

TFT-LCD 산업에서의 품질마이닝 시스템

이 현 우¹⁾ · 남 호 수²⁾ · 최 경 호³⁾

(주)한국신뢰성기술서비스¹⁾ · 동서대학교 시스템경영공학과²⁾ · 전주대학교 여론통계학과³⁾

Keywords : Data Mining, Quality Mining System, SPC, TFT-LCD

Abstract

Data mining is a useful tool for analyzing data from different perspectives and for summarizing them into useful information. Recently, the data mining methods are applied to solving quality problems of the manufacturing processes.

This paper discusses the problems of construction of a quality mining system, which is based on the various data mining methods. The quality mining system includes recipe optimization, significant difference test, finding critical processes, forecasting the yield.

The contents and system of this paper are focused on the TFT-LCD manufacturing process. We also provide some illustrative field examples of the quality mining system.

1. 서론

21세기의 컴퓨터의 발달은 우리들의 일상생활에 많은 영향을 주었다. 며칠이 걸려서 주고받던 소식들도 단 몇 초안에 많은 양의 정보를 서로 주고받을 수 있으며, 시간을 들여 백화점이나 시장을 가지 않더라도 원하는 물건을 인터넷이라는 도구로 구매할 수 있게 되었다. 이러한 모든 생활들이 컴퓨터와 연결이 되면서 업계에서는 CRM이라는 새로운 이슈가 도래되었고, CRM 업무 중에 가장 중요한 역할을 하는 것 중 하나가 데이터마이닝이다. 데이터마이닝을 통해 고객들의 성향을 파악하고 고객에 맞는 상품을 추천하거나, 우수고객이 이탈가능성을 파악하여 이탈하지 않도록 사전에 마케팅을 한다.

이러한 변화는 제조업에서도 마찬가지로 일어나고 있다. 전 공정의 상황을 자동으로 계측하고 이들 데이터를 DB화해서 실시간으로 공정상태를 파악할 수 있게 되었다. 또한 제품의 고장이나 결점의 상태를 파악하고 이들의 패턴을 감지함으로써 무결점의 완벽한 제품을 만들기 위하여 많은 시간을 들이고 있다. 고장이나 각종 결점에 대한 데이터를 통계분석하고 분석결과를 이용해

서 결점의 원인을 해결하고, 다시 이러한 결점이 일어나지 않도록 예방보전하는 과정을 벌이고 있다. 이런 방대한 데이터를 가공, 분석하는데 중요하게 활용될 수 있는 기법중의 하나가 데이터마이닝이다.

제조 공정에서의 마이닝의 적용은 CRM에서의 마이닝보다 현실적으로 많은 어려움이 존재하기 때문에 그다지 활성화 되어있지 않은 게 사실이다. 그러나 이러한 어려움에도 불구하고 무결점의 최상의 제품을 만들기 위해서는 반드시 필요한 과정이 데이터 마이닝이며, 이를 효율적으로 활용하기 위해서는 데이터 마이닝 시스템의 구축이 필수불가결하다고 할 수 있을 것이다.

본 논문에서는 LCD패널제조분야에서 공정을 효율적으로 운영하기 위한 데이터마이닝 시스템에 대한 현황과 방법론을 설명하고 실제 구축된 사례를 통해서 구축내용과 활용사례를 설명하기로 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 데이터 양이 기하급수적으로 증대함에 따라 이전과 같은 수작업에 의한 데이터 분석 방법으로는 도저히 불가능에 가까운 일이라고 할 수 있다.(장남식, 1999). 따라서 방대한 양의 데이터로부터 유용한 지식을 얻도록 사람을 지능적이고 자동적으로 지원하는 새로운 기법과 도구가 필요하게 되었다.(Fayyad 등, 1996).

백준결 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. 에칭율에 영향을 미치는 공정변수(기계 내의 압력, 온도, 에칭가스 밀도, 웨이퍼 위치 등)를 속성으로 정의하고 에칭율에 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 혼합형태로 공정변수가 이루어지며 품질특성치로서 불량률 데이터가 주어진 경우에 최적 공정조건에 대한 정보추출을 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.

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.

이현우 등(2005)은 LCD산업에서의 제조공정을 간략히 소개하고, 복잡한 제조 공정에서 데이터마이닝을 적용함에 있어서 발생할 수 있는 여러 가지 문제점을 제시하고 이를 해결하기 위한 방법들을 제시하였다.

3. 품질마이닝 시스템 구축

앞서 언급하였다시피 제조공정에서의 데이터마이닝의 목적은 불량이나 결점에 영향을 미치는 요인을 찾는 것과 수율을 향상 시킬 수 있는 방안을 찾고자 하는데 있다. 즉 품질에 관련된 주요요인들을 찾고 이들 요인들 중에서 어떤 조건에서 제품을 생산할 경우에 최고의 제품을 생산할 수 있는지 혹은 품질관리에서 주요공정에 대한 적용을 위하여 데이터마이닝을 활용할 수 있다. 다시 말해 기존의 품질관리는 단변량적인 분석 방법이라면 이를 한 차원 높은 고급 분석을 통한 공정관리에 적용한 것이 데이터마이닝이라 할 수 있다. 따라서 본 논문에서는 이런 제조 공정의 데이터마이닝을 품질마이닝이라 한다.

최고의 제품이나 불량률의 원인을 정확히 파악하기 위해서는 일회성의 분석으로만 가능한 것이 아니라 지속적인 분석과 분석된 결과를 공정에 반영하고 반영된 이후의 변화를 감지하고 변화된 이후에도 더 좋은 공정을 만들기 위한 이러한 과정들을 반복함으로써 이루어진다. 이를 위해서는 품질마이닝을 위한 시스템이 구축

되어야 한다. 우선 품질마이닝 시스템 구축을 위하여 갖추어야 할 기본적인 내용 몇 가지 사항들을 살펴보면 다음과 같다.

3.1 시스템구축의 목적(필요성)

세계적으로 LCD산업의 시장 변화는 대량화, 무한경쟁화, 급속한 기술발전 등 무한 경쟁체제가 갖추어 졌다고 보아야 한다. 국내뿐만 아니라 국제적으로 그 시장의 규모가 커진 것뿐만 아니라 경쟁역시 치열한 상황이다. 이러한 무한 경쟁의 시장 상황에서 살아남거나 최고가 되기 위해서는 결국 품질 경쟁력을 확보하는 것이 중요하다. 이를 위해서는 효과적인 수율 관리 체계와 무결점 공정 라인을 유지하는 것이라 하겠다. 기존의 SPC시스템에서 한발 나아가 수율 향상과 무결점 공정을 위한 체계화, 구체화, 자동화된 생산라인과 공정의 이상 발생시 보다 효율적인 관리 체계 구축을 위해서는 품질마이닝 시스템의 구축이 필수일 것이다. 또한 다양한 품질/공정 관련 데이터의 효과적인 통합관리를 하고, 엔지니어가 발견하지 못했던 인자간의 패턴 및 유의차나, 공정 불량 의 원인 또는 주요인자들을 발견해서 이를 효율적으로 관리 할 수 있어야 한다.

3.2 주요내용

기존의 공정관리를 위하여 SPC 시스템을 구축한 사례는 많이 있었다. 품질마이닝 시스템을 SPC 시스템과 비교한다면 SPC 시스템을 OLAP차원의 분석이라면 마이닝 시스템은 그보다 한 차원 높은 분석이 가능한 시스템이라고 할 수 있으며, 따라서 품질마이닝 시스템의 구성 요소 역시 SPC시스템보다 차원 높은 분석 기법들을 포함하고 있어야 할 것이다. 마이닝 시스템 구축에 있어서 고려되어야 할 주요 내용을 살펴보면 다음과 같다.

1) 주요변수 추출

LCD 공정의 데이터 들은 CRM이나 제조공정에서 반도체 공정과는 달리 수많은 결측치들을 포함하고 있다. 그 이유로는 공정의 복잡성으로 인하여 모든 제품들에 측정을 한다는 것이 불가능하며 따라서 모든 측정은 표본을 통한 측정이 이루어진다. 또 다른 이유로는 과도한 측정은 오히려 측정 자체만으로도 불량의 원인이 될 수 있기 때문에 많은 결측치는 어쩔 수 없는 현상이라 할 수 있다. 따라서 이러한 많

은 결측치를 내포한 데이터의 분석을 위해서는 주요인자를 찾아 그 변수들만을 이용한 모델링을 함으로써 결측치로 인해 발생하는 문제점들을 다소 해결 할 수 있으며, 모델링의 효과를 상승시킬 수 있다. 반응변수(불량, 결점 등)에 영향을 미칠 수 있는 수많은 원인변수들 가운데 크게 영향을 미칠 수 있는 설명변수들을 선택해 내는 프로세스로서 반응변수의 종류에 따라 다양한 통계적 방법(독립성검정(카이제곱검정), 분산분석 등)을 이용하게 된다.

2) 유의차분석

앞에서도 언급하였다시피 LCD공정은 CRM이나 반도체 공정보다 훨씬 복잡한 공정으로 이루어져 있다. 이런 복잡한 공정을 제대로 관리하기 위해서는 어떤 공정의 어떤 수준에서 문제가 자주 발생하는지를 파악하여 그 수준에 대한 적절한 조치가 이루어져야 한다. 반응변수에 영향을 미칠 수 있는 요소(주로 4M1E ; 설비, 작업자, 소재, 방법, 환경조건)들의 복수 수준에서 반응변수들의 값들이 통계적으로 의미 있는 차이를 나타내는가를 분석하고, 이에 기초하여 4M1E 요소에 대하여 적절한 조치를 취하게 하며, 이를 통하여 대량의 부적합을 예방하는 효과를 기대 할 수 있다.

유의차분석 도구로는 반응변수의 분류에 따라 분산분석, 독립성검정 및 의사결정트리 등이 활용될 수 있다.

3) 주요공정파악

반응변수에 영향을 크게 미치는 주요공정을 통계적으로 분석하여 해당공정의 중요도(기여도)를 평가한다. 이를 통하여 공정별 관리의 엄격도에 대한 계획을 수립/운영할 수 있다.

4) 최적경로 탐색

LCD공정의 최대 목표는 무결점 제품 생산이나 최고의 수율을 위한 여러 공정들을 관리하고자 하는 것이라고 할 수 있다. 이를 위해서 공정변수들의 어떠한 조건에서 결점이 최소화 되고, 수율이 극대화 되는지를 탐색하는 기능으로서 최적경로 탐색을 주목적으로 이용한다.

5) 예측 및 추정

여러 가지 환경 내지 조건들을 변화시켰을 때 수율 및 결점 수 등의 반응변수가 어떻게 변화하게 되었는지를 추정 또는 예측하는 내용으로 통계적 모델링을 통하여 반응변수의 추정

/예측 값을 제공한다.

6) 고장이나 결점 발생의 패턴인식

결점에 영향을 미치는 변수가 어떤 변수이고 어떤 수준에서 혹은 어떤 조건에서 발생되어 진다는 식의 내용도 중요하지만 여러 가지 변수들이 조합적으로 어떤 환경에서 고장이나 결점이 발생하는 경향이 나타난다면 더욱 중요한 정보가 될 것이다. 이는 결국 전자의 경우가 OLAP차원의 분석이라면 후자의 경우는 품질마이닝을 통한 분석이라고 할 수 있을 것이다.

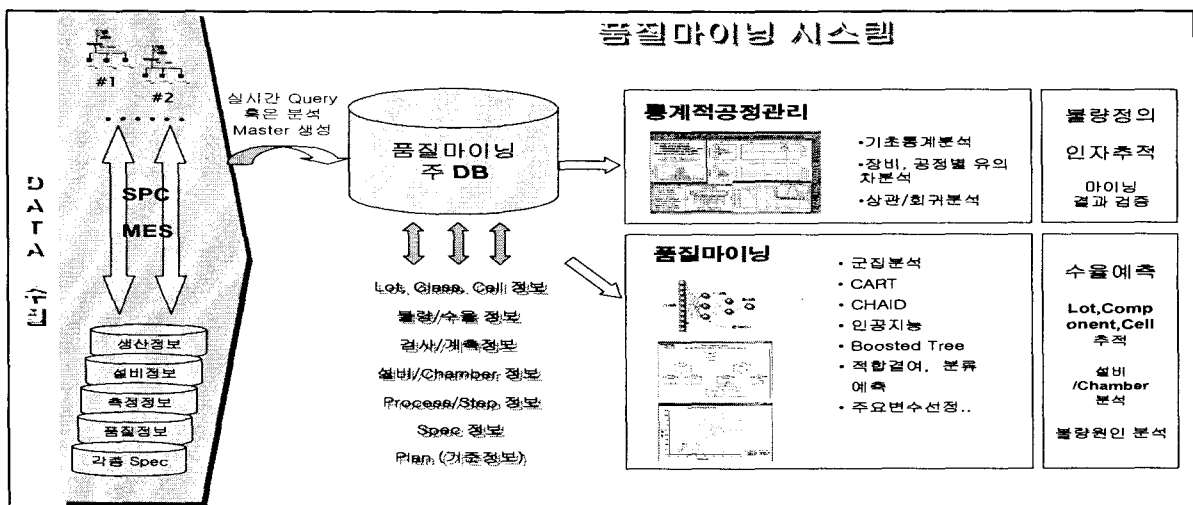
7)기타

그 외 마이닝을 하기 위한 몇 가지 중요한 기능들로 구성되어 진다. 간단히 살펴보면 가장 기본적으로 데이터를 관리 가공, 정제할 수 있는 기능과 각종 도표나 그래프를 통하여 분석할 수 있는 기능 그리고 이러한 통계적인 분석 기능들을 이용하여 분석한 결과를 이용해서 보고서를 작성할 수 있는 기능들로 구성되어 있어야 할 것이다.

3.3 품질마이닝 시스템의 구성

마이닝 시스템의 구성은 데이터를 수집해서 마이닝을 할 수 있는 데이터를 제공해주는 DB부분, 각종 간단한 통계와 각종 도표 그리고 그래프 등을 볼 수 있는 SPC부분 및 여러 가지 주제별로 마이닝 기법들을 이용하여 마이닝 분석을 할 수 있는 마이닝 부분으로 구성이 된다.

1) DB부분



<그림 3.1> 마이닝 시스템의 구성도

마이닝 시스템에서 가장 기본이 되고 중요한 부분일 수 있다. 필요한 데이터를 수집하고 마이닝에 필요한 데이터로 수정, 정제, 변환을 하는 작업도 이루어 져야하며, 분석결과 각 주제별로 유용한 정보들이 빠짐없이 분석에 포함될 수 있도록 설계되어야 한다. 또한 분석의 결과에 신경을 쓰지 않고 너무 많은 변수들을 DB에 포함 시킬 경우 불필요한 변수들이 너무 많아 발생하는 결측값이 많아지는 부분과 분석을 위해 요구되는 분석 시간을 엄청나게 필요로 하기 때문에 효율적이 되지 못하는 경우가 많다. 따라서 반응변수에 영향을 많이 미치는 변수들로 효율적인 DB화가 중요한 작업이라고 할 수 있다.

2) SPC 부분

별도의 SPC시스템이 구축되어 있으면 이를 연동하여야 할 필요성이 있다. 마이닝 시스템에 가장 중요한 부분은 아니라고 할지라도 없어서는 안 될 중요한 기능들을 포함하며, 전체 공정의 운영을 간략히 파악하기 위해서는 중요한 역할을 한다. SPC시스템은 대부분 통계패키지를 이용하여 구축이 되어있고, 여러 가지 통계량을 활용하는 경우도 있지만 대부분 각종 그래프를 이용한 분석이 많은 부분을 차지하기 때문에 그래픽에 탁월한 기능을 갖고 있는 통계패키지의 선택이 중요한 사안일 수 있다.

3) 마이닝 부분

주제별로 필요한 변수들이 달라질 수 있고, 여러 가지 상황별로 선택되어지는 옵션이 변화

될 가능성이 많다, 따라서 마이닝 부분의 구축 시 반드시 선행되어야 할 사항은 통계학적인 관점에서의 분석 혹은 컨설팅이다. 해당 변수들을 연결해서 마이닝 결과를 도출할 수 있는 기능을 보유하는 것이 중요한 것이 아니라 어떤 주제에 어떤 옵션을 사용하여 최적의 조건을 만들어 공정에 효율화를 극대로 상승하는 것이 궁극적인 목적이 될 것이다. 따라서 마이닝에 필요한 모든 조건들을 검토해서 시스템에 반영하여야 하며, 필요시 각 주제별로 선행 분석을 통해서 필요한 절차별로 정리해서 시스템에 반영하는 것도 하나의 방법이 될 수 있을 것이다. <그림 3.1>은 이러한 내용의 마이닝 시스템의 구성도를 제시하고 있다.

4) 통계패키지

품질마이닝 뿐만 아니라 CRM에서의 마이닝에서도 가장 중요한 부분은 단순 마이닝 시스템의 구축이 아니라 분석이다. 이를 위해서는 보다 신뢰성이 있고, 사용하기 편한 통계패키지의 선택은 매우 중요하다고 볼 수 있다. 현재 이런 시스템을 구축하기 위하여 가장 많이 사용되고 있는 통계패키지로는 SAS, SPSS, STATISTICA, S-Plus 등이 있다. 모든 패키지들이 장단점이 있지만, 단순한 통계량을 제시해주는 기능보다는 데이터들을 보다 효율적으로 이해하기 위해서는 그래픽 면에서 강점이 있고 마이닝 분석을 초보자도 쉽게 활용할 수 있도록 응용이 편하게 되어지는 통계패키지의 선택이 좋을 듯하다.

4. 구축사례

A사의 경우 품질에 영향을 미치는 인자들을 발견/관리 하고, 패턴을 발견하여 최적의 수율을 올리기 위해 마이닝 시스템을 구축하였다. 그 세부내용은 다음과 같다.

4.1. 시스템 목적

품질마이닝 시스템은 복잡하고 다양한 분석을 사용자의 편의를 위해 일부 정형화된 분석 기법을 설정하고, 이에 정규분석에 4가지 주제별 분석과 일정주기 분석에 수율/불량에 관련한 분석을 Web 환경으로 구축 하였다.

또한 기존의 분석에 사용되었던 공정 관련 데이터 수집 방법을 대폭 개선하여 사용자가 쉽게 분석에 사용 할 수 있도록 본 시스템에서

지원할 수 있도록 하였다.

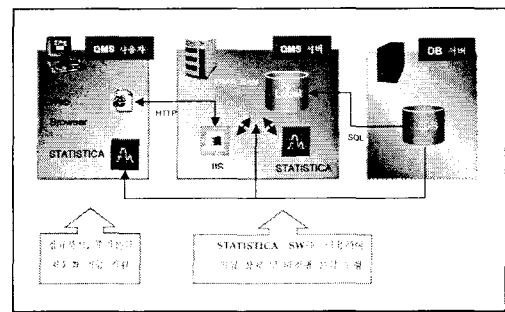
주요 분석업무 (4가지 정규분석)는

1. 반응변수에 크게 영향을 미치는 주요변수(공정)는?
2. 반응변수에 유의하게 영향을 끼치는 4M1E 요소는 무엇이며, 어떤 4M1E요소? 긍정적/부정적으로 특성에 차이를 발생시키는가?
3. 주요변수의 어떤 조건(조합)에서 특성이 최적화(recipe optimization) 되는가?
4. 조정(최적화)된 조건에서 반응변수의 추정/예측값은?

이며, 이를 효과적으로 분석/조명해 볼 수 있는 시스템이다.

4.2. 시스템의 구성

본 품질마이닝 시스템은 정규분석 / 주기분석 두 Theme 별 분석을 Web환경으로 구성 되었다. <그림 4.1>에서는 시스템의 구성도를 나타내고 있다.



<그림 4.1> 품질마이닝 시스템 구성도

본 시스템은 분석 대상 및 목록에 적절한 Table을 Query 할 수 있도록 해주는 DB 서버와 필요한 데이터를 추출하여 STATISTICA를 이용하여 분석할 수 있는 품질마이닝 서버와 웹상에서 사용자가 로그인 후 본 시스템을 활용할 수 있는 사용자 모드로 구성되어 있다.

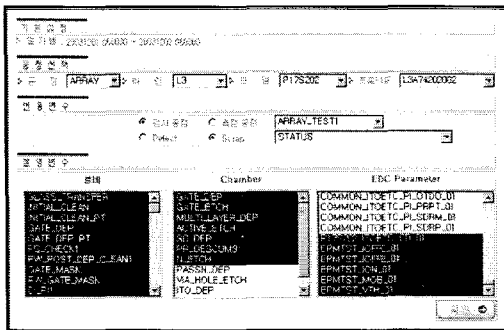
4.3 분석을 위한 주요내용

분석의 내용은 앞에서 언급한 4가지 정규분석과 주기적으로 분석이 필요한 부분을 별도로 분류하여 크게 두 부분으로 나누어 구성하였다. <표 4.1>는 주요 분석주제에 대한 내용을 도시하였다.

<표 4.1> 주요분석 주제

구분	주요내용	세부분석	
정규분석	기본설정		
	세부설정		
	Feature Selection	유의차분석	
		주요공정파악 최적경로탐색 예측 및 추정	
주기분석	기본설정		
	주요분석	수율분석 유의차분석	

사용자는 웹 UI 화면에서 승인(사용자 권한 설정)절차를 거쳐 본 시스템에 들어온다. 분석을 위해서 제시된 분석을 수행하기 위해서는 분석에 필요한 기본적인 내용 (데이터의 종류 : 일자, 생산라인, 설명변수, 반응변수 등)을 선정하고 분석 내용을 클릭하는 것만으로 분석을 할 수 있도록 구성하였다. <그림 4.2>에서는 기본적인 사항들을 선택할 수 있는 화면이다.



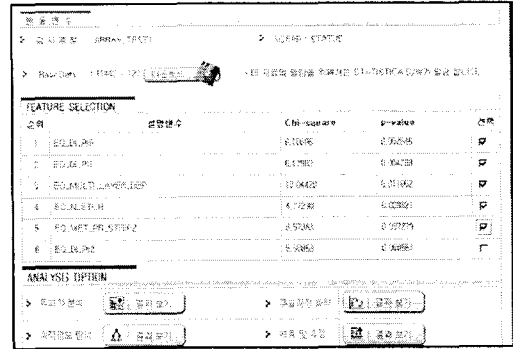
<그림 4.2> 기본사항 선택화면

4.4 주요분석 방법

기본적인 사항들을 선택한 후 <표 4.1>에서 언급된 내용의 분석을 위해서는 공정에 영향을 크게 준다고 예상되는 변수들을 선택하여 주기 분석의 4가지 내용중 하나의 분석을 실시하면 된다.

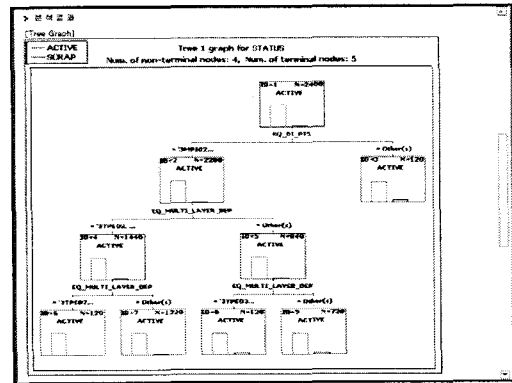
Feature Selection 결과는 각 설명변수 별로 반응변수의 차이분석결과를 p-value순으로 나타낸다. 각 설명변수 리스트 우측의 체크박스는 해당하는 설명변수를 4대 분석에 포함시킬 것인지 여부를 결정한다.

<그림 4.3>에서는 Feature Selection 의 화면을 나타내고 있다.



<그림 4.3> Feature Selection 화면

화면 우측 상단의 유의차 분석/주요공정 파악/최적경로 탐색/예측 및 추정 4가지 버튼과 화면 하단의 4가지 분석 버튼을 클릭함으로써 4대 분석을 수행할 수 있다. <그림 4.4>에서는 최적경로 탐색을 위하여 의사결정나무의 화면을 나타낸다.



<그림 4.4> 최적경로 탐색 결과화면

5. 결론

본 논문에서는 TFT-LCD 공정에서의 데이터 마이닝 시스템에서 반드시 필요로 하는 주요기능에 대한 설명과 실제 시스템을 구축한 사례를 들어 설명하였다. TFT-LCD공정은 공정 자체의 복잡성과 각 공정별로 데이터 획득에 대한 문제점 등 아직까지 완벽한 데이터 마이닝 시스템을 구축하기 위해서는 많은 문제점이 존재하고 있고, 이를 해결하기 위한 많은 노력들이 지속적으로 이루어져야 원하는 시스템에 원하는 분석기능들이 제대로 작동할 수 있으리라 생각한다.

또한 어느 분야에서도 마찬가지로 이지만 데이터 마이닝은 데이터를 분석하기 위한 하나의 방법 혹은 도구이지 만능은 아니다. 따라서 이러한 시스템을 도입하고 지속적인 발전을 위한

노력을 병행하고, 각종 옵션을 변경/선택하며 여러 가지 시각으로 분석을 하지 않고 시스템을 도입했으니 시스템이 모든 것을 해결해 줄 것이며, 시스템에서 제시되는 결과가 정답이라 생각하고 시스템에만 의존 한다면 많은 오류를 양산할 가능성이 존재한다.

추후 연구에서는 본 논문을 통하여 제시된 방법과 현재 구축된 시스템을 어느 일정기간 동안 운영 결과를 통하여, 향후 보다 정확하고 효율성있는 품질마이닝 시스템에 대한 보완점을 제시하고자 한다.

variation diagnosis", *Journal of materials processing technology*, Vol. 129, pp 315~320.

References

- [1] 백준결, 김강호, 김성식, 김창욱(2000), "실시간 기계상태 데이터베이스에서 데이터 마이닝을 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] 이현우, 남호수, 강중철(2005), "A Study on Data Mining Application Problem in the TFT-LCD Industry", *한국데이터정보과학회지*, V16, No4, pp 823~833.
- [5] 장남식(1999), "성공적인 지식경영을 위한 핵심정보 기술 : 데이터마이닝", *대청미디어*
- [6] Braha, D., and Shmilovici, A.(2002), "Data mining for improving a cleansing process in the semiconductor industry," *IEEE transactions on semiconductor industry*, Vol. 15, No. 1, pp91~101
- [7] 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.
- [8] 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