

**DIAGNOSIS USING BAYESIAN NETWORKS TO ENABLE VALUE-ADDED  
TROUBLESHOOTING TO MANUFACTURING PROCESS PROBLEM**

**LOKMAN MOHD FADZIL**

**UNIVERSITI SAINS MALAYSIA**

**2007**

**DIAGNOSIS USING BAYESIAN NETWORKS TO ENABLE VALUE-ADDED  
TROUBLESHOOTING TO MANUFACTURING PROCESS PROBLEM**

by

**LOKMAN MOHD FADZIL**

**Thesis submitted in fulfillment of the  
requirements for the degree  
of Master of Science**

**SEPTEMBER 2007**

## **ACKNOWLEDGEMENTS**

All gratitude and praise be to Allah, the Almighty who granted the author unlimited bounties and grace with which this work becomes a reality.

The author wishes to thank his beloved parents (Haji Mohd Fadzil Haji Saidon & Hajjah Sharifah Haji Noor) for their encouragement, care and prayers. Love and special thanks to his lovely wife, Nor Azimah Ismail, daughters Siti Farah, Siti Farhana, Siti Farhah, and Siti Farzana, sons, Ahmad Farhan, and Ahmad Faiz Syakirin, for their patience, sacrifice, understanding and support, especially when he needs them most, and also to dedicate in memory of deceased daughter Siti Farhain.

The author also wishes to express his utmost gratitude to his supervisor, Dr. Zahurin Samad, ex-supervisors, Dr. Wan Mansor Wan Muhamad, and Encik Jamaluddin Abdullah, and Dr. Azman Samsuddin, for their continuous guidance, ideas, and support. The author also feel indebted to lecturer Adam Baharum, student assistants, Nik Nor Hishamul Azma @ Azman, Tengku Azeezoo, Abdul Mannan Mansor and Armizal Azwan Alias, Universiti Sains Malaysia, particularly USM Institute of Postgraduate Studies, USM School of Mechanical Engineering, USM technical support personnel, Hishamuddin Endan, and Mohd Ali Shahbana Mohd Raof, fellow students, Wan Azhar Wan Yusoff, Mohd Nizam Ahmad, Zakri Ghazali, and Mohd Zulkarnain, and those who are not able to be named here, who have provided valuable assistance for completion of this thesis and research, and made it possible for the author to complete his graduate studies.

## TABLE OF CONTENTS

	Page
<b>ACKNOWLEDGEMENTS</b>	ii
<b>TABLE OF CONTENTS</b>	iii
<b>LIST OF TABLES</b>	iv
<b>LIST OF FIGURES</b>	v
<b>LIST OF SYMBOLS</b>	v
<b>LIST OF ABBREVIATION</b>	vi
<b>LIST OF PUBLICATIONS &amp; SEMINARS</b>	vi
<b>ABSTRAK</b>	vii
<b>ABSTRACT</b>	ix
<b>CHAPTER ONE : INTRODUCTION</b>	
1.0 Introduction	1
1.1 Problem Statement	4
1.2 Research Objectives	5
1.3 Scope and Limitation	6
1.4 Approach	8
1.5 Thesis Outline	12
<b>CHAPTER TWO : LITERATURE REVIEW</b>	
2.0 Introduction	13
2.1 Failure and Downtime	27
2.2 Artificial Intelligence Techniques	32
2.3 Theoretical Considerations	37
2.4 Summary of Literature Review	45
<b>CHAPTER THREE : THEORETICAL FRAMEWORK</b>	
3.0 Introduction	47
3.1 Data Collection Stage	49
3.2 Data Algorithm Stage	57
<b>CHAPTER FOUR : DEVELOPMENT OF THE DIAGNOSIS SYSTEM</b>	

4.0	Introduction	71
4.1	Data Processing Stage	72
	4.1.1 Conversion to Standard Troubleshooting Steps	72
4.2	Data Analysis Stage	80
	4.2.1 Troubleshooting Steps Mapping into Bayesian Networks Diagram	81
4.3	Summary	90
 <b>CHAPTER FIVE : RESULTS AND DISCUSSION</b>		
5.0	Comparison	91
5.1	Results	91
5.2	Discussion	96
 <b>CHAPTER SIX : CONCLUSION AND RECOMMENDATIONS</b>		
6.0	Conclusion	99
6.1	Recommendations	100
 <b>BIBLIOGRAPHY</b>		102
 <b>APPENDIX A</b>		108

## LIST OF TABLES

	Page	
2.1	Branches of Artificial Intelligence (Jones 2003)	36
2.2	Two of the Five Probability Tables (Buntine 1996, p.196)	41
2.3	Conditional Probabilities P (K P1, P2) For All Possible Values (Scheiterer and Obradovic 2005)	44
3.1	List of Selected Failure Modes with respective MTTR, Downtime and Occurrence	57
3.2	Raw Equipment Downtime Data Source for BIN 9 (Intel Corporation ca. 2003, approx. 18 lines)	60
3.3	Raw Equipment Downtime Data Source for BIN 18 (Intel Corporation ca. 2003, approx. 14 lines)	62
3.4	Raw Equipment Downtime Data Source for BIN 21 (Intel Corporation ca. 2003, approx. 31 lines)	64
4.1	Sample Table for Standard Troubleshooting Steps With The Degree Of Completion	74
4.2	Bin 18 Standard Troubleshooting Steps With Respective Degree Of Completion	75
4.3	Bin 9 Standard Troubleshooting Steps With Respective Degree Of Completion	76

