development*, vol. 1, issue 4, 2007, pp. 137-145.

[8] Z. Ge and Z. Song, "Semiconductor manufacturing process monitoring based on adaptive substatistical PCA," *IEEE Transactions on Semiconductor Manufacturing*, vol. 23, no. 1, 2010, pp. 99-108.

[9] B. E. Goodlin, D. S. Boning, H. H. Sawin, and B. M. Wise, "Simultaneous fault detection and classification for semiconductor manufacturing tools," *Journal of Electrochemical Society*, vol. 150, no. 12, 2003, pp. G778-G784.

[10] Q. P. He and J. Wang, "Fault detection using the k-nearest neighbour rule for semiconductor manufacturing processes," *IEEE Transactions on Semiconductor Manufacturing*, vol. 20, no. 4, 2007, pp. 345-354.

[11] A. M. Ison, W. Li, and C. J. Spanos, "Fault diagnosis of plasma etch equipment," in *Proc. IEEE Int. Symposium on Semiconductor Manufacturing*, 1997, pp. B49-B52.

[12] A.M. Ison and C. Spanos, "Robust fault detection and fault classification of semiconductor manufacturing equipment," in *Proc. Int. Symposium on Semiconductor Manufacturing*, Tokyo, 2-4 October 1996.

[13] K. Kerdprasop, and N. Kerdprasop, "A data mining approach to automate fault detection model development in the semiconductor manufacturing process," *International Journal of Mechanics*, Vol. 5, Issue 4, 2011, pp. 336-344.

[14] H.D. Lin, C.Y. Chung, W.T. Lin, "Principal component analysis based on wavelet characteristics applied to automated surface defect inspection," *WSEAS Transactions on Computer Research*, vol. 3, issue 4, 2008, pp. 193-202.

[15] G. S. May and C. J. Spanos, *Fundamentals of Semiconductor Manufacturing and Process Control*, John Wiley & Sons, 2006.

[16] M. McCann, Y. Li, L. Maguire, and A. Johnston, "Causality challenge: benchmarking relevant signal components for effective monitoring and process control," in *Proc. JMLR Workshop*, Canada, 12 December 2008, pp.277-288.

[17] S.J. Qin, G. Cherry, R. Good, J. Wang, and C.A. Harrison, "Semiconductor manufacturing process control and monitoring: a fab-wide framework," *Journal of Process Control*, vol.16, 2006, pp.179-191.

[18] L. Rokach, R. Romano, and O. Maimon, "Mining manufacturing databases to discover the effect of operation sequence on the product quality," *Journal of Intelligent manufacturing*, vol. 19, 2008, pp.313-325.

[19] N. Ruschin-Rimini, O. Maimon, and R. Romano, "Visual analysis of quality-related manufacturing data using fractal

geometry," *Journal of Intelligent manufacturing*, vol. 23, 2012, pp.481-495.

[20] S. Sarawagi, "Sequence data mining," in S. Bandyopadhyay, U. Maulik, L.B. Holder, and D.J. Cook (eds.), *Advanced Methods for Knowledge Discovery from Complex Data*, pp.153-187. Springer, 2005.

[21] G. Spitzlsperger, C. Schmidt, G. Ernst, H. Strasser, and M. Speil, "Fault detection for a via etch process using adaptive multivariate methods," *IEEE Transactions on Semiconductor Manufacturing*, vol. 18, no. 4, 2005, pp. 528-533.

[22] S.K. Subramaniam, S.H. Fusin, R.S.S. Singh, and A.H. Hamidon, "Production monitoring system for monitoring the industrial shop floor performance," *International Journal of Systems Applications, Engineering & Development*, vol. 3, issue 1, 2009, pp. 28-35.

[23] E. Tafazzoli and M. Saif, "Application of combined support vector machines in process fault diagnosis," in *Proc. American Control Conference*, 2009, pp. 3429-3433.

[24] G. Verdier and A. Ferreira, "Fault detection with an adaptive distance for the k-nearest neighbor rule," in *Proc. Int. Conference on Computer & Industrial Engineering*, Troyes, France, 6-9 July 2009, pp.1273-1278.

[25] G. Verdier and A. Ferreira, "Adaptive Mahalanobis distance and k-nearest neighbour rule for fault detection in semiconductor manufacturing," *IEEE Transactions on Semiconductor Manufacturing*, vol. 24, no. 1, 2011, pp. 59-68.

[26] Z. Wang, S. Ji, L. Zhang, Y. Huan, and G. Yang, "Research on automatic detecting technology of Razor's assembling quality based on the machine vision system," *WSEAS Transactions on Computers*, vol. 5, issue 6, 2006, pp. 1319-1324.

[27] Z.W. Zhong and N.G. Shankar, "A clustering technique for defect inspection," *WSEAS Transactions on Electronics*, vol. 2, issue 4, 2005, pp. 185-188.

Kittisak Kerdprasop is an associate professor at the School of Computer Engineering and one of the principal researchers of Data Engineering Research Unit, Suranaree University of Technology, Thailand. He received his bachelor degree in Mathematics from Srinakarinwirot University, Thailand, in 1986, master degree in computer science from the Prince of Songkla University, Thailand, in 1991 and doctoral degree in computer science from Nova Southeastern University, USA, in 1999. His current research includes Data mining, Machine Learning, Artificial Intelligence, Logic and Functional Programming, Probabilistic Databases and Knowledge Bases.

Nittaya Kerdprasop is an associate professor and the director of Data Engineering Research Unit, School of Computer Engineering, Suranaree University of Technology, Thailand. She received her B.S. in radiation techniques from Mahidol University, Thailand, in 1985, M.S. in computer science from the Prince of Songkla University, Thailand, in 1991 and Ph.D. in computer science from Nova Southeastern University, U.S.A., in 1999. She is a member of IAENG, ACM, and IEEE Computer Society. Her research of interest includes Knowledge Discovery in Databases, Data Mining, Artificial Intelligence, Logic and Constraint Programming, Deductive and Active Databases.