

단일 클래스 분류기법을 이용한 반도체 공정 주기 신호의 이상분류

조민영 · 백준걸[†]

고려대학교 산업경영공학과

One-class Classification based Fault Classification for Semiconductor Process Cyclic Signal

Minyoung Cho · Jun-Geol Baek

School of Industrial Management Engineering, Korea University

Process control is essential to operate the semiconductor process efficiently. This paper consider fault classification of semiconductor based cyclic signal for process control. In general, process signal usually take the different pattern depending on some different cause of fault. If faults can be classified by cause of faults, it could improve the process control through a definite and rapid diagnosis. One of the most important thing is a finding definite diagnosis in fault classification, even-though it is classified several times. This paper proposes the method that one-class classifier classify fault causes as each classes. Hotelling T2 chart, kNNDD(k-Nearest Neighbor Data Description), Distance based Novelty Detection are used to perform the one-class classifier. PCA(Principal Component Analysis) is also used to reduce the data dimension because the length of process signal is too long generally. In experiment, it generates the data based real signal patterns from semiconductor process. The objective of this experiment is to compare between the proposed method and SVM(Support Vector Machine). Most of the experiments' results show that proposed method using Distance based Novelty Detection has a good performance in classification and diagnosis problems.

Keyword: fault classification, semiconductor process, one-class classification, cyclic signal, multi-class classification

1. 서론

컴퓨터를 비롯한 각종 디지털 정보기기의 수요가 증가함에 따라 반도체 산업의 중요성은 증대되고 있다. 기계의 휴대성 증대 및 많은 데이터 처리를 위한 고집적(high integrated) 반도체 제품의 생산은 반도체 산업의 당면과제로 고집적 반도체

생산을 위해 초미세 공정의 안정적인 운영은 필수적이다. 기존의 반도체 생산 공정에 비해 고집적을 위한 나노공정의 운영은 보다 정밀함이 요구되며 작은 변동에 따른 불량 발생 위험도 증가하였다. 또한 반도체 시장의 치열한 경쟁은 단위 생산비를 절감하기 위한 방안으로의 수율 증대를 추구함에 따라 불량률 감소를 목표로 삼게 하였다.

본 과제는 정보통신산업진흥원의 SW공학 요소기술 개발과 전문 인력 양성사업의 결과물임을 밝힙니다. 본 연구는 지식경제부 및 정보통신산업진흥원의 정보통신연구기반구축사업의 연구결과로 수행되었음(NIPA-2011-(B1110-1101-0002)).

[†]연락처: 백준걸 교수, 136-701 서울시 성북구 안암동 5가 고려대학교 산업경영공학과, Fax : 02-929-5888, E-mail : jungeol@korea.ac.kr
투고일(2011년 12월 16일), 심사일(1차 : 2012년 03월 03일), 게재확정일(2012년 03월 03일).

공정 장비에서 발생하는 신호를 바탕으로 현재 진행되는 공정의 이상을 탐지하고 원인을 분류(FDC : Fault detection and classification)하여 불량 발생을 예방하기 위한 신속한 조치를 취하는 것은 불량률을 감소시키는 가장 중요한 방안이다(Goodlin *et al.*, 2003). 이를 통해 현재 진행되고 있는 공정을 개선하여 생산성을 향상 시키고 불량률을 감소시켜 이윤 창출과 기업 경쟁력 강화에 크게 기여할 수 있다(Kymal *et al.*, 2006). 따라서 공정 장비의 상태를 지속적으로 관리하여 공정을 개선하기 위한 노력으로 여러 공정관리 방법들이 제안되었다.

공정을 관리하기 위한 방법으로는 공정에서 발생하는 자료를 통계적으로 관리하는 통계 공정 제어(Statistical Process Control, SPC) 기법이 대표적이다. 이는 생산된 제품의 관측치의 평균이 정규분포를 근사적으로 따른다는 가정 하에 구축되는 Shewart 관리도, 공정 자료의 자기상관성이 존재할 때 이용하는 CUSUM 관리도(CUmulative SUM) 및 EWMA(Exponentially Weighted Moving Average) 등으로 구성된다. 이러한 다양한 SPC 관리도들을 이용하여 공정의 모니터링을 통해 이상을 감지하여 제품의 품질을 향상시킨다(Montgomery, 1996; Chiang *et al.*, 2001).

컴퓨터 기술의 발전은 공정의 완료로 발생하는 결과물을 바탕으로 공정 상태를 진단하는 대신 공정과 동시에 발생하는 주기 신호를 바탕으로 공정 상태를 진단하는 것을 가능하게 하였다. 주기신호 데이터란 하나의 주기를 기준으로 특정패턴을 반복하는 공정 데이터이다. 반도체 제조 공정은 여러 단위 공정으로 구성되어 있으며 단위공정에서 일정한 주기마다 특정 패턴을 갖는 주기 신호가 센서(sensor)를 통해 획득된다(Park, *et al.*, 2010). 공정 상황이 정상일 경우 주기신호 데이터는 일정한 패턴을 유지하지만 비정상일 경우 정상과는 다른 패턴을 보인다. 또한, 공정에 이상이 있을 경우 발생하는 원인에 따라 특정한 형태를 따르는 주기신호가 나타난다. 이와 같이 많은 정보를 포함하고 있는 주기 신호 데이터를 이용하면 보다 효과적인 품질관리를 진행할 수 있다. 공정의 이상탐지를 통해 불량이 발생하기 전 이상을 미리 찾아내고 발생한 이상의 분류를 거쳐 원인이 무엇인지를 파악하여 공정의 진단 및 개선이 가능하다(Lee *et al.*, 2007).