4.4	Bin 21 Standard Troubleshooting Steps With Respective Degree Of Completion	77
4.5	Bin 18 Standard Troubleshooting Steps Occurrences And Present Percentage	80
4.6	Bin 18 Belief Report Generated From Netica™ Software	83
4.7	Bin 9 Standard Troubleshooting Steps Occurrences And Present Percentage	86
4.8	Bin 9 Belief Report Generated From Netica™ Software	87
4.9	Bin 21 Standard Troubleshooting Steps Occurrences And Present Percentage	88
4.10	Bin 21 Belief Report Generated From Netica™ Software	90
5.1	Comparison Table for Netica™-generated Bin 18 Troubleshooting Steps vs. Existing Troubleshooting Guide	93
5.2	Comparison Table for Netica™-generated Bin 9 Troubleshooting Steps vs. Existing Troubleshooting Guide	94
5.3	Comparison Table for Netica™-generated Bin 21 Troubleshooting Steps vs. Existing Troubleshooting Guide	95

## LIST OF FIGURES

	Page	
1.1	Universal Data Feedback Model (Kindree, et al. 1994, p.67)	7
2.1	A Sample Token Ring Local Area Network (LAN) (Bentlage et al. 2001, p.216)	17
2.2	Block Diagram of Decision Flow	21
2.3	Simple Bayesian model	39
2.4	Another Simple Bayesian model (Buntine 1996, p.196)	40
2.5	Bayesian Multiple Fault Model	42
2.6	Cause-To-Effect Model Of Simple Domain With Two Problems And A Single Common Symptom (Scheiterer and Obradovic 2005)	44
3.1	A Typical product lifecycle phases for semiconductor manufacturing (Geng 2005, p 301)	52
3.2	Sample Test Program Flow for Bin 9, Bin 18 and Bin 21 (Intel Test Methodology Handbook 2002, p.101)	54
4.1	Summary Flow Chart of the Research Methodology	71
4.2	Bin 18 Netica™ Bayesian Networks Diagram	83
4.3	Bin 9 Netica™ Bayesian Networks Diagram	86
4.4	Bin 21 Netica™ Bayesian Networks Diagram	89

## LIST OF SYMBOLS

	Page	
1.1	Π - PI	38-39

## LIST OF ABBREVIATION

	Page
1.1 SPC - Statistical Process Control	17, 20
1.2 HTTP – Hyper Text Transfer Protocol	16, 17
1.3 LAN - Local Area Network	16, 17
1.4 AI - Artificial Intelligence	32
1.5 AVI - Automated Visual Inspection	33
1.6 MTTR – Mean Time to Repair	51

## LIST OF PUBLICATIONS & SEMINARS

- 1 Wan Muhamad, W. M., Fadzil L. M., and Samsudin A. (2000) The Application Of Statistical Methods Using Web Infrastructure To Enable World Class Manufacturing Processes. RESQUA2000 Regional Symposium On Quality & Automation Proceedings, 1st, 113-117.
- 2 Wan Muhamad, W. M., Subari K, and Fadzil L M (2002). A Strategy Towards Global Competitiveness: Internet-Based Quality Assessment. Conference On Quality Proceedings, 77.

# DIAGNOSA MENGGUNAKAN KAEDAH BAYESIAN NETWORKS UNTUK MENGUPAYAKAN BAIKPULIH KE ATAS PERMASALAHAN PROSES PEMBUATAN

## ABSTRAK

Di dalam proses perkilangan, juruteknik selalunya melakukan pembaikpulihan mesin-mesin pembuatan menggunakan piawaian Panduan Pembaikpulihan yang sedia ada. Apabila mesin-mesin pembuatan menjadi lebih canggih dan kerosakan mesin-mesin yang berkaitan menjadi bertambah sukar untuk dianalisa dan diselesaikan, pengalaman juruteknik menjadi bertambah penting bagi menggantikan piawaian Panduan Pembaikpulihan di dalam membaikpulih kegagalan dan kerosakan mesin-mesin yang kompleks dan serious. Penyelidikan ini bertujuan mendefinisi dan mengaplikasikan satu kaedah algorithma untuk menterjemahkan data pembaikpulihan mesin-mesin pembuatan kepada rangkaian langkah-langkah pembaikpulihan yang piawai. Algorithma ini membantu menentukan setiap langkah-langkah pembaikpulihan yang dihasilkan mempunyai maklumat kebarangkalian dan rangkaian untuk diproses oleh Netica™, satu perisian Bayesian Networks. Perisian ini akan memberikan cadangan rangkaian langkah-langkah pembaikpulihan berdasarkan maklumat yang diberikan. Perbandingan dibuat di antara parameter yang umum seperti Waktu Rosak (Downtime), Purata Waktu Untuk Baiki (MTTR) dan Peratusan Kekesanan (Percentage Effectiveness) dengan kes-kes pembaikpulihan yang terdahulu yang menggunakan piawaian Panduan Pembaikpulihan yang sedia ada untuk meninjau keberkesanan teknik tersebut. Keputusan menunjukkan piawaian Panduan Pembaikpulihan yang dihasilkan oleh Netica™ memberikan peningkatan dari segi Purata Waktu Untuk Baiki sebanyak 12.67% untuk Bin 18, 32.43% untuk Bin 9, 34.93% untuk Bin 21; peningkatan dari segi Waktu Rosak sebanyak 21.89% untuk Bin 18, 39.10% untuk Bin 9, 35.42% untuk Bin 21; peningkatan dari segi Peratusan Kekesanan tidak berubah untuk Bin 18, 26.67% untuk Bin 9, 28.24% untuk Bin 21, berbanding dengan piawaian Panduan Pembaikpulihan yang sedia ada. Satu trend/turunan yang dilihat meningkat di

