

## A Study on Data Mining Application Problem in the TFT-LCD Industry

Hyunwoo Lee<sup>1)</sup> · Ho Soo Nam<sup>2)</sup> · Jung-chul Kang<sup>3)</sup>

### Abstract

This paper deals the TFT-LCD process and quality, process control problems of the process. For improvement of the process quality and yield, we apply a data mining technique to the LCD industry. And some unique quality features of the LCD process are also described.

We describe some preceding researches first and relate to the TFT-LCD process and the problems of data mining in the process. Also we tried to observe the problems which need to solve first and the features from description below hazard must be considered a quality mining in LCD industry.

**Keywords** : 데이터마이닝, 불량률, 수율, CRM, TFT-LCD

### 1. 서론

21세기 들어 컴퓨터의 발달은 정보체계의 발전과 더불어 일상생활에 너무나 많은 변화를 불러일으키고 있는 게 사실이다. 우리나라 어떤 지역에 어떤 사건이 발생하였는가를 실시간으로 알아볼 수 있을 뿐만 아니라 전 세계 거의 모든 지역에서 일어나는 중요한 사건들도 거의 실시간으로 알아볼 수 있다. 예전에는 중요한 자료를 찾기 위해서는 대형 백과사전을 찾거나 큰 도서관에 가서 혹은 외국 도서관에 의뢰를 해서 정보를 찾아야만 했다. 그러나 지금은 그러한 것들을 인터넷 이라는 도구를 이용해서 훨씬 많은 정보를 자신의 책상에 앉아 아주 짧은 시간에 조사가 가능해졌다.

산업 전반에도 많은 변화가 일어났다. 집안에서 시장을 볼 수 있고, TV를 보면서 전화로 혹은 인터넷으로 옷을 구매할 수 있다. 회사에서는 이런 고객들의 성향을 파악해서 고객이 원하는 제품을 추천하고, 고객의 구매 패턴을 분석해서 마케팅에 활용

- 
- 1) 제1저자 : 서울시 송파구 오금동 33-10 (주)한국신뢰성기술서비스, 부장  
E-mail : hyunwoo@korts.co.kr
  - 2) 부산시 사상구 주례2동 산 69-1, 동서대학교 시스템경영공학과, 부교수
  - 3) 부산시 부산진구 엄광로 995 동의대학교 상경대학 금융보험학과, 교수

한다. 이러한 활동을 하기 위해서는 수많은 고객들에 대한 정보와 그 고객들의 행동 하나하나를 데이터화 하고 분석하여야 한다. 그러나 이러한 분석은 불과 10년 전만 하더라도 그리 쉬운 일은 아니었다. 결국 컴퓨터의 발달이 이런 일을 가능하게 만들어 준 것이다.

이러한 활동은 제조업에서도 마찬가지로 일어나고 있다. 전 공정상에 상황을 자동으로 검사하고 이들 데이터를 DB화해서 실시간으로 전 공정의 상태를 파악할 수 있게 되었다. 또한 제품의 고장이나 결점의 상태를 파악하고 이들의 패턴을 감지함으로써 무결점의 완벽한 제품을 만들기 위한 노력들을 많은 시간을 들여 하고 있다. 고장이나 각종 결점에 대한 데이터를 통계분석하고 분석결과를 이용해서 결점의 원인을 해결하고, 다시 이러한 결점이 일어나지 않도록 예방보전을 해주는 과정을 벌이고 있다. 이런 방대한 데이터를 통계처리 할 수 있는 기법이 데이터마이닝이다.

데이터마이닝의 이론이 최근 들어 새롭게 나온 것은 아니다. 이미 1960년 초부터 데이터마이닝의 이론이 등장했으나 대량데이터의 정의와 이를 처리할 수 있는 컴퓨터의 용량 문제로 이를 제대로 활용할 수 없었으나, 최근 급속도로 발전하고 있는 IT의 발전이 이를 가능하게 하였다. 또한 약 10여 년 전부터 급성장한 CRM은 데이터마이닝의 확산에 크게 기여하였다. 반면에 제조 공정에서의 데이터마이닝의 적용은 CRM에서의 데이터마이닝 보다 현실적으로 많은 어려움 때문에 그다지 활성화 되어있지 않은 게 사실이다. 그러나 이러한 어려움에도 불구하고 무결점의 최상의 제품을 만들기 위해서는 반드시 필요한 과정이 데이터마이닝이며, 이를 효율적으로 활용하기 위해서는 데이터마이닝 시스템의 구축이 필수불가결하다고 할 수 있을 것이다.

본 논문에서는 LCD분야에서 공정을 효율적으로 운영하기 위한 데이터마이닝을 적용함에 있어서 발생하는 문제점과 그 해결책을 제시하고, LCD산업에서의 데이터마이닝 적용 사례를 소개하기로 한다.

## 2. 선행연구

정보기술의 빠른 발전은 업무의 자동화를 촉진시켜 방대한 양의 데이터를 전자적으로 수집하고 보관하는 것을 가능하게 하였다. 이는 CRM 뿐만 아니라 제조업체의 품질관련 분야에서도 동일한 현상을 보이고 있다. 데이터의 양이 오늘날과 같이 방대하지 않았던 과거에는 소수의 전문가들이 통계기법이나 질의(Query)를 통해 데이터를 분석하고 요약된 결과를 보고서 형식으로 제공하였다. 그러나 데이터양이 기하급수적으로 증대함에 따라 이전과 같은 수작업에 의한 데이터 분석방법으로는 도저히 불가능에 가까운 일이라고 할 수 있다.(장남식 등, 1999). 따라서 방대한 양의 데이터로부터 유용한 지식을 얻도록 사람을 지능적이고 자동적으로 지원하는 새로운 기법과 도구가 필요하게 되었다.(Fayyad 등, 1996).

백준걸 등(2000)은 반도체 에칭(etching) 공정에서 실시간으로 수집되는 기계이력 데이터의 분석을 통해 현재의 기계 상태를 정확히 반영한 기계 고장 정보를 추출해 낼 수 있는 점진적 의사결정나무 구축을 제시하였다. 에칭 율에 영향을 미치는 공정 변수(기계 내의 압력, 온도, 에칭가스 밀도, 웨이퍼 위치 등)를 속성으로 정의하고 에칭 율에 따라 결과 클래스(LE : Low Etching rate, NE : Normal Etching rate, HE : High Etching rate)를 정의한 이력 데이터를 기반으로 에칭율의 이상 상황을 예측할

수 있는 의사결정나무를 작성하여 활용하였다.

안진석 등(1999)은 자동차 휠 주조공정에서 범주형 데이터와 연속형 데이터의 혼합 형태로 공정변수가 주어지며 품질 특성치로서 불량률 데이터가 주어진 경우에 최적 공정조건에 대한 정보 추출을 위해서 연속형 변수의 범주화와 유사빈도의 군집화 기법을 제시하였다.

Braha 등(2002)은 반도체 클린징 공정(cleansing process)에서 미세 오염물질 제거를 위해 레이저 빔을 이용하는 advanced wafer cleansing이라 불리는 건조 클린징 기술(dry cleansing technology)의 개선을 위해 데이터마이닝 기법을 적용하였다.

Kusiak 등(2001)은 rough set theory를 이용하여 printed-circuit board(PCB) 조립 공정 중 일부 구성 요소 밑에서 발생하는 불량률의 원인을 찾고자 하였다. 기존의 통계적 방법이나 실험계획적 방법으로는 불량률의 원인을 명확히 파악하지 못하였다. 이 논문에서는 stencil 배합, stencil 두께, 위치, 진공여부 등 14개의 속성을 이용하여 불량률의 원인을 찾는 방법을 제안하였다.

Lian 등(2002)은 판금(sheet metal) 제품의 규격 변동(dimensional variation)의 원인을 진단하기 위해 데이터마이닝 기법과 지식발견기법을 적용한 의사결정 방법을 제안하였다. 자동차 차체의 규격 정확도는 이후 공정에 큰 영향을 미치므로 규격변동의 원인을 신속, 정확하게 파악하는 것은 매우 중요하다. 이 논문은 우선 방대한 다차원 측정 데이터로부터 변동 폭이 큰 그룹을 추출하기 위해 상관분석과 최대나무(maximal tree) 기법을 적용하였으며, 주된 변동 패턴을 발견하기 위해 다변량 통계분석을 이용하였다. 또한 변동 원인의 진단을 위해 제품과 조립공정 지식을 기반으로 한 의사결정나무 접근을 하였다.

백동현 등(2003)은 반도체 FAB공정의 복잡성과 생산현장에서 수집되는 방대한 기술데이터로 인해 기존의 통계적 방법이나 엔지니어의 경험적 분석 방법으로는 미처 파악하지 못하는 수율저하 문제를 FAB 공정을 마친 웨이퍼에 불량 칩(chip)이 지리적으로 특정 위치에 집중적으로 발생하는 현상을 육안검사 대신 군집분석을 이용하여 데이터로부터 자동 판별할 수 있는 방법을 제시하고, 연속패턴분석, 분류분석, RBF(Radial Base Function) 기법을 적용하여 수율저하의 원인이 되는 문제 장비나 문제 파라미터를 신속, 정확하게 파악할 수 있도록 해 줄 뿐만 아니라 공정진행 중인 제품의 미래 수율을 예측할 수 있도록 지원하는 방법을 제안하였다.

### 3. LCD 제조공정

LCD공정에서의 데이터마이닝 적용을 설명하기 위하여 우선 LCD공정이 얼마나 복잡하고 따라서 데이터를 수집하는 과정에서 어떤 문제점을 내포하고 있는지를 이해하기 위해서 우선 LCD공정에 대한 이해가 선행 되어야 한다. 그래서 먼저 LCD제조 공정에 대한 설명을 하고, 데이터마이닝 적용방법에 대해서 설명하기로 한다.

TFT-LCD는 초박막 액정표시장치(Thin Film Transistor Liquid Crystal Display)의 약자로, 액체와 고체의 중간 특성을 가진 액정의 상태 변화와 편광판의 편광성질을 이용하여 통과하는 빛의 양을 조절함으로써 정보를 표시하는 첨단 디지털 디스플레이이다. TFT-LCD는 Color Filter와 TFT가 형성된 두 장의 유리 기판과 그 사이에 주입된 액정(Liquid Crystal), 그리고 광원인 Back Light Unit로 구성되어 있다. TFT

는 액정을 제어하기 위해 초박형 유리기판 위에 반도체 막을 형성한 회로이며, 이 회로는 영상의 기본 단위인 화소(pixel)를 하나하나 제어하는 역할을 한다. Color Filter는 빛의 삼원색인 Red, Green, Blue 화소가 유리판 위에 코팅되어 컬러영상을 구현해주는 역할을 한다. TFT와 Color Filter 사이에 빛의 투과율을 조절하는 액정이 주입되면 TFT-LCD의 첨단기술이 집약된 LCD패널이 완성되며, 여기에 광원이 되는 Black Light Unit, 영상신호를 전달하는 Drive-IC와 PCB, 그리고 케이스가 조립되면 TFT-LCD 모듈이 탄생된다.

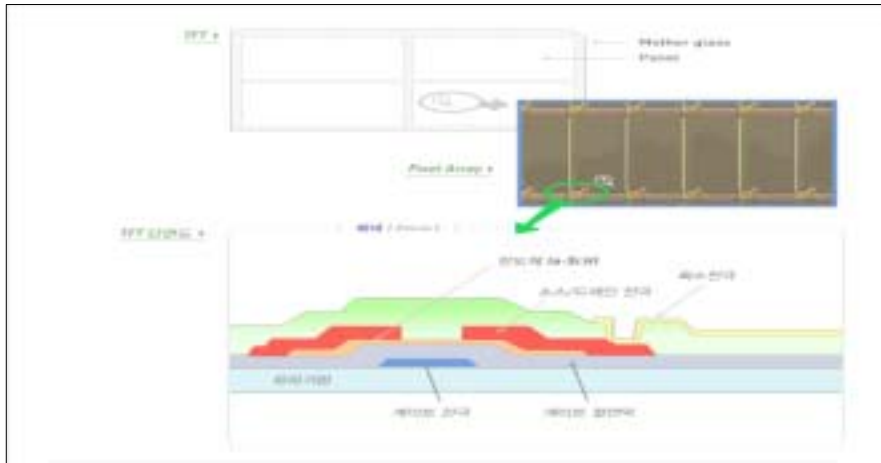
TFT-LCD 공정은 다음과 같이 크게 네 부분으로 나눌 수 있다.

- TFT Array
- Color Filter 공정
- 액정 공정
- Module 공정

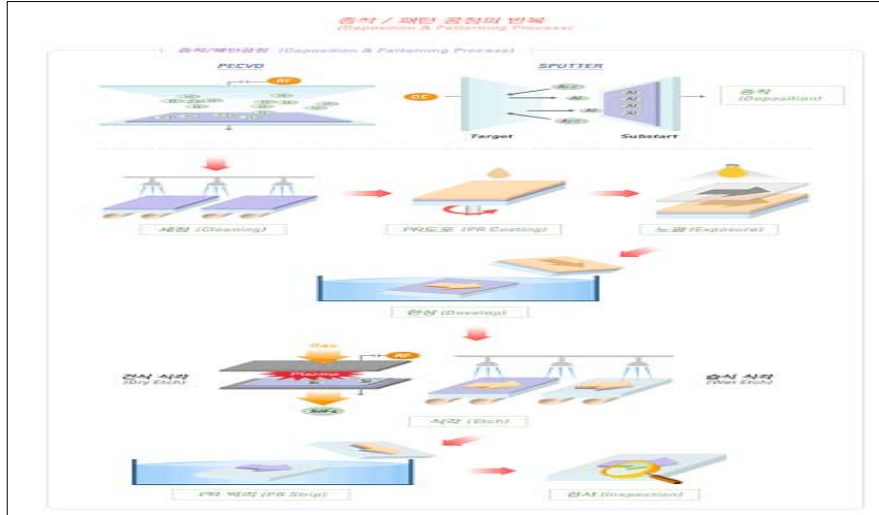
각 공정의 일반적인 제조과정과 TFT-LCD 공정의 특징을 간략하게 살펴보면 다음과 같다.

### 3.1. TFT Array 공정

TFT Array 공정은 TFT(박막 트랜지스터) 기판을 형성하는 공정으로, Silicon 반도체 공정과 매우 흡사하다. TFT 기판은 종류와 기능이 서로 다른 박막(Thin Film)들이 여러 층(Layer)을 형성하고 있으며, 각 층을 이루고 있는 박막들은 각각의 패턴에 따라 패턴을 형성하는 공정을 거치게 된다. 각 박막 층의 패턴을 형성하는 공정은 증착(Deposition), 세정(Cleaning), 감광물질(Photo Register, 이하 PR) 코팅, 노광(Exposure), 현상(Develop), 식각(Etching 공정), PR 박리(Strip 공정), 검사(Inspection) 등의 과정으로 이루어져 있으며, 각각의 박막마다 이러한 과정을 반복적으로 거쳐 각각의 패턴을 형성하고 여러 층을 이루어 TFT 기판을 형성하게 된다.



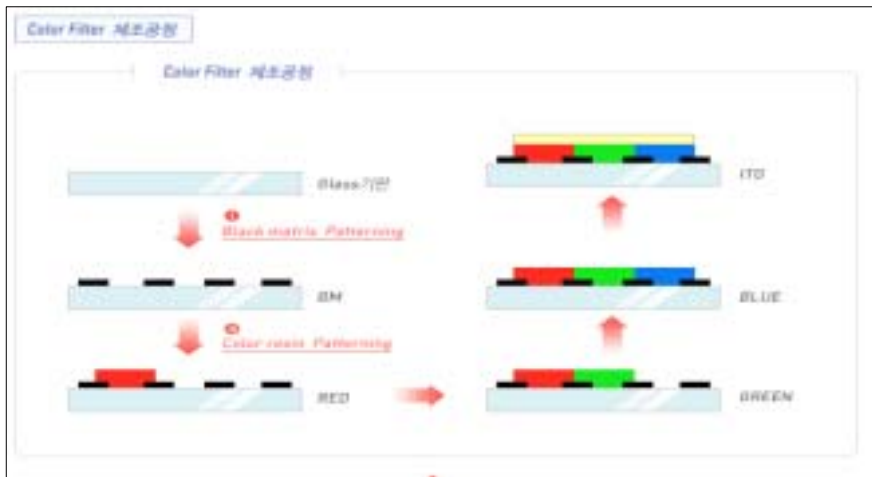
<그림 3.1> TFT Array 공정



<그림 3.2> 증착/패턴공정

### 3.2. Color Filter 공정

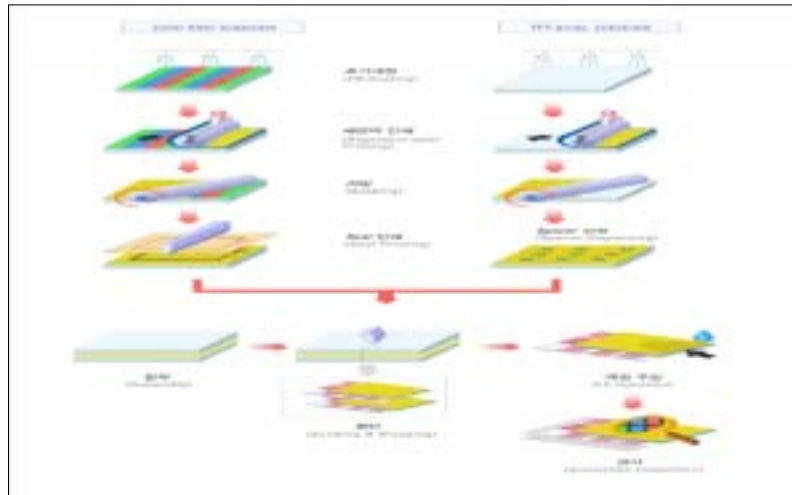
Color Filter 공정은 색상을 표현할 수 있는 화소를 배열한 기판을 형성하는 공정으로, TFT Array 공정과 그 과정이 유사하다. TFT-LCD는 각 셀이 자체 발광할 수 없기 때문에 Back Light의 빛을 통하여 발광하게 된다. 하지만, Back Light는 백색광이기 때문에 색상을 구현하기 위해 Color Filter를 필요로 하며, Color Filter에는 색상을 구현하기 위해 RGB를 표현할 수 있는 화소 기판을 형성하게 된다.



<그림 3.3> Color Filter 제조공정

### 3.3. 액정 공정

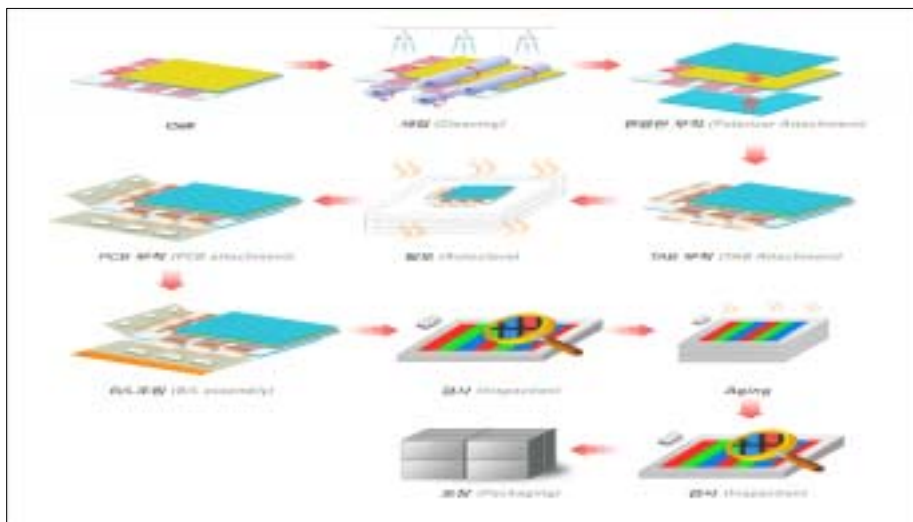
액정 공정은 TFT Array 공정을 거친 TFT 기판과 Color Filter 공정을 거친 Color Filter 기판을 합착하고, Panel 크기대로 절단한 후 TFT 기판과 Color Filter 기판 사이에 액정(Liquid Crystal)을 주입하는 과정이다.



<그림 3.4> 액정공정

### 3.4. Module 공정

Module 공정은 액정 공정까지 거친 TFT-LCD Panel에 신호를 전달하기 위한 PCB 기판을 연결하고, Back Light Unit을 부착하여 Panel을 완성하는 과정이다.



<그림 3.5 Module 공정>

### 3.5. TFT-LCD 공정 특징

TFT-LCD Panel은 매우 미세한 Pixel로 이루어져 있으며, 각 Pixel 내에 하나의 TFT가 존재하여 TFT의 크기는 Pixel 보다 더 작아지게 된다. 게다가 R, G, B 각 Pixel마다 TFT가 위치하는데, 일반적으로 표시하는 해상도에 따른 Pixel의 수는 R, G, B 3개의 Pixel을 하나의 Pixel로 규정한 단위이기 때문에 실제로는 3배의 Pixel이 한 Panel에 존재하게 된다. 예를 들어, TFT-LCD Panel의 최대 해상도가  $1280 \times 1024$  일 경우 Pixel의 수가  $1280 \times 1024 = 1,310,720$ 개가 된다. 하지만 하나의 Pixel은 R, G, B 3개의 Pixel을 하나의 Pixel로 규정한 것이므로 실제 존재하는 Pixel의 수는  $1,310,720 \times 3 = 3,932,160$  개가 되며, 이 3,932,160개의 Pixel 내에 TFT가 각각 존재하는 것이다. 결국, 동일한 크기의 Panel이라면 해상도가 높을수록 Pixel은 더 미세한 크기로 존재하게 된다. 따라서, TFT-LCD 공정은 매우 높은 정밀도(Accuracy)를 요하는 공정이라 할 수 있다.

TFT-LCD Panel은 비교적 넓은 평판인 경우가 많기 때문에 Panel의 극히 일부에라도 불량화소나 얼룩 등의 결점이 존재한다면 이 하나의 Panel은 불량품이 되게 된다. 반도체 Wafer의 경우도 비슷하지만 비교적 그 크기가 작기 때문에 불량 존재하는 셀을 제외한 나머지는 정상적으로 활용이 가능한 경우가 많은 것이 차이가 있다고 할 수 있겠다. 따라서, TFT-LCD 공정은 한 제품 내의 위치에 대한 품질의 균일도(Uniformity)가 타 산업에 비해 더 중요시 되는 공정이라 할 수 있다.

TFT-LCD 공정은 타 산업의 공정에 비해 상당히 복잡하고, 제조에 소요되는 시간이 상당히 긴 공정이라고 볼 수 있다. 공정이 복잡한 만큼 불량이 나타날 가능성이 많고 이전의 단위 공정이 이후의 단위 공정에 영향을 줄 가능성이 매우 많다고 할 수 있다. 특히, TFT Array 공정의 경우, 각 층(Layer)마다 패턴 공정이 반복되므로 먼저 패턴이 형성된 층의 작업 상태가 이후에 패턴이 형성될 층에 영향을 줄 수밖에 없게 된다.

## 4. 데이터마이닝 적용 시 발생하는 문제점

현재 LCD공정에서는 대부분 SPC시스템을 구축하여 고장이나 결점 현황들을 파악하고 조치하는 수준으로 이루어져 있으나 데이터마이닝의 적용 사례는 그리 많지 않다. 현재까지 파악된 데이터마이닝 적용 시 발생하는 몇 가지 문제점에 대해서 언급하고 그 해결책을 제시 하고자 한다.

### 4.1 결측치

데이터마이닝을 하고자 하는 가장 큰 목적은 결점의 원인이나 수율의 저하를 일으키는 원인을 보다 효율적으로 그리고 보다 빨리 찾아내서 빠른 대응을 통한 무결점 고 수율의 공정을 만들어 나가기 위함이다. 그러기 위해서는 하나의 제품이 만들어지는 전체 공정에서 공정 상태를 체크하고 이를 데이터화 한 후 이를 분석해서 원인을 찾아야 한다. 그러나 전체 공정에서 공정 상태를 측정하는 것은 그리 쉬운 일은 아니다. 측정을 한다는 것은 공정 중에 자동적으로 수많은 측정요소들을 측정할 수 있는 것이 아니기 때문에 공정 중에 측정을 위해 라인 밖에서 측정하여야 할 필요성도 있고, 경우에 따라서는 제품에 나쁜 영향을 미치면서 까지 측정을 실시하여야 한다. 따

라서 하나의 제품이 만들어 지는 과정 처음부터 마지막 공정까지 전체 공정에 대한 측정이 이루어진다는 것은 거의 불가능하며, 각 공정별로 측정을 위하여 일부 제품만 샘플을 하여 측정하여야 한다. 따라서 각 공정별로 선정된 샘플이 서로 다르기 때문에 전체 공정 차원에서 보면 상당량의 결측치 존재하게 된다. 그러나 분석을 위한 데이터의 형태는 하나의 제품에 대하여 전 공정에서 모든 측정값이 있기를 요구한다. 결국 상당량의 결측치가 존재하지 않을 수 없다. 어찌면 이 요인이 LCD공정에서 데이터마이닝을 함에 있어서 가장 큰 어려움일 수 있다. 이를 해결하기 위한 방안은 몇 가지가 있다. 하나는 분석 시에 사용하는 변수의 수를 가급적 줄이는 방법이다. 여러 가지 통계기법들을 이용해서 일차적으로 관련이 없는 불필요한 변수들을 분석에서 제외시키거나 주요공정의 탐색을 통해 불량에 직접적으로 영향을 미치는 변수라고 판단되는 주요변수들만 추출하여 분석을 하면 결측치의 양을 줄일 수 있다. 다른 하나는 각 제품별 결측치가 존재 할 때 타 제품과 비교 분석하여 유사한 제품의 값으로 대체하는 방법이다. 그러나 이 경우 결측치가 너무 많을 경우 그리 효용 있는 방법은 아니다. 또 다른 방법으로는 공정 전체에 대한 분석 보다는 공정별 원인과 각 공정별 최적 조건을 찾아 해당 공정 내에서의 결점을 줄이거나 최적 조건을 찾아냄으로써 전체 공정의 수율을 올리는 방법이 있을 수 있다. 이 경우 전체 공정에 대한 분석이 아니라 단위 공정별로 분석이 이루어지므로 절대적으로 결측치가 줄어들게 된다.

#### 4.2 결점/무결점의 비

최근 들어 결점을 줄이려는 시도 보다는 한결음 더 나아가 무결점이 되도록 공정을 관리하고 있다고 보아도 될 정도로 결점을 줄이기 위해서 부단한 노력을 기울이고 있고, 그 노력의 결과로 결점의 수가 상당량 줄어 있다. 그 결과 데이터마이닝 분석에서의 가장 중요한 Target 변수인 결점여부를 나타내는 변수의 결점이 무결점에 비해 엄청나게 낮은 비율로 발생한다. 이를 해결하기 위한 방법으로 데이터마이닝 분석과정 중 Sampling단계에서 결점과 무결점의 비를 조절하여야 한다. 결점과 무결점의 Sample의 비를 적절히 조절 한 다음 모델링 후 해석 시에 판정기준값(Cut Off Point)을 조절 하여 해석 하여야 한다.

#### 4.3 공정별 패턴의 교락 현상

앞서 소개한 LCD제조공정과 같이 LCD공정은 크게 4가지 공정을 거쳐 하나의 LCD가 만들어지고 각 공정마다 증착/패턴 공정을 반복하는 세부공정이 이루어지며, 각 세부 공정별로 몇 개의 설비가 있고 제품들은 서로 다른 설비를 거치게 되고 1000여개가 넘는 수많은 측정요소(아이템)들로 전체 Data가 구성되어 진다. 또한 각 공정이나 특정 설비 하나에 의한 불량이 발생할 수도 있지만 몇 개의 아이템이나 몇 개의 설비가 복합적으로 불량에 영향을 미칠 가능성도 있다. 즉, 하나의 제품이 만들어지기까지는 수많은 세부 공정을 거쳐야 하며, 각 세부 공정별 과정을 거치면서 서로 교락 현상이 발생되어 각 공정의 아이템들이 최종 Target 변수인 양품/불량의 여부에 미치는 영향들이 어떤 패턴을 갖기는 어려운 일이다. 어떤 패턴이 발생된다 하더라도 다른 공정의 또 다른 패턴으로 인해 그 패턴이 어떤 영향을 미치는지 해석하기에도 상당한 어려움이 있다. 따라서 전체 공정 Data와 Target 변수간의 관계를 보는 것보다는 각 공정별로 불량여부를 측정하여 각 공정별 데이터마이닝 작업을 하는 것이 보다 효율적인 운영 방법이 될 수 있다.



#### 4.4. 다양한 불량 유형

LCD공정을 완료하고 최종 검사단계에서 나타나는 불량 유형은 거의 1000여 가지에 이른다. 아주 미세한 점에서 빛이 세어 나와 도저히 활용할 수 없는 불량도 있고, 전체적으로 기준치 이하의 밝기나 그 반대의 경우도 있고, 가로나 세로로 줄이 발생되어 불량이 되어지는 경우도 있다. 이렇게 수 많은 종류의 불량을 단순히 불량이라는 하나의 값으로 정의하고 분석하기에는 발생하는 불량의 유형의 가지 수 만큼이나 그 원인도 다양하다. 따라서 발생원인에 대한 분석을 함에 있어서 Target변수를 단순히 불량/양품으로 하고 분석할 수 만은 없을 것이다. 따라서 동일 유형의 불량 발생 빈도가 높은 것은 각 유형별로 분석이 이루어져야 하며, 그렇지 않은 것은 비슷한 유형끼리 그룹을 형성한 후 분석이 이루어져야 한다.

### 5. 사례 연구

아래 사례는 모 LCD제품에 대하여 가로줄 얼룩이 발생하는 불량에 대한 원인을 조사하기 위하여 분석된 결과를 사례로 제시한다. 참고로 본 내용은 LCD공정 전체에 대한 보안 관계상 제조사나 공정에 대한 세부 내용을 정확히 밝힐 수 없음을 미리 언급한다.

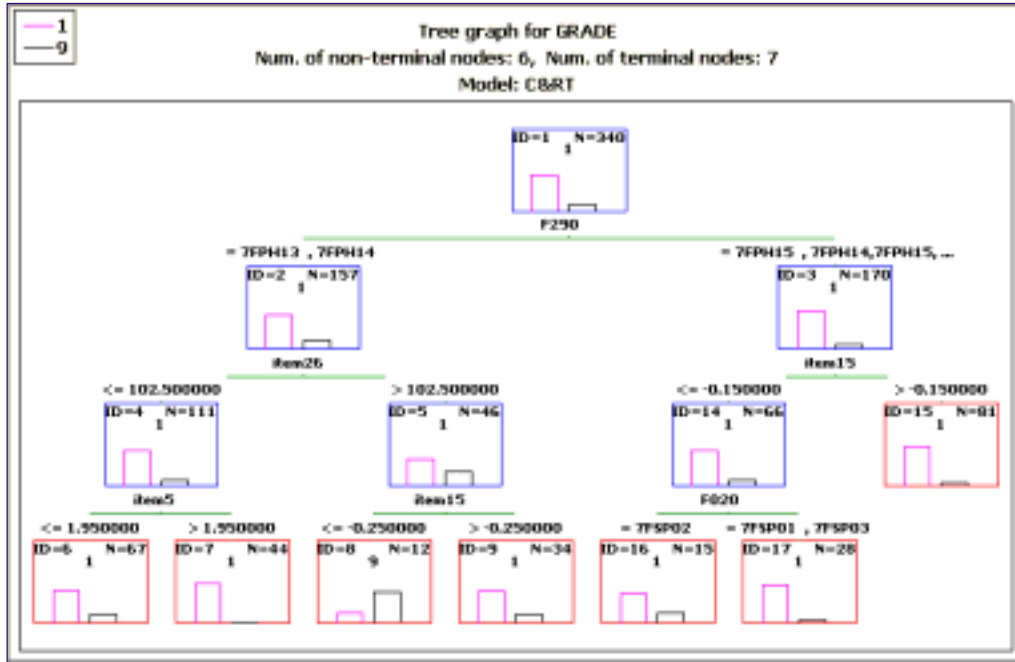
일정 기간 동안 발생한 가로줄 얼룩이 발생한 제품 전체와 불량이 되지 않는 제품 중에서 일부 표본을 뽑아 분석에 사용하였다. 본 분석의 목적은 TFT-array 공정에서 가로줄 얼룩이 발생하는 원인이 무엇이며, 어떤 조건에서 많은 불량이 발생되고, 그렇지 않은 조건은 무엇인지를 찾고자 하는 것이다.

전체 데이터 중에서 30%이상의 결측을 보이는 샘플과 변수를 제외시키고 남은 데이터를 분석에 활용하였다. 분석에 사용된 데이터의 수는 340개이며 이중 양품의 개수는 291개, 가로줄 얼룩이 발생한 제품이 49개인 데이터로 구성되었으며, 초기 분석에 사용된 독립변수는 18개의 설비변수와 147개의 환경성 변수를 사용하였다.

생산 과정에서 일부 공정에서는 결함이 발견되면 공정 자체를 한 번 더 거치는 과정이 있다. 예를 들어 세정공정을 거치고 난후 완벽한 세정이 이루어지지 않았다고 판단되는 일부는 추가적인 세정작업이 이루어진다. 따라서 이러한 제품의 경우 세정 후에 발생하는 데이터는 다른 데이터와 달리 동일 셀에 두 개의 데이터가 존재하게 된다. 이를 공정에는 rework이라고 하고 분석을 위하여 rework이 불량에 원인인지를 고려하기 위하여 rework의 여부를 정리하는 변수를 새로 생성하고, rework의 영향을 고려한 분석이 이루어 졌다. 기본적인 기술 통계량과 각 설비별 불량률의 차이를 고려하여 차이가 있는 설비와 환경변수들을 대상으로 Feature Selection을 실시하였다. 이 결과 3개의 설비 변수와 3개의 환경 변수가 불량에 주요 영향을 미치는 변수로 채택이 되었으며, 채택된 6개의 독립변수와 가로줄 얼룩의 여부인 Target 변수를 이용하여 Decision Tree를 통계패키지인 STATISTICA를 이용하여 분석한 결과 아래 <그림5.1>과 같다.

분석결과 불량 유형이 현저히 떨어지는 노드로는 7, 15, 17번 노드이며, 그 반대로 불량률이 현저히 높은 노드는 8번 노드로 나타났다. 불량률이 현저히 높아지는 원인변수인 item15는 세정 시간을 나타내는 변수로서 기준 시간에 비해 세정 시간이 어느 정도에 미치지 못할 경우 다량의 불량이 발생되어 즉각적인 조치가 이루어 졌다. 또한

불량이 현저히 떨어지는 조건들에 대하여 세부 환경적인 요인과 추가적인 분석을 통하여 가로줄 얼룩에 발생에 대한 추가적인 분석들이 이루어졌으며, 이러한 노력에 의하여 공정을 개선한 결과 가로줄 얼룩의 발생률을 약 6%p정도 개선 효과를 발생하였다.



<그림 5.1> TFT-array 공정에서의 Decision Tree 분석 결과

## 6. 결론

최근 들어 LCD 산업의 중요성이 부각 되면서 전체 제품의 수율향상이나 결점을 줄이기 위한 부단한 노력들이 계속되고 있다. 이 과정에서 LCD 산업에서 제품의 불량에 영향을 주는 요인들을 찾기 위한 수많은 노력들이 이루어졌고, 최근 들어 데이터마이닝을 적용하는 단계에까지 발전하였다.

본 논문에서는 TFT-LCD 제조 공정에 대하여 간략히 소개하고, LCD 공정의 특징에 관하여 설명하였다. LCD산업에서 데이터마이닝을 적용하기 위해서는 우선적으로 해결 되어야 하거나 반드시 고려되어야 할 몇 가지 사항들을 제시 하였다. 제시된 문제들에 대한 뚜렷한 해결책은 각 제조 회사별 특성에 맞게 방법을 찾아야 할 것이며 이러한 고려 없이 데이터마이닝을 하거나 데이터마이닝 시스템을 도입한다면 CRM에서 분석을 고려하지 않고 시스템을 도입함으로써 발생하였던 실패의 사례를 그대로 답습할 가능성이 높다고 할 수 있다.

앞에서 제시된 몇 가지 문제점들에 대한 해결책은 명확하고 정확한 해결책이라고는

보기 힘들 것이다. 보다 정확한 해결책을 제시하기 위해서는 보다 많은 연구와 보다 많은 분석 경험을 구축함으로써 이루어질 수 있고, 이러한 활동을 통해서 새롭게 대두될 수 있는 문제점들을 파악하고, 그러한 문제점들을 해결하고자 하는 노력을 지속되어질 때 최고의 수율을 올릴 수 있는 공정이 이루어질 것이다.

### 참고문헌

1. 백준걸, 김강호, 김성식, 김창욱(2000), "실시간 기계 상태 데이터베이스에서 데이터마이닝을 위한 적응형 의사결정트리 알고리즘," *대한산업공학회지*, Vol. 26, No. 2, pp 171~183.
2. 백동현, 한창희(2003), "데이터마이닝을 이용한 반도체 FAB공정의 수율개선 및 예측", *한국지능정보시스템학회논문지*, 제 9권 1호, pp 157~177
3. 안진석, 고용민, 장중순(1999), "데이터마이닝을 이용한 최적공정조건 탐색", *대한설비관리학회지*, Vol. 4, No. 2, pp129~144.
4. 장남식(1999), "성공적인 지식경영을 위한 핵심정보기술 : 데이터마이닝", *대청미디어*
5. Braha, D., and Shmilovici, A.(2002), "Data Mining for Improving a Cleasing Process in the Semiconductor Industry," *IEEE Transactions on Semiconductor Industry*, Vol. 15, No. 1, pp91~101
6. Kusiak, A. and Kurasea, C., "Data Mining of Printed-circuit Board Defects," *IEEE Transactions on Robotics and Automation*, Vol. 17, No. 2, pp 191~196.
7. Lian, J., Lai, X. M., Lin, Z, Q., and Yao, F. S.(2002), "Application of Data Mining and Process Knowledge Discovering in Sheet Metal Assembly Dimensional Variation Diagnosis", *Journal of Materials Processing Technology*, Vol. 129, pp 315~320.
8. Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P.(1996) "From Data Mining to Knowledge Discovery in Databases," *Advances in Knowledge Discovery and Data Mining*, AAAI Press/MIT Press, pp 1~34.

[ 2005년 9월 접수, 2005년 10월 채택 ]