주기신호의 특징을 파악하고 이상을 진단하는 것은 공정을 효과적으로 관리하기 위한 매우 중요한 과제이다. 특히 반도체 공정과 같이 제품생산에 많은 요인들이 영향을 미치는 경우, 이상이 발생되었다고 하더라도 원인을 찾아내기는 매우 어렵기에 이상 분류를 통한 진단이 필수적이다. 반도체 공정 주기신호를 이용한 이상 탐지와 관련된 연구는 많이 진행된 반면 이상 분류의 경우 직접적으로 분류를 통해 얻어지는 효과를 예측하기 어렵기에 상대적으로 연구된 바가 적었다.

본 연구에서는 이상 원인에 따른 신호의 패턴이 알려져 있다는 가정을 바탕으로 공정의 이상을 분류(classification)하는 방법을 제시하고자 한다. 일반적으로 공정의 이상이 탐지되었다면 이상의 원인을 찾아 개선하고자 해당 공정 시스템을 검사

하게 된다. 이때 어떤 원인에 의해 이상이 발생했는지에 대한 정보가 존재한다면 이상의 원인을 찾는 데 걸리는 시간이 발생하지 않아 원인을 해결하는 데까지의 소요시간이 감소한다. 즉, 이상 원인에 따라 신호의 패턴이 알려져 있다면 이상 분류를 통해서 보다 신속한 이상 원인의 진단 및 해결이 가능하다.

신속한 이상 원인의 진단을 위해서는 단순히 발생한 이상이 어떠한 종류인지 분류하는 성능을 높이는 데 그치는 것이 아니라 오분류되는 경우 적합한 원인을 찾는 데까지 소요되는 시간을 줄이는 것이 필요하다. 이를 위해 단일 클래스 분류(One-Class Classification : OCC) 기법을 활용할 수 있다. 동일한 이상 원인에 따라 발생한 신호를 바탕으로 기준 모델을 만들고 이상이 발생했을 때 해당 이상인지 아닌지를 분류하여 가장 적합한 이상 클래스부터 그렇지 않은 이상 클래스별로 순서를 정하여 이상의 진단을 차례로 진행하는 것이다. 즉, 이상 분류의 정확도와 동시에 이상의 원인을 찾기 위한 오분류 횟수를 줄이는 것을 목적으로 한다.

논문의 구성은 다음과 같다. 본 논문의 제 2장에서는 Hotelling의 T^2 와 k-NNDD(k-Nearest Neighbor Data Description), Distance based Novelty detection 등의 단일 클래스 분류 기법에 대해 기술한다. 또한, PCA 변환을 거친 자료에 대해 분류 알고리즘을 활용한 반도체 공정 주기 신호의 이상 분류 방법을 제안하였다. 제 3장에서는 기존 SVM과 제안 기법의 성능을 비교를 위해 수행한 시뮬레이션 실험 결과에 대해 언급하며 마지막으로 결론과 추후 연구의 내용을 제시한다.

2. 단일 클래스 분류 기반 이상 분류

본 연구는 보다 신속한 이상 원인의 진단을 위해 가장 적합한 이상 클래스부터 그렇지 않은 이상 클래스별로 이상 진단 순서를 결정한다. 다중 클래스 분류를 이용하여 이상 분류를 할 경우 가장 적합한 하나의 이상 클래스만 결과 값으로 도출된다. 때문에 본 연구에서는 각각의 동일한 이상 원인에 따라 발생한 신호를 바탕으로 단일 클래스 분류 모델을 만들어 이상 분류를 진행한다. 각각의 단일 클래스 분류 모델을 이용하여 가장 적합한 이상 클래스부터 이상 진단 순서를 결정한다. 본 연구에서 적용한 단일 클래스 분류 기법으로는 Hotelling의 T^2 관리도, k-NNDD, Distance based Novelty Detection으로 대표적인 분류 기법이다. 제 2.1절에서는 Hotelling의 T^2 관리도, k-NNDD, Distance based Novelty Detection 등의 단일 클래스 분류 알고리즘에 대해 기술한다.

2.1 단일 클래스 분류 알고리즘

(1) Hotelling's T^2 관리도

새로운 관측 값과 기존 정상 관측 값과의 유사성을 활용한

다면 정상인지 비정상인지를 구분할 수 있다. 이때 기존 관측 값과의 유사성이 높을수록 관측 값이 정상일 가능성이 크고 반면, 유사성이 적을수록 비정상일 가능성이 높아진다. 이때 유사성을 어떻게 측정하며 유사하다고 판단할 수 있는 기준 근거를 어떻게 구하는 지에 대한 기준이 필요하다.

이때 주로 사용된 방법은 두 관측치 사이의 유클리드 거리를 사용하는 방법인데 다변량 공정의 모니터링에서는 단순 거리뿐만 아니라 변수간의 상관관계를 동시에 고려해야 한다. 다변량 관리도를 이용해 모니터링 할 경우 상관관계가 고려되지 않기 때문에 정확한 이상탐지를 하는데 어려움이 있다. 때문에 상관관계가 존재하는 다수의 관측치들을 관리하기 위한 다변량 관리도가 사용되었다. 그 중에서도 Hotelling's T^2 통계량을 이용한 T^2 관리도가 널리 사용된다(Montgomery, 1996). Hotelling's T^2 통계량은 아래 식 (6)에서 보이는바와 같이 계산할 수 있다.

$$T^2 = (X - \bar{X})^T S^{-1} (X - \bar{X}) \quad (6)$$

여기서 \bar{X} 와 S 는 관측 값들로부터 계산한 표본의 평균과 분산을 의미한다. 중심으로부터 얼마나 떨어져 있는 지를 나타내는 $X - \bar{X}$ 와 이를 각 변수의 공분산 행렬의 역행렬을 곱하여 각 변수의 영향을 보정하여 조금 더 설명력이 높은 유사도를 계산하게 하였다.

$$UCL = \frac{p(n+1)(n-1)}{n^2 - np} F_{\alpha, p, n-p} \quad (7)$$

위의 식 (7)과 같이 T^2 관리도의 관리 한계선을 계산할 수 있다. p 는 다변량 데이터의 변수의 개수이고 n 은 총 관측 값의 수이다. T^2 값은 항상 양의 값이기 때문에 관리 하한선은 언제나 0이 된다. 식 (7)의 관리 상한선보다 통계량이 작다면 해당 관측치는 이상으로 분류된다.

(2) k-Nearest Neighbor Data Description

kNNDD(k-Nearest Neighbor Data Description)은 kNN(k-nearest neighbor) 방법을 기반으로 k 번째 nearest neighbor와의 거리를 사용한다(Wang *et al.*, 2004). kNNDD는 kNN을 이용하여 데이터의 local density를 추정함으로써 단일 클래스 분류 문제에 적용된다(Breunig *et al.*, 2000; Tax, 2001). kNN을 통하여 데이터 x 에 대한 k 개의 nearest neighbor를 찾고 이를 $N_i(x)$ 라 한다. 즉, $N_i(x)$ 는 i 번째 nearest neighbor를 의미한다. 각각의 $N_i(x)$ 에 대한 k 개씩의 nearest neighbor를 $N_j(N_i(x))$ 라고 할 때, $N_i(x)$ 에 대한 x 의 local density는 식 (8)과 같고, $N_j(N_i(x))$ 에 대한 $N_i(x)$ 의 local density는 식 (9)과 같다.

$$d(x) = \frac{k/N}{V \|x - N_i(x)\|} \quad (8)$$

$$d(N_i(x)) = \frac{k/N}{V \|N_i(x) - N_j(N_i(x))\|} \quad (9)$$

위의 식에서 N 은 전체 데이터의 개수를 의미한다. kNNDD는 x 를 분류할 때 $N_i(x)$ 에 대한 x 의 local density와 $N_j(N_i(x))$ 에 대한 $N_i(x)$ 의 local density의 비율을 이용하는데 식 (10)과 같다.

$$\frac{d(x)}{\sum_{i=1}^k d(N_i(x))} = \frac{\sum_{j=1}^k \|N_i(x) - N_j(N_i(x))\|}{\sum_{i=1}^k \|x - N_i(x)\|} \quad (10)$$

<Figure 1>는 k 가 2일 때 kNNDD의 예제를 나타낸다. 이때 k 의 크기는 kNNDD의 성능에 영향을 미치는데(Sukchotrat *et al.*, 2010) k 의 수의 적절한 범위는 10개에서 50개 사이라고 연구되어졌다(Breunig *et al.*, 2000). 이는 분류를 분석하는 데이터에 따라 다르며 해당 k 값은 training data를 분석하며 정하는 것이 일반적이다.

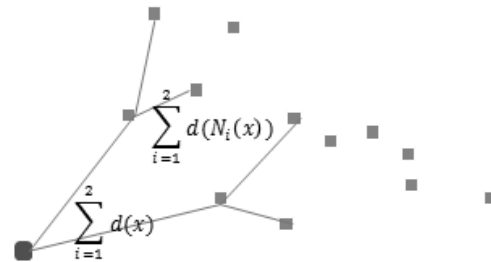


Figure 1. Example of kNNDD(k = 2)

(3) Distance based Novelty Detection

Novelty detection은 정상적인 데이터의 분포를 학습하여 주어진 패턴을 분류, 판단하는 기법이다(Kang and Cho, 2008). Novelty detection은 크게 분포추정 기법, 단일 클래스 분류 기법, 거리 기반 기법 등으로 연구 되어져 왔다.

분포추정 기법은 학습 데이터의 특정한 분포와 그 모수를 찾고 확률이 낮은 데이터를 이상치로 판단하는 기법이다. 단일 클래스 분류 기법은 정상 데이터만을 이용해서 학습을 수행하여 정상 영역에 경계면을 생성해 이상치를 분류하는 기법이다. 대표적으로 One-class SVM이 있다(Schölkopf *et al.*, 2001).

거리기반의 Novelty detection 방법들은 kNN을 이용한다. kNN을 이용하여 한 집단으로 이루어진 학습된 데이터 중 k 개를 뽑아서 그들의 거리기반의 novelty score를 계산하여 집단에 속하는지 여부를 결정한다. novelty score란 거리기반으로 새로운 관측치와 이웃 패턴 간의 거리를 비율적으로 계산한 분류 척도이다. 정상 데이터 간에 즉, 새로운 관측치와 이웃 패턴간의 거리가 멀수록 이상치로 분류할 확률을 높게 추정하는 것이다(Kang and Cho, 2009). 거리기반 novelty detection 방법은 앞서 설

명한 분포추정 기법과 one-class classification 기법에 비해 더 정확히 정상데이터의 영역을 추정할 수 있고, 데이터의 추가에 따라 재학습하지 않아도 된다는 장점이 있다(Kang and Cho, 2008).

<Figure 2>는 k가 5일 때 Average Distance의 예제를 나타낸다. 이 기법은 Knorr *et al.*(2000)에 의해 처음 제안되었고, 변형된 여러 가지 종류의 거리기반 novelty score가 제안되었다.

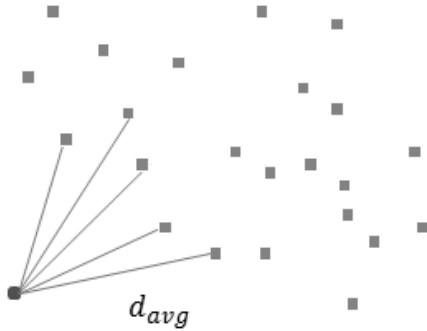


Figure 2. Example of Average Distance(k = 5)

식 (11)은 Average distance(Angiulli and Pizzuti, 2005)로 계산한 novelty score 식이다.

$$d_{avg} = \frac{1}{k} \sum_{i=1}^k \|X_n - X_n^i\| \quad (11)$$

Average distance는 위의 식 (11)에서 볼 수 있듯이 k개의 nearest neighbor들의 거리의 평균값을 novelty score로 정의했다. 이는 k개의 neighbor의 정보를 모두 사용한다는 장점을 가진다.

2.2 주성분 분석을 이용한 Feature Selection

주성분 분석(Principal Component Analysis)은 다변량 분석에서 각 차원 간에 상관성이 있는 데이터를 다룬다. 고차원 정보를 유지하면서 저차원으로 차원을 축소하는 다변량 데이터 처리 방법 중 하나이다. 이는 데이터에서 중요한 정보를 추출하는 것을 목적으로 하며 데이터를 한 개의 축으로 사상(projection)시켰을 때 그 분산이 커지는 축을 순서대로 놓이도록 데이터를 선형 변환하여 차원을 축소시킬 수 있다(Han, 2009).

$$\mu = \frac{1}{N} \sum_{n=1}^N X_n \quad (12)$$

$$\Sigma = \frac{1}{N-1} \sum_{n=1}^N (x_n - \mu)(x_n - \mu)^T \quad (13)$$

위의 식 (12)과 식 (13)를 통해 구한 공분산 행렬을 기저 벡터와 기저 값으로 분해한다. 식 (14)를 통해 고유분해를 행한다. 주성분 분석은 정해진 기저 벡터를 갖지 않으며, 기저 벡터는 데이터의 특성에 따라 달라진다.

$$\Sigma = UAU^T \quad (14)$$

선택된 기저 값과 관련된 기저벡터를 구하고, 연결하여 변환행렬 A를 식 (15)과 같이 만든다.

$$A = [u_1, u_2, \dots, u_n] \quad (15)$$

특징 벡터를 다음 식 (16) 변환 식으로 변환한다.

$$Y = A^T \cdot X \quad (16)$$

X는 원래의 데이터 행렬이고 Y는 score 행렬이다. Y는 n by p 행렬로 n은 데이터 수를 의미하고 p는 PC의 수를 뜻한다. 여기서 데이터의 정보를 잘 반영하는 PC를 뽑아 차원을 축소시킨다. 그리고 새로운 관측 값들을 식 (16)의 X에 대입하여 같은 방법으로 score 행렬을 계산할 수 있다.

본 연구에서는 주기길이가 100인 고차원 다변량 데이터를 사용함으로 주성분 분석을 이용하여 이상분류를 위한 단일 클래스 변수 알고리즘의 효율을 높이고자 한다.

2.3 단일 클래스 분류 기반 이상 분류 기법

본 연구에서는 반도체 공정에서 이상이 발생했을 때 어떤 종류의 이상인지를 분류하는 방법을 제안한다. 이상을 잘 분류하는지에 대한 분류 정확도가 중요한 동시에 이상의 원인을 파악하기 위한 시간을 줄이는 것도 중요하다. 이상의 원인을 파악하기 위해서 반도체 공정의 엔지니어는 검사를 실행하는데 평소 이상의 비율이 가장 컸던 종류순으로 검사를 실행한다. 검사를 실행하는데 시간과 비용이 소요되기 때문에 이상의 원인을 잘 분류해 내어 검사 횟수를 감소시키는 것이 중요하다. 검사 횟수를 줄이기 위해 가장 가능성이 높은 이상 클래스로부터 적은 것 까지 순서를 나열한다면 검사 횟수를 줄일 수 있을 것이라 기대된다.

본 연구에서는 단일 클래스 분류기법 기반의 이상 분류를 제안한다. 각각의 이상 클래스를 기반으로 단일 클래스 분류 모델을 만들어 각각의 모델의 Novelty score가 큰 순서대로 정렬하여 클래스에 속할 가능성에 따라 정렬하도록 하였다.

Hotelling's T^2 통계량을 이용하여 분류 모델을 설정하였다. 위의 제 2.2절에서 언급한바와 같이 Hotelling's T^2 값은 다변량에서 변수의 상관관계를 고려한 거리이다. 각각의 이상 클래스의 데이터를 이용하여 새로운 관측 치에 대한 거리를 T^2 값을 식 (7)과 같이 계산한다. 이때 T^2 값이 클수록 거리가 멀리 떨어져 있다는 것을 의미하고 작을수록 거리가 가깝고 클래스에 속할 확률이 높아지므로 T^2 값이 작은 순으로 검사 순서를 설정한다.

또한 kNNDD를 이용하여 T^2 를 이용한 방법과 같이 새로운

관측 치를 기준으로 각각의 클래스의 데이터와 가장 근접한 k 개의 $N_i(x)$ 와 각각의 $N_i(x)$ 에 대한 k 의 $N_j(N_i(x))$ 비율을 식 (11)과 같이 계산할 수 있다. 각 클래스에 대해 계산한 관측 치의 비율이 낮을수록 그 클래스에 속할 가능성이 크다는 것을 의미하므로 비율이 낮은 순으로 검사 순서를 설정한다. 위의 <Figure 3>는 class가 5개이고 k 가 2일 때의 경우이다. 여기에서 가장 근접한 2개의 $N_i(x)$ 와 각각의 $N_i(x)$ 에 대한 2의 $N_j(N_i(x))$ 비율 가장 작은 클래스 1, 4, 5, 2, 3순으로 가능성을 이용한다.

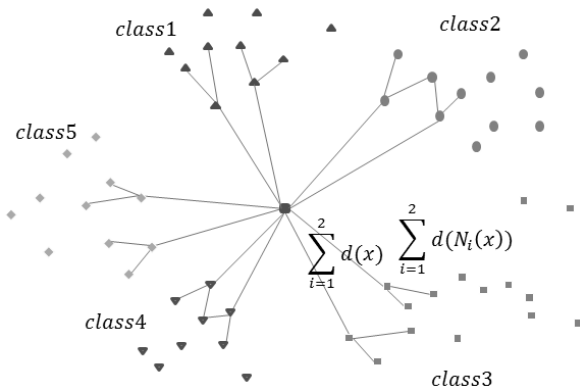


Figure 3. k-NNDD(k = 2)

마지막은 Distance based Novelty detection을 이상 분류에 적용시키는 방법이다. kNN을 이용하여 각각 이상 클래스의 관측치로부터 근접한 k 개의 nearest neighbor를 뽑아서 그들의 average distance 기반의 novelty score를 계산하여 novelty score가 낮은 순으로 검사 순서를 설정한다. 위의 <Figure 4>는 class가 5개이고 k 가 3일 때의 예제이다. 여기에서 거리가 가장 짧은 클래스 1, 4, 5, 3, 2순으로 가능성을 이용한다.

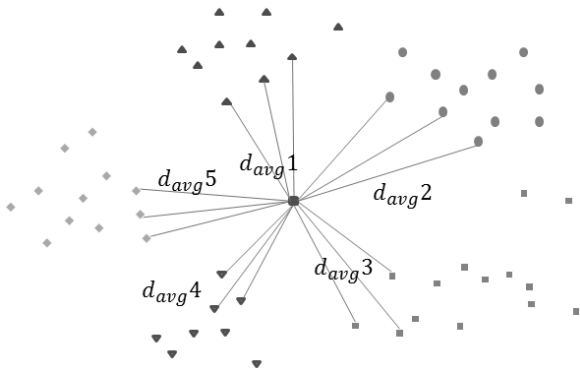


Figure 4. Distance based novelty detection(k = 3)

본 연구는 주성분분석을 통해서 신호의 특성을 추출하고 Hotelling's T^2 , kNNDD, Distance based Novelty detection을 통해서 novelty score을 계산하여 검사 순서를 배열하여 검사 횟수를 줄임을 목적으로 한다. 즉, 단일 분류 클래스 기법들을 적용하

여 통계적 거리가 가까운 순서로 검사순서를 배열하고, 그 순서대로 검사를 진행함으로써 검사 횟수를 감소시킬 수 있다. 검사 횟수 감소는 이상 원인을 찾아 개선하고자 공정을 검사하는 시간을 감소시킬 수 있고, 검사 비용을 감소시킬 수 있는 효과가 있다. 제안한 방법들은 기존의 다분류 문제 분류기 보다 분류 정확도는 더 높지 않을 수 있지만 검사 횟수는 더 낮을 것으로 예상된다.

3. 실험 및 결과 분석

본 연구에서 제안하는 이상분류의 성능을 평가하기 위해 실제 반도체 공정에서 발생하는 주기 신호의 형태를 모방한 데이터를 생성하여 실험을 수행하였다. 실제 현업에서 사용하는 데이터는 기업의 보안상의 이유로 본 연구는 실제 데이터를 가지고 실험할 수 없었다. 이에 본 연구에서는 반도체 제조 공정 중에서 가장 많은 장비가 필요하며 긴 시간이 소요되는 FAB 공정에서 발생하는 주기신호를 본떠 데이터를 만들었다. 그 중에서도 가장 대표적인 3종류 주기 신호의 특성을 반영하여 실험 신호 데이터를 생성하여 실험을 수행하였다. 본 실험에서 생성한 데이터는 Park et al.(2011), Han et al.(2010)이 이전 연구에서 생성한 데이터를 참고하여 생성 하였다.

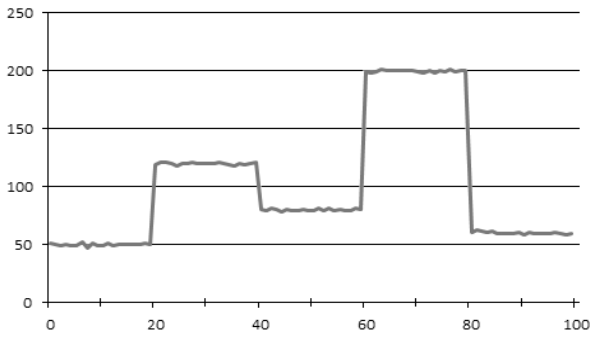
3.1 설계

본 실험에서 사용한 주기신호 데이터 중 첫 번째 step 신호(Type1)는 주로 광전자 센서에서 측정되는 전압 데이터를 모방한 것이다. 두 번째 Steep 신호(Type2)는 Diffusion 공정의 온도 센서에서 측정되는 데이터이며, 마지막 신호(Type3)는 Smooth 신호로 Photo-lithography 공정의 온도 데이터이다.

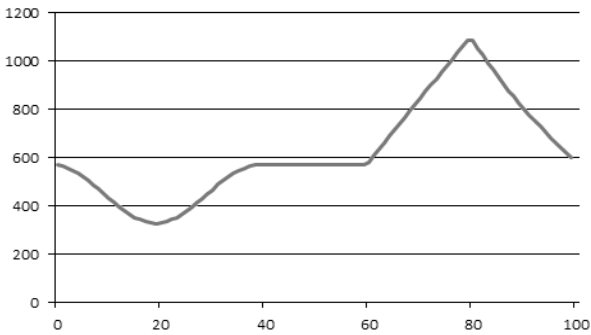
실제 반도체 공정에서 발생하는 신호는 원래의 신호에 화이트 노이즈(white noise)가 포함된다. 실제로 반도체 공정에서 발생하는 주기신호의 특성을 반영하기 위하여 길이 100의 주기신호 데이터에 평균이 0이고 분산이 1인 정규분포를 따르는 노이즈를 생성하여 더하였다. 생성한 세 종류의 주기신호는 <Figure 5>와 같은 형태를 가진다.

한 주기 신호는 5개의 step으로 구성되며 본 실험은 구성된 신호의 step에 shift나 drift값이 더해진 이상 신호만을 이용해 실험을 수행한다. 이상 신호의 패턴 그룹은 <Table 1>과 <Table 2>에 언급한 5개의 종류이다. <Table 1>은 Type1과 Type3 신호의 이상그룹정보이고 <Table 2>는 Type2 신호의 이상 그룹 정보이다.

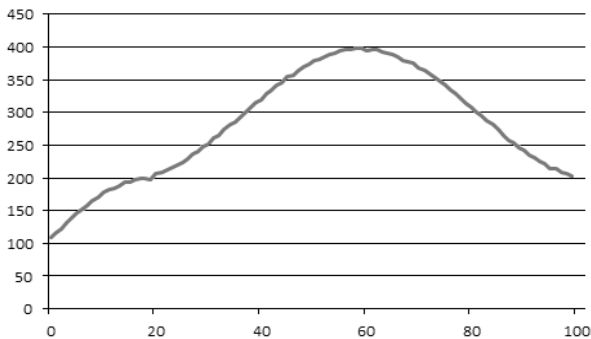
Type1과 Type3 신호에 비해 Type2 신호는 현저하게 큰 평균의 차이를 보정하기 위해서 위와는 각각 다른 shift 값과 drift 값이 주어진다. Type1과 Type3의 shift 값은 0.4이고 Type2의 shift는 0.6이다. Type1과 Type3의 drift의 기울기는 0.007이고 Type2의 drift 기울기는 0.01이다. 첫 번째 이상 클래스는 upshift가 발생



(a) Type1 : Step Signal



(b) Type2 : Steep Signal



(c) Type3 : Smooth Signal

Figure 5. Characteristic of the three type of Signal

되는 이상 데이터이며 step 2와 3에서 shift 값이 존재한다. 두 번째 이상 클래스는 step 2, 3, 5에 대해서 downshift가 발생한다. 세 번째 이상 클래스는 step 4, 5에 대해 updrift가 발생하며, 네 번째 이상 클래스는 step 3, 4, 5에 대해서 downdrift가 발생한다. 다섯 번째 이상그룹은 전 step에 대해서 white noise의 표준 편차가 1.2배인 경우이다.

각각의 데이터 Type 별로 1000개의 주기신호를 생성하고 이상패턴 종류의 비율을 <Table 3>과 같이 각각 두 가지 비율로 발생했다.

본 연구에서는 3가지의 Fault classification 방법을 제안하였고 one-against-one을 이용한 다분류 SVM과 비교 분석을 진행하였다. 주성분분석 기법에 의해 추출된 특질에 대해서 위에서 제

Table 1. Description of the Error for Type 1, Type 3 signal

Fault Class	Description
1	Upshift(+0.4) in step 2, 3
2	Downshift(-0.4) in step 2, 3, 5
3	Updrift(+0.007) in step 4, 5
4	Downdrift(-0.007) in step 3, 4, 5
5	High variance(1.2 times) in all step

Table 2. Description of the Error for Type 2

Fault Class	Description
1	Upshift(+0.6) in step1, 4
2	Downshift(-0.6) in step 2, 3, 5
3	Updrift(+0.01) in step 1, 3
4	Downdrift(-0.01) in step 2, 3, 4
5	High variance(1.2 times) in all step

Table 3. Fault Mix Rate

Fault Class	Type1/Type2/Type3	
	Mix1(%)	Mix2(%)
1	20	60
2	20	20
3	20	10
4	20	5
5	20	5

안한 3가지 단일 클래스 분류 기반 이상 분류 기법(Hotelling's T^2 , kNNDD, Distance based novelty detection)을 가지고 클래스의 검사순서를 나열하여 분류 정확도와 검사횟수를 계산한다. 즉, 예를 들어 본래 클래스가 2인 데이터가 3-2-1-4-5의 순서대로 검사를 진행한다는 결과가 나온다면 검사 횟수는 2가 된다.

주성분분석 기법에 의해 추출된 특질을 이용하여 다분류 One-against-one 기반의 SVM이 대조군으로 이용되었다(Hsu, and Lin. 2002). 커널 함수는 대표적으로 많이 쓰이는 Gaussian 커널 함수와 Polynomial 커널 함수 중 이 타입의 문제에서 Polynomial 커널 함수의 성능이 더 좋았기 때문에 Polynomial 커널을 사용하였다. 이 때 검사 횟수는 먼저 분류된 클래스의 검사 우선순위를 제일 앞으로 두고 나머지 클래스를 빈도순으로 배열하여 계산한다. 예를 들어 SVM에 의해 클래스 2라고 분류한 신호가 있다고 가정하자. 실제 이 신호가 실제로 클래스 3이었고 클래스 별 빈도가 같은 <Table 3>의 Mix 1과 같다고 할 때 검사순서는 2-1-3-4-5로 나열할 수 있다. 또 이때의 검사 횟수는 3이 된다. 이렇게 계산한 평균값을 구한 것이 <Table 4>의 검사 횟수이다.

각 이상 클래스의 비율을 알고 있기에 이를 이용하면 분류 정확도와 검사 횟수의 하한을 계산할 수 있는데 이는 <Table 4>에서의 Lower bound로 나타난다.

3.2 결과

전체 5가지 종류의 이상 패턴에 따라 생성한 데이터에 대해 5-fold Cross validation으로 이상 분류 방법의 성능을 평가한 결과가 <Table 4>와 같다. 전체적으로 볼 때, Distance based novelty detection(PC-Averagedist)이 가장 우수한 성능을 보이고 있다. 분류 정확도면에서는 SVM이 좋은 성능을 보이는 경우도 있었지만 검사 횟수 측면에서는 모든 경우 Distance based novelty detection이 가장 우수한 성능을 보였다.

Distance based novelty detection을 제외하고 본 연구에서 제안한 나머지 방법들인 Hotelling's T^2 와 kNNDD는 모든 실험 종류에 대해서 SVM보다 분류 정확도가 떨어졌다. 그러나 Hotelling's T^2 는 SVM보다 모든 데이터에서 검사 횟수가 적었으며, kNNDD 역시 6개의 데이터 중 4개의 데이터에서 SVM보다 검사 횟수가 적은 결과를 보였다.

결과적으로 본 연구에서 제안한 방법은 SVM과 유사한 분류 정확도를 보장하는 동시에 이상 원인을 파악하는 시간을 의미하는 검사 횟수에서 전체적으로 더 좋은 결과를 보였다.

검사 횟수 면에서 가장 큰 성능 차이를 보이는 Type1, Mix1 데이터를 보면 Distance based novelty detection과 SVM의 검사 횟수 차이가 평균 0.432만큼 나는 것을 볼 수 있다. 이는 공정에서 1000개의 이상이 탐지 되었다고 가정 했을 때, Distance based novelty detection는 SVM보다 432번의 검사 횟수를 줄여 공정 개선을 위한 시간을 단축시키고 비용을 감소시킬 수 있는 효과이다.

4. 결론

본 논문에서는 단일 클래스 분류기법을 활용하여 다중 이상

분류 방법을 제시하였다. 주성분분석을 이용해 추출된 특징 데이터에 Hotelling's T^2 , kNNDD 및 Distance based novelty detection의 단일 클래스 분류 기법을 적용하였다. 각 이상 클래스별 분류 모델을 생성하여 모델의 novelty score를 이용해 검사 순서를 배열하는 방법은 SVM과 유사한 분류성능을 나타낼 뿐만 아니라 적은 검사 횟수를 가짐을 확인할 수 있었다.

검사 횟수의 감소는 이상이 탐지 되었을 때 이상의 원인을 찾아 개선하고자 해당 공정을 검사하는 시간을 감소시킬 수 있다. 본 논문에서는 검사순서를 제안함으로써 오분류되는 경우에도 적합한 원인을 찾는 데 까지 소요되는 시간을 줄일 수 있다. 즉, 신속한 이상 원인의 진단을 위해서 검사시간을 줄일 수 있고, 검사 비용을 감소시킬 수 있는 효과가 있다.

본 논문에서 제안한 방법 중 Distance based novelty detection에서 novelty score로 사용한 average distance는 k개의 nearest neighbor의 관측치의 척도로 사용하였다. 하지만 nearest neighbor의 패턴모양에 따라 구분을 하지 못하는 문제가 발생한다(Harmeling et al., 2006). 추후 연구에서 이점을 고려한다면 더욱 향상된 결과를 기대할 수 있을 것이다. 또한 인공지능영역을 이용하여 단일 클래스 분류 기법을 통합한 형태의 다중 클래스 분류기법을 적용하는 것은 검사 횟수를 줄이는 동시에 분류의 정확도를 높일 것이라 기대된다.

참고문헌

Ahn, H. C. and Kim, K. J. (2009), Corporate Bond Rating Using Various Multiclass Support Vector Machines, *Asia pacific journal of information systems*, 19(2), 157-178.
 Angiulli, F. and Pizzuti, C. (2005), Outlier mining in large high-dimensional data sets, *IEEE Transactions on Knowledge and Data Engineering*, 17(2), 203-215.

Table 4. Accuracy of Fault Classification and the Number of Test

		Mix1(0.2/0.2/0.2/0.2/0.2)		Mix2(0.6/0.2/0.1/0.05/0.05)	
		Classification accuracy	검사 횟수	Classification accuracy	검사 횟수
Type1	PC-SVM	0.7330	1.7320	0.8930	1.3190
	PC- T^2	0.7290	1.3130	0.7740	1.2540
	PC-kNNDD	0.7150	1.3310	0.6890	1.3440
	PC-Averagedist	0.7380	1.3000	0.8820	1.1510
Type2	PC-SVM	0.9570	1.1200	0.9750	1.0780
	PC- T^2	0.9470	1.0560	0.9360	1.0670
	PC-kNNDD	0.9520	1.0510	0.9410	1.0640
	PC-Averagedist	0.9600	1.0400	0.9760	1.0280
Type3	PC-SVM	0.9020	1.2550	0.9500	1.1650
	PC- T^2	0.8960	1.1120	0.8600	1.1480
	PC-kNNDD	0.8770	1.1300	0.8240	1.1820
	PC-Averagedist	0.8980	1.1080	0.9460	1.0700
Lower Bound		0.2	3	0.6	1.75

- Breunig, M. M., Kriegel, H. P., Ng, R. T., and Sander, J. (2000) LOF : identifying density-based local outliers, in *Proceedings of the ACM SIGMOD 2000 International Conference on Management of Data*, 93-104.
- Chiang, L. H., Russell, E. L., and Braatz, R. D. (2001), Fault detection and diagnosis in industrial systems, *Springer-Verlag*.
- Goodlin, B. E., Boning, D. S., Sawin, H. H., and Wise, B. M. (2003), Simultaneous Fault Detection and Classification for Semiconductor Manufacturing Tools, *Journal of Electrochemical Society*, 150(12), 778-784.
- Han, A. H., Park, C. S., Kim, S. S., and Baek J. G. (2010), A Method to Adjust Cyclic Signal Length Using Time Invariant Feature Point Extraction and Matching (TIFEM), *Journal of the Korea Society for Simulation*, 19(4), 111-122.
- Han, H. Y. (2009), Theory of pattern recognition, *Hanbit media*, 247-250.
- Harmeling, S., Dornhege, G., Tax, D., Meinecke, F., and Muller, K.-R. (2006), From outliers to prototypes: ordering data, *Neurocomputing*, 69(13~15), 1608-1618.
- Hsu, C.-W. and Lin, C.-J. (2002), A comparison of methods for multi-class support vector machines, *IEEE Transactions on Neural Networks*, 13, 415-425.
- Kang, S. P. and Cho, S. J. (2008), Locally Linear Reconstruction for Novelty detection, *IE/MS*.
- Kang, P. and Cho, S. (2009), A hybrid novelty score and its use in keystroke dynamics-based user authentication, *Pattern Recognition*, 42, 3115-3127.
- Knorr, E. M., Ng, R. T., and Tucakov, V. (2000), Distance-based outliers: Algorithms and applications, *The VLDB Journal*, 8(3/4), 237-253.
- Kymal, C. and Patiyasevi, P. (2006), Semiconductor Quality Initiatives : How to maintain quality in this fast-changing industry, *Quality Digest*, 26(4), 43-48.
- Lee, J. H., Kim, J. H., Hwang, J. B., Kim, S. S. (2007), A Study on Fault Detection of Cycle-based Signals using Wavelet Transform, *Journal of the Korea Society for Simulation*, 16(4), 13-22.
- Montgomery, D. C. (1996), Introduction to statistical quality control, 3th edition, Johan Wiley and Sons, Inc.
- Park, J. H., Kwon, I. H., Kim, S. S., and Baek, J. G. (2011), Spline regression based feature extraction for semiconductor process fault detection using support vector machine, *Expert Systems with Applications*, 38(5), 5711-5718.
- Park, S. H., Kim, J. S., Kim, S. S., Park, C. S., and Baek, J. G. (2010), A Fault Detection of Cyclic Signals Using Support Vector Machine-Regression, *KSQM*, 354-362.
- Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., and Williamson, R. C. (2001), Estimating the support of a high-dimensional distribution, *Neural Computation*, 13(7), 1443-1471.
- Sukchotrat, T., Kim, S. B., and Tstmg, F. (2010), One-Class Classification-based Control Charts for Multivariate Process Monitoring, *IIE Transactions*, 42, 107-120.
- Tax, D. M. J. (2001), One-class classification: concept - learning in the absence of counter-examples, *PhD thesis, Delf University of Technology, The Netherlands*.
- Wang, Q., Lopes, L. S., and Tax, D. M. (2004), Visual Object Recognition Through One-Class Learning, *Lecture notes in computer science*, 3211, 463-470.



조민영

건국대학교 산업공학과 학사
 현재 : 고려대학교 산업공학과 석사과정
 관심분야 : Advanced process control, Data mining applications, FDC



백준길

고려대학교 산업공학과 학사
 고려대학교 산업공학과 석사
 고려대학교 산업공학과 박사
 현재 : 고려대학교 산업경영공학과 부교수
 관심분야 : 첨단공정제어, 지능형 이상 진단, 데이터마이닝 응용