kalangan parameter-parameter tersebut menunjukkan potensi cadangan rangkaian langkah-langkah pembaikpulihan sebagai Panduan yang praktikal untuk kegunaan juruteknik.

# **DIAGNOSIS USING BAYESIAN NETWORKS TO ENABLE VALUE-ADDED TROUBLESHOOTING TO MANUFACTURING PROCESS PROBLEM**

## **ABSTRACT**

In current manufacturing process, technicians normally perform equipment troubleshooting using standard Troubleshooting Guides. As equipment gets more sophisticated and associated failures become more difficult to analyze and to solve, experience becomes a better substitute than existing Troubleshooting Guides in troubleshooting complex failures. This research explores the definition and application of an algorithm to translate historical maintenance repair data to generate sequence of standardized equipment troubleshooting steps. The algorithm help determine each troubleshooting steps probability and sequence information to be processed by Netica™, a Bayesian Networks modeling software. The software produces recommended sequence of troubleshooting steps based on the given information. Comparison is made between common parameters i.e. Downtime, MTTR, and Percentage Effectiveness data with repair cases using existing Troubleshooting Guides. Results show that Netica™-generated troubleshooting steps represent an MTTR improvement over existing Troubleshooting Guide steps of 12.67% for Bin 18, 32.43% for Bin 9, 34.93% for Bin 21, respectively; Downtime improvement of 21.89% for Bin 18, 39.10% for Bin 9, of 35.42% for Bin 21, respectively; % Effectiveness improvement of zero change for Bin 18, of 26.67% for Bin 9, of 28.24% for Bin 21, respectively over the troubleshooting steps using existing Troubleshooting Guide. In summary, an improving trend seen across these parameters seen across the failure subgroups shows potential for usage as practical troubleshooting steps for use by maintenance technicians.

# CHAPTER 1 INTRODUCTION

## 1.0 Introduction

In Introduction, a brief discussion touches on key motivating factors for this research on developing methods and algorithms for resolving equipment failures using Bayesian Networks model. To begin with, there are rising trends in industries, which demonstrate the importance of making informed decision-making by having highly structured data infrastructure and the method to process and utilize the data in the data repository. This transition to structured data infrastructure begins with the transformation of corporate information from paper in its various forms into digital format. The new digital platform enables companies to start engaging promising business models over traditional ones (Schneier 1994). Significant enhancements in technology have profoundly changed virtually every type of business, and the way people do business. In other words, improvement in computers and communication technologies brings up new businesses opportunities (Brynjolfsson and Kahin, 2000).

By installing sophisticated communications and technology systems to provide value-added businesses information, a company increases its ability to make better and more informed business decisions. As time changed, and computer networks become widely used, standalone computers or workstations started to connect in form of networks and the digital data inside those computers shared among employees in the enterprise. Eventually, the digital data traffic started to increase, not just within the confines of the company, but also outside the company to its customers and suppliers through the Internet or World Wide Web as a method to disseminate information.

Due to increasing connectivity between computers and the information infrastructure, previously untapped information for businesses now becomes readily

available (Pesko 1999, p.11). As this information is critical to get job done and to enable businesses to seize new opportunities, it creates demand for appliances with specialized access to information—hence the name “information appliance” (Want, and Borriello, 2000, p.24). This poses new challenges and opportunities to businesses.

These challenges include low cost and technology advantage that are important to enable companies to successfully compete with and force other players out of the market (Masatsugu 2000, p.178). Slywotzky (1999, p.94) noted that computer industry is an excellent example where companies i.e. Compaq reinvented new models to achieve cost and technology leadership to drive giants i.e. IBM out of PC markets. To maintain competitiveness, technology and cost are key factors to today’s companies.

Miltenburg (2005) cited a case study on Rolls-Royce which used to be the most advanced engineering company in the world but eventually lost its cost and technology leadership. Companies realize the need to promote data exchange for efficient decision-making to drive down cost and introduce new technologies. In turn, effective data flow across their supply chains will ensure that business information communicated effectively and timely to the correct audience.

Organizations without well-defined data infrastructure (Werner and Hermansson 2002, p.131) will find it difficult to maintain and control their own processes, resources and inventory in rapidly changing technology industries. As Bentlage et al. (2001, p.215) stated, it is not possible to share data between processes, customers and suppliers, merge data for analysis, and scrutinize the data for improvements.

The process to enable making informed decisions starts with Data Collection Process, where data originating from a certain process is collected. The flow continues with Data Analysis Process where the collected data is processed and analyzed for

patterns or trends. The analyzed data is used as a base for making decisions under Decision-making Process as demonstrated by Foy (1996, p.25).

Making good decisions is a necessity in maintenance business as well to enable better troubleshooting decisions and higher-performing equipment. Luxhoj et al. (1997, p.437) pointed out that changing manufacturing requirements today makes it imperative for maintenance management to contribute to cost and service improvements to the organization.

With this in mind, Luxhoj et al. (1997, p.437) and Tu and Yeung (1997, p.453) reviewed two major maintenance-benchmarking studies to understand current issues in maintenance industry. The purpose of the Scandinavian and US benchmarking studies was to identify issues and general maintenance trends for improving maintenance.

The survey indicates that industrial firms' maintenance costs have increased with Scandinavian firms at 0.5% per year and US firms (10-15%). Wireman (1990) noted that the total "waste" in excessive maintenance expenditures was approximately 200 billion dollars in 1990, which equaled the total maintenance costs in 1979.

Wireman (1990) concluded that the survey results illustrate the need for better maintenance planning and the need for more maintenance research and development. One of the key needs identified in the studies includes the move toward computer-based maintenance systems as the result of emerging developments and advances in maintenance technology, information and decision technology, and maintenance methods. This requires development of expert systems or decision support tools, as advocated by Werner and Hermansson (2002, p.131) to support equipment maintenance that will be key driving forces to address issues in maintenance industry.

## **1.1 Problem Statement**

Based on a brief introduction on maintenance benchmarking and case studies mentioned in Introduction, the problem statement summarized as “a need to aid current troubleshooting process using artificial intelligence techniques to generate more accurate decisions to solve a tester equipment non-productive downtime issues.” This research attempts to address an issue when using the troubleshooting guides to diagnose equipment issues. The guides are static information designed to fix specific failures and cannot be utilized to resolve failures different from those defined in the guides. As different machines perform differently over time, the ongoing troubleshooting information captured by the technicians might reveal certain patterns or trends that might be able to help troubleshoot future failures.

According to Bloch, H.P. (2001 p.74), the causes of all process-machinery failures, without exception, can be put into one or more of seven categories: faulty design, material defects, fabrication or processing errors, assembly or installation defects, off-design or unintended service conditions, maintenance deficiencies, whether procedural or due to neglect, and improper operation. Kindree, et al. (1994, p.66) provides a similar troubleshooting concept but the focus is more on the 5 phases of machine and equipment life cycle, namely concept, development design, build and install, operation and support, and conversion and/or decommission, where upfront engineering diagnosis and continuous improvement done in each of the process.

For this research, the tester equipment non-productive downtime issues defined under the problem statement mostly focus on material defects with some issues on assembly or installation defects.

## 1.2 Research Objectives

The main objective of this research is to develop a technique to diagnose or troubleshoot non-productive downtime issues for tester equipment used in a manufacturing process to enable value-added troubleshooting by using Bayesian Networks.

As reflected by the Problem Statement in Section 1.1, the benchmarking studies identifies an established need based on increasing maintenance costs, high unforeseen maintenance, and high ratio of lost production due to maintenance cost for Scandinavian and United States organizations (Luxhoj et al. 1997, p.437). The benchmarking studies clearly recommended using artificial intelligence techniques to generate decisions that are more accurate. In doing so, troubleshooting process can be improved and made more effective in solving tester equipment non-productive downtime issues.

To achieve the main objectives of this research, the detailed research objective are defined as follows:

- Develop a model using Bayesian Networks technique by using existing real-world repair or maintenance data to produce practical troubleshooting steps for use by maintenance technicians
- Utilize downtime data embedded in the repair or maintenance data to compare troubleshooting steps generated from Netica™ software with existing Troubleshooting Guide currently in use to evaluate the effectiveness of this technique

### **1.3 Scope and Limitation**

The scope of this research is limited to the use of the data source derived from the content of a Web-based tester equipment downtime database, accessible at <http://e-db.png.intel.com/> developed by Intel Corporation for ten production machines for the year 2003.

The database is used to capture or document failures that occur during a particular tester equipment operation. How a tester equipment fails and how the failures are captured can be explained per the generic high-level Universal Data Feedback Model process, defined by Kindree, et al. (1994, p.67). This process comprises a close-loop system of various functions. The first function is the data collection function where the tester equipment failure information is captured. The second function is the analysis function where failure information is analyzed. The third function is the feedback function where recommendations are provided to fix the tester equipment issues and improve the tester equipment reliability and maintainability. In addition, it also specifies how to handle failure in a typical tester equipment process, as shown in Figure 1.1.

While the tester equipment is in operation (shown as 'Equipment Operation'), a failure happens (shown as 'Equipment Failure'). A decision box poses a question whether to replace the tester equipment part, with a subsequent need to document the failure. This process step, known as 'Document the Failure', produces the data source used in this research.

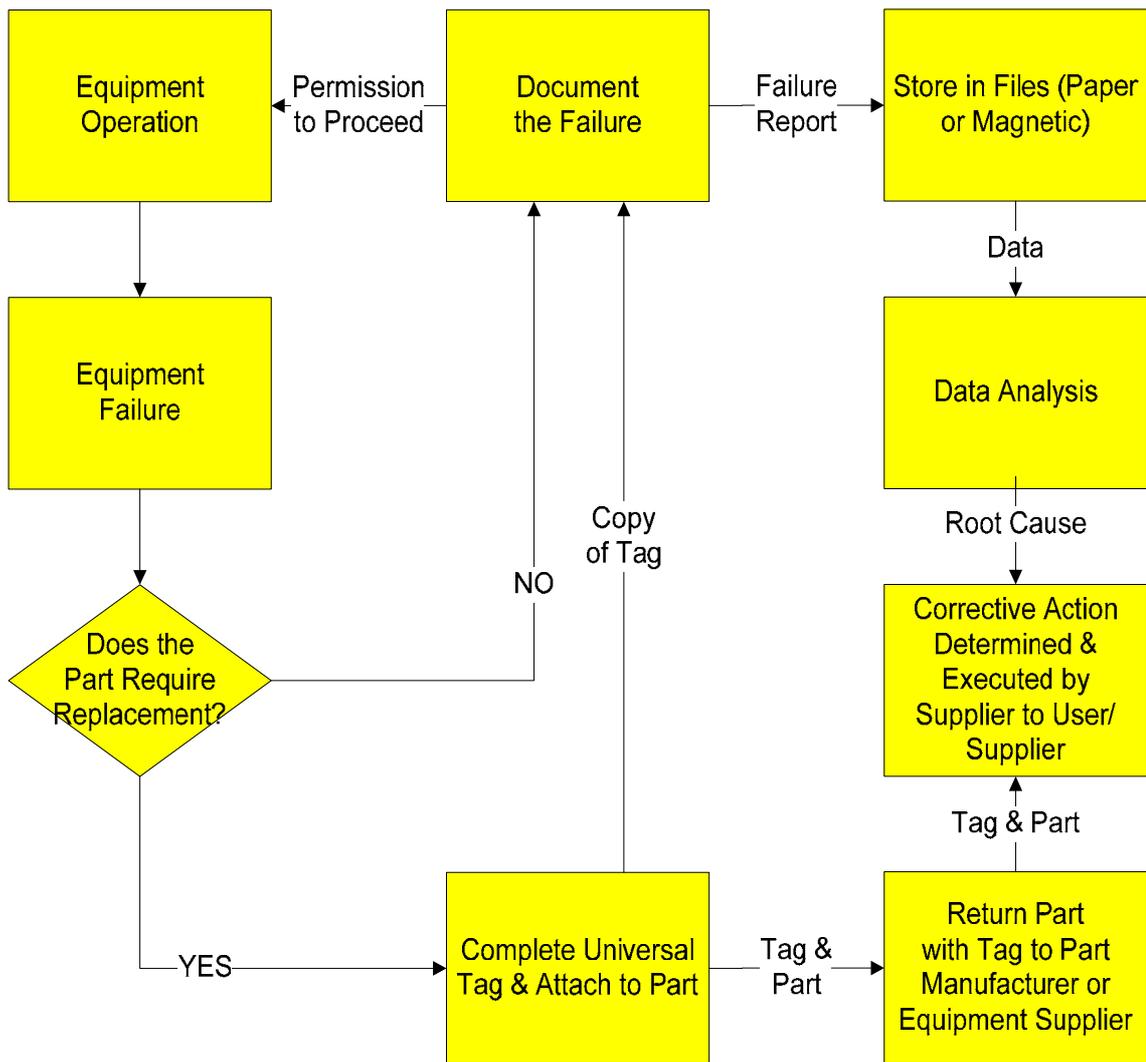


Figure 1.1 Universal Data Feedback Model (Kindree, et al. 1994, p.67)

This process of capturing the tester equipment failure information, analyzing failure information, and providing feedback to fix the tester equipment issues and improve the tester equipment reliability and maintainability proposed by Kindree, et al. adequately simulates a real-world manufacturing process. During a particular tester equipment operation, a problem may occur unexpectedly, which either will stop the process from running entirely or cause the process not to run per the expected performance specifications. The operator of the machine running this process will then stop the machine and notify the technician to fix the problem.

The technician will attempt to diagnose the problem, determine, and execute the most appropriate troubleshooting steps to fix the problem. Bloch, H.P. (2001 p.74) cited the strategy of rational thinking, referring to equipment documentation, and occasional recollection of the simpler laws of physics, to result in failure identification, and point to future failure avoidance. In a factory where proper the tester equipment record keeping are maintained, the technician will record down the machine information, failure signature, and the troubleshooting steps to fix the problem for future reference.

This research is only limited to 'Document the Failure' process step where the tester equipment failure information is being captured and 'Data Analysis' process step where the tester equipment failure information is being analyzed. The process to define troubleshooting steps recommended to fix the tester equipment issues is similar to 'Corrective Action Determined & Executed by Supplier to User/Supplier' process step.

However, in this step, the troubleshooting steps recommendations are provided to the tester equipment technician, not to the User/Supplier as described in the process step. All the other process steps, i.e. the tester 'Equipment Operation', 'Equipment Failure', Store in Files (Paper or Magnetic), Return Part With Tag to Part Manufacturer or Equipment Supplier', Complete Universal Tag & Attach to Part' and 'Does the Part Require Replacement?' decision box are not included in this research. In short, the scope and limitations are clearly defined in this research.

#### **1.4 Approach**

This research approach is divided into several milestones. The first step is to analyze current issues facing the tester equipment in manufacturing processes. This is accomplished by reviewing the current literature on the topic.

Whitney (2004) noted that all current maintenance issues apply to any one of three major approaches concerning equipment: strategic, technical, and economic. The strategic issues focus on choice of method of accomplishing the manufacturing — manual, robotic, and so on — plus part presentation, flexibility, inspection, and throughput. The technical problems involve detailed technology choice and assurance of proper performance, mainly achieved via an error analysis. Economic analysis is concerned with choosing a good combination of alternative methods of achieving assembly and controlling error. The focus of this research work will be more on technical standpoint as selection of techniques is a matter of technology choice. Proper performance will be monitored by a set of pre-defined metrics (to be explained in Chapter 4).

As current issues are understood, the second step is to determine the focus of research work to be undertaken to address one of the issues selected for study. Once the research focus has been determined, the third step is to understand other researches done in this field of study and comprehend how those researchers address the issues. At this point, a research is proposed that tackles a different aspect of the issue or propose improvements on existing research.

The fourth step is to make use of existing data collected on target tester equipment in the manufacturing process. The process of selecting a technique to define an algorithm to process the collected data and to subject to a model to predict systematic troubleshooting guides is considered. This output of predicted troubleshooting steps is compared to the existing manual troubleshooting guide using pre-defined criteria to determine the effectiveness of the technique. The pre-defined criteria for comparison will indicate whether the so-called the Bayesian Networks

models is better at predicting appropriate troubleshooting steps than existing manual troubleshooting guide to solve a particular tester equipment problem.

The data source derived from the content of a Web-based tester equipment downtime database is extracted into Microsoft® Excel-based files. All the tester equipment failure information is captured per the 'Document the Failure' process step in Section 1.3 in these files. The decision to limit to a number of so-called failure data bins for particular tester equipment is due to several factors to be explained further in Chapter 3.

Commenting on this, Liu and Desmarais (1997 p.991) agreed that constructing a valid knowledge representation is a time-consuming task and there are issues on insufficient empirical or meaningful data and/or the complexity involved in the network induction in real-life applications. Jäger and Bertsche (2004 p.91) also cautioned that the accumulated information might be very imprecise If the amount of information becomes quite large.

Buntine (1996, p. 200) also agreed on the difficulties in determining the number of cases required for the sample, which is referred to as Sample Complexity, and the time or space required for optimizing the sample into the model, which is referred to as Computational Complexity. Buntine quoted the computational learning theory where there are roughly three distinct phases depending on the quantity of cases obtained for the sample, namely the small sample, the medium sample, and the large sample phases. Initially with a small sample, Buntine used the term 'learning' or absorption of information to correspond to one's biases or prior information. With a large sample, learning is close to the "true" model possible with high probability. Since an error rate known as Bayes optimal error rate is inherent in all three different types of samples,

this indicates that using a small sample is possible. From literature, Bayes optimal error rate is proven not to influence the validity of the sample data.

These files are in the form of spreadsheet that comprises a number of columns. The listing contains the associated tester equipment type and number, the root cause of equipment failure, equipment type and number (secondary supporting equipment), equipment downtime and up-time dates and times. It also contains Mean Time to Repair (or MTTR) in hours, down-time (or D/T) in hours, equipment failure mode, the employee number of the technician who attended to the problematic equipment, technician work shift hours, and comments or troubleshooting steps taken by the technician to fix the problem.

From these files, the Failure Mode column (column M) is filtered to reflect a list of output binning called Bin 9, Bin 18, and Bin 21 (to be explained in Chapter 3). They will be used to be taken as sample of the equipment failure mode that contribute to the downtime are taken for further analysis to determine the standard troubleshooting steps taken to resolve the problem.

A Theoretical Framework for the algorithm to process the data for the model will be defined. Once a systematic troubleshooting has been determined, this data is plugged into the Norsys Software Corp Netica™ Bayesian Networks software that will recommend the most effective troubleshooting set of steps to rectify the problem based on evidence using Bayesian Networks algorithm. The troubleshooting set of steps derived from the Bayesian Networks model will be compared against the existing troubleshooting steps found in the guides.

## **1.5 Thesis Outline**

This thesis is organized into six main chapters: Chapter 1 starts with Introduction that briefly explains the challenges faced by companies doing business on a global level, especially on technology and cost, and results of maintenance benchmarking studies that advocates effective manufacturing equipment management and usage of data for decision-making in manufacturing process improvement. Chapter 2 focuses on Literature Review with an academic treatment and definition of the ideas, terminology, and equations for Failure and Downtime, Artificial Intelligence Techniques and Theoretical Considerations for Bayesian Networks used in this research.

Chapter 3 describes the Theoretical Framework or the definition on the algorithms and the data manipulation steps necessary to process the raw data for the proposed system. Chapter 4 deals with the Development of the Diagnosis System on how this research is pursued together with the assumptions, processing and transforming the data in the model for analysis. Chapter 5 discusses on Comparison and Results where tabulated data is compared and interpreted, and wraps up with Discussion where analysis made on the data is being discussed. In Chapter 6, a Conclusion is made to support the objectives of this research and Recommendation for future proposal is included to provide future directions for this research.

## **CHAPTER 2**

### **LITERATURE REVIEW**

#### **2.0 Introduction**

Literature review is divided into four sections in this chapter. The first part is on introducing rising trends in industries on making informed decision-making by having highly structured data infrastructure and the method to process the data. It also includes results of maintenance benchmarking studies that identify the maintenance industry need for development of expert systems or decision support tools to support equipment maintenance. The second part is on academic treatment of maintenance terms, i.e. Failure and Downtime including definitions, equations, and related concepts. The third part is on the method to process the data, namely Artificial Intelligence Techniques with definitions, related concepts and review on available AI techniques with special focus on Bayesian Networks. The fourth part is on Theoretical Considerations for Bayesian Networks with definitions, concepts, and applications.

This research paper deals with the effort to develop methods and algorithms for resolving equipment failures using Bayesian Networks model and validate them using standard industry metrics. Emerging trends in computers and industry, case studies, and importance of structured data infrastructure for informed decision-making in businesses are discussed, leading to a review of two maintenance-benchmarking studies necessary to provide an understanding of current maintenance issues in industry before proceeding to the next section.

Initially, company information has been accumulated on paper in various forms; i.e. paper ledgers, logbooks, balance sheets, manuals, directories, data sheets, records, and other analog forms. Goldsmith (2003) states one case study where advances in other American economic sectors in applying digital information and

communications technologies are not proliferated successfully to some industries. Decision-making in the new millennium remains glued to paper, the telephone, and practitioners' memories. This includes paper records, often-unreadable paper prescriptions, paper orders, paper lab reports, paper telephone message slips, fax paper verifications, and paper bills of questionable accuracy. Retrieving useful data from this compendium can be an intimidating effort. Quite often, due to the considerable amount of time spent searching for data, that data holds less value when it is found for use. As Foy (1996, p.24) summed it up, the current model for information acquisition, storage, and access in today's corporations is hopelessly out-of-date. This is the state of an early 1970s information environment.

With the advancement of computer technologies available at a much lower cost, the availability of inexpensive yet powerful computer hardware and software reduces the costs of setting up new types of businesses, for example e-business, and expands the possibilities for setting up electronic portals or Web sites to conduct business (Brynjolfsson and Kahin, 2000). Companies have started engaging on new business models on digital platform over traditional channels. Example companies include VeriSign and GTE CyberTrust that have recently emerged as Certification Authorities (CAs), or third party companies, to provide software authentication services. These companies issue digital certificates based on technological mechanisms such as the public key cryptography, equivalent to software key, to access secure Web sites and databases (Schneier 1994). Another company, Red Hat, functions as a digital intermediary company, or 'middle-man' company, which adds value by testing and assembling customized software components for consumers (Brynjolfsson and Kahin, 2000). In short, improvement in computer technologies enables new businesses opportunities.

Slywotzky (1999, p.94) commented that some companies suffered through hard ways, i.e. missing profitability projections, or losing out market share to competition, etc before taking the digital transition. Intel, for example, decided to invest \$300 million to digitize its product-development process in computer-aided design and computer-aided manufacturing (CAD/CAM) following a \$203 million loss in 1986. That is a key investment as becoming digital in the design and production of chips improves competitive performance. Slywotzky also cited Wal-Mart making similar investments digitizing its logistics system at about the same time. By installing sophisticated communications and technology systems to provide real-time sales-and-ordering information, the company moved from atoms to bits. As a result, Wal-Mart outperformed its competitors by offering the right products at the right stores, by cutting costs, by integrating its operations with its suppliers, and by capturing valuable information about its customers. Twenty years ago, Wal-Mart and Intel were already digitizing their way of doing business.

This is demonstrated by the fact that a typical Fortune 500 company keeps an average of 8 Gigabytes of digital information in 1970, and steadily increases to 27,000 Gigabytes in 1990 and expected to reach 400,000 Gigabytes in 2000 (Foy 1996, p.23). However, a lot of this information resides on standalone computers or workstations, stored in a certain format understood by and benefit only a small number of staff. This will make other people difficult to access and analyze the data. At that time, networked computers were too costly and complex and are only in the domain of very large companies, which were financially able and willing to pay for its high cost. This is the first wave of digital evolution where there is a surge in industry transition to digital environment.

As time changed, and computer networks become more prevalent, these standalone computers or workstations started to connect in form of networks and the

digital data are shared among employees around the office. Pesko (1999, p.11) noted that by end of the millennium there are 50 million computers in the U.S. workplace and 12.3 million networks in operation, providing broader access to information. Within the companies, the digital data traffic started to increase, but still confines to the small periphery of the network topology it was being designed. In reality, what this means is that for example, office computer networks cannot speak the language used by the computer networks in the manufacturing floor, and vice versa. Hence, these networks are like 'islands of automation', with little or no connection between processes (Bentlage et al. 2001, p.215). This is the second wave of digital evolution where those seemingly separate digital environments start communicating to each other.

With the explosion of Internet in 1994, many companies, organization, governments, as well as individuals have started to tap into the power of Internet with the idea that this novel communication medium has the potential to spread the information quickly and in a standardized fashion. The open structure of the Internet now allows small firms to conduct businesses previously available only to a select few who had access to EDI (Electronic Data Interchange) which is the exchange of electronic business documents between two or more businesses.

With the advent of Internet or World Wide Web, publishing information becomes easier with increasing use of HTTP or Hyper Text Transfer Protocol, one of the most popular protocols. More than 100 million Internet users surf the Web around the world with 27.5 million people exchanging e-mail messages every day (Pesko 1999, p.11). The efficient distribution of content - any information that enabled to be communicated electronically, made possible by virtually removing all physical barriers of content distribution, promising bright future on the Web (Mccandless 1996, p.8). Most network configurations already employed in companies in either LAN (Local Area Network) as shown in Figure 2.1, or dial-up computer connected to the Internet have started using

World Wide Web as a method to disseminate information. This typifies the third wave of digital evolution where digital communication is used for serious business purposes.

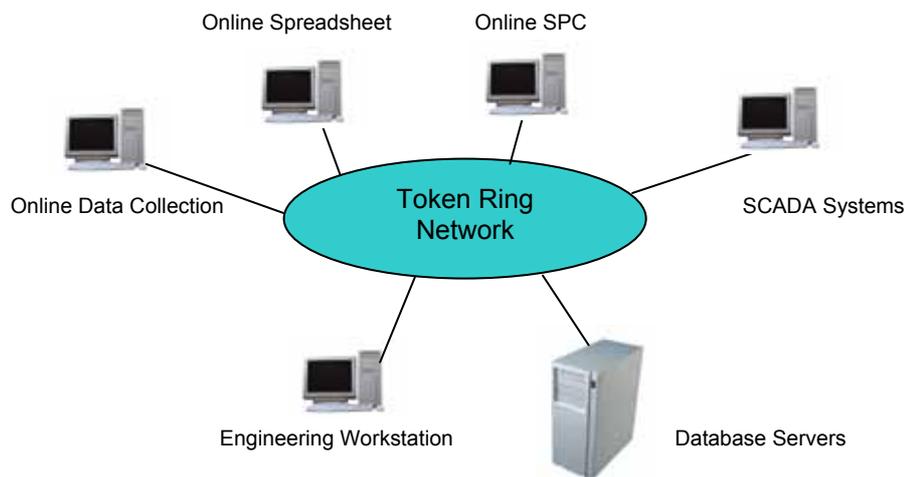


Figure 2.1: A Sample Token Ring Local Area Network (LAN) (Bentlage et al. 2001, p.216)

For the next phase, the fourth wave of digital evolution envisions all types of electrical or electronic appliances and equipments available in offices or homes to be connected to the network using the familiar HTTP Protocol. Information can be shared easily, quickly and understood by common people to make decisions. The intimate connection of a computer with the information infrastructure creates the demand for an appliance that can provide specialized access to information—hence the name “information appliance.” The real value of an information appliance is the ability to connect to the global repository of information, the Internet and the World Wide Web (Want, and Borriello, 2000, p.24). The Internet-enabled manufacturing is one breakthrough concept in an attempt to get the manufacturing equipment and machines in the factories connected to a company-wide network and have the key production parameters and controls hardwired to the central infrastructure. The fourth wave indicates ubiquitous computing or any time and any-place computing.

Today, a single multinational company can have a multitude of factories and offices spreading across continents. These factories usually deploy numerous manufacturing equipments producing a variety of products in various stages of production. With the increase in the number of factories spread in various geographical regions, the numbers of machines used for production without doubt will increase. This phenomenon poses new challenges to businesses.

As Michael and Thomas (2006) noted, the challenges faced are "... capacity allocation strongly influences supply chain performance and profitability. As with so many other supply chain considerations, it is a balancing act for manufacturers with multiple locations. Allocating too little capacity to a facility creates inability to meet demand and loss of sales. Saddling a facility with having to carry too much capacity results in low utilization rates and higher supply chain costs."

Other challenges include low cost and technology advantage, which are important criteria that enable successful companies to compete with and force other companies out of the market (Masatsugu 2000, p.178). Slywotