

國立交通大學

電機與控制工程學系

碩士論文

應用貝氏網路於

工業製程之診斷與預測



Application of Bayesian Networks to
Diagnosis and Prognosis of Manufacturing Process

研究生：林志憲

指導教授：周志成 副教授

中華民國九十三年九月

應用貝氏網路於
工業製程之診斷與預測

Application of Bayesian Networks to
Diagnosis and Prognosis of Manufacturing Process

研究生：林志憲

Student : Chin-Hsien Lin

指導教授：周志成

Advisor : Chi-Cheng Jou

國立交通大學



A Thesis

Submitted to Department of Electrical and Control Engineering

College of Electrical Engineering and Computer Science

National Chiao-Tung University

in Partial Fulfillment of the Requirements

for the Degree of Master

in

Electrical and Control Engineering

September 2004

Hsinchu, Taiwan, Republic of China

中華民國九十三年九月

應用貝氏網路於 工業製程之診斷與預測

學生：林志憲

指導教授：周志成

國立交通大學電機與控制工程學系碩士班

摘 要

工業製程中常常有上百個步驟，每個步驟都會包含相當多的測量項目，所以我們會得到相當龐大的原始數據。而這些數據中包含著一些可能導致製程異常的原因或造成產品異常的因素，因此我們選擇使用聯合樹模型的方法來建立模型，提供工程師或分析師來分析根據聯合樹診斷出來的異常狀態。

本論文以晶圓製程為例，將資料先經過一些處理建立聯合樹模型和應用聯合樹，接著就影響應用的因素作討論，討論樣本數量的影響和變數數量的影響並測試其計算時間跟計算的極限，接著討論被解釋變數分類的問題，最後驗證模型是否正確。

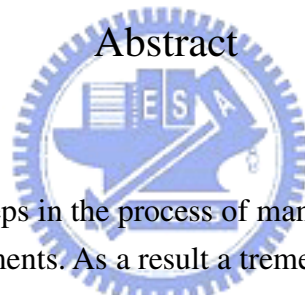
同樣的方法可應用在很多領域上，如類似製程的自動化科技上、大量病歷資料的醫學上、大量數據的氣象上...等等各種需要處理大量資料的領域，都可以應用此方法來進行診斷和預測的工作。

Application of Bayesian Networks to Diagnosis and Prognosis of Manufacturing Process

Student: Chin-Hsien Lin

Advisor : Chi-Cheng Jou

Department of Electrical and Control Engineering
National Chiao Tung University



There are hundreds of steps in the process of manufacturing operation. Every step contains lots of measurements. As a result a tremendous amount of data is available. These data maybe contain reasons that case abnormal states of manufacturing process. We use Junction Tree Algorithm to establish Junction Tree Models, in order to provide engineers or analysts to diagnose abnormal states.

Take manufacture of silicon wavers for example. After establishing the junction tree model, we use the model to find abnormal states. Then we will discuss factors such as number of cases and variables and the limit of calculation time and the quantization of yield variable. At last we will verify the junction tree model.

This method also can be applied to many areas that need to handle a tremendous amount of data, for example in the industrial, medical or meteorological area etc.. We can use this method to diagnose and prognose results after the junction tree is built.

誌 謝

首先感謝指導教授周志成老師在學習的過程中給予協助與鼓勵，除了讓我在專業領域上有更深刻的體會外，在待人處世和為學態度方面也都獲得相當大的啟發，在此謹致上最誠摯的敬意與謝忱。

另外要感謝實驗室同學郭宇豪與郭光華提供研究上的意見與經驗，還有同學黃得裕與邱俊傑在上活上的互相幫忙。

最後感謝家人的支持與關懷，讓我能專心順利完成學業，謹以此篇論文獻給你們。



林志憲 于新竹

中華民國九十三年九月

目 錄

中文摘要.....	i
英文摘要.....	ii
誌謝.....	iii
目錄.....	iv
表列.....	vi
圖列.....	ix
第一章 緒論.....	1
1.1 研究動機與目的.....	1
1.2 研究方法.....	1
1.3 論文架構.....	6
第二章 貝氏網路簡介.....	7
2.1 機率.....	7
2.1.1 貝氏原理.....	7
2.1.2 連鎖貝氏原理.....	7
2.1.3 條件獨立.....	8
2.1.4 聯合機率.....	9
2.2 貝氏網路.....	9
第三章 聯合樹演算法.....	13
3.1 轉換.....	14
3.1.1 Moral Graph.....	14
3.1.2 三角化圖形.....	15
3.1.3 分群圖形.....	19
3.1.4 聯合樹.....	23

3.2	傳播	24
3.2.1	位能初始化	24
3.2.1.1	建立初始化位能	25
3.2.2	傳播位能	27
3.2.2.1	傳遞訊息	28
3.2.2.2	全體傳播	29
3.2.2.3	排斥	31
3.2.3	更新機率	31
3.2.3.1	將證據編碼成一個近似函數	32
3.2.3.2	擁有觀察資料之位能初始化	33
3.2.3.3	輸入觀察資料於網路中	34
第四章	實驗與結果	35
4.1	實驗資料描述	35
4.2	模型發展程序	35
4.2.1	特徵選取	35
4.2.2	圖形化模型	36
4.2.3	節點排序	37
4.2.4	量化	37
4.2.5	統計計算	40
4.2.6	聯合樹	40
4.3	模型的應用	41
4.4	影響應用因素之探討	52
4.4.1	樣本數分析	52
4.4.2	變數個數分析	54
4.4.3	良率變數量化分析	56
4.4	模型驗證	58

第五章 結論60

參考文獻.....62



表 列

表1-1 量化.....	5
表3-1 分群 $\{a, d\}$ 的排斥位能.....	28
表3-2 一個簡單的訊息傳遞.....	29
表3-3 排斥的例子.....	32
表4-1 特徵選取（對良率變數比較有關係的變數11個）.....	36
表4-2 節點排序.....	37
表4-3 對應變數名稱.....	37
表4-4 8個變數經k-means分三類，每各類別的中心數值.....	39
表4-5 條件機率 $P(V_3 V_1, V_2)$	39
表4-6 良率變數量化結果.....	43
表4-7 條件機率 $P(V_1 V_8)$	44
表4-8 條件機率 $P(V_2 V_8)$	44
表4-9 低良率跟變數最有關係的狀態.....	45
表4-10 (a)更新過後的變數的事前機率.....	48
表4-10 (b)更新過後的良率機率值.....	48
表4-10 (c)良率期望值.....	48
表4-11 變數對良率的影響力排名.....	48
表4-12 $V_1 \sim V_5$ 的異常狀態.....	49
表4-13(a) 去除 V_1 和 V_2 不同組合的狀態所得的良率、良率期望值和改良率.....	50
表4-13(b) 去除 V_1 和 V_2 不同組合的狀態所得的良率、良率期望值和改良率.....	51
表4-14 50、100和152個樣本數的良率變數於k-means三個類別的平均值（或中心點）.....	52

表4-15 三個樣本數量不同的低良率跟變數最有關係的狀態	52
表4-16 三個樣本數量不同的變數對良率的影響排名.....	52
表4-17 (a)：低良率與每一個變數最有關係的狀態（3~6個變數）	54
表4-17 (b)：（調整過後）良率期望值（3~6個變數）	54
表4-17 (c)：影響良率的變數排名（3~6個變數）	54
表4-18 分析時間.....	55
表4-19 k-means分類結果和事前機率和良率期望值	56
表4-20 兩種良率分類中良率變數中的L狀態跟其他變數中最有關係的狀態.....	57
表4-21 良率期望值.....	57
表4-22 良率分類的變數對良率的影響力排名	57
表4-23 跟低良率比較有關係的變數狀態.....	59
表4-24 $V_s = L$ 的情形之下，其他變數的狀態.....	59
表4-25 跟低良率比較有關係的變數狀態.....	59



圖 列

圖1-1 知識工程：專家知識轉化為電腦程序.....	3
圖1-2 研究方法.....	4
圖1-3 (a) 無方向的圖形；(b) 方向性非循環圖形.....	5
圖2-1 簡單方向性非循環圖.....	10
圖2-2 全體變數聯合機率.....	10
圖2-3 簡單診斷例子.....	11
圖3-1 聯合樹演算法流程.....	13
圖3-2 左：貝氏網路圖形；右：moralized graph.....	15
圖3-3 未三角化的非方向性圖形.....	16
圖3-4 運用最大基數搜尋演算法將邊線填入.....	18
圖3-5 運用最大基數搜尋演算法將節點編號.....	18
圖3-6 另外兩種運用最大基數搜尋演算法將節點編號.....	19
圖3-7 三角化的圖形.....	20
圖3-8 分群.....	22
圖3-9 分群圖形.....	22
圖3-10 聯合樹.....	24
圖3-11 貝氏網路.....	26
圖3-12 將圖3-11轉換成樹狀結構聯合樹的圖形.....	26
圖3-13 初始化分群{A,C,E}跟交集{C,E}.....	27
圖3-14 全體傳播.....	30
圖3-15 有觀察資料的聯合樹演算法流程.....	33
圖4-1 圖形化模型.....	36
圖4-2 貝氏網路.....	38

圖4-3 Yield量化結果	38
圖4-4 moralized graph (也是三角化圖形)	40
圖4-5 聯合樹	41
圖4-6(a) V_1 和 V_2 的資料分佈情況.....	45
圖4-6(b) V_4 和 V_5 的資料分佈情況	46
圖4-6(c) V_6 和 V_7 的資料分佈情況.....	46
圖4-7 分析時間.....	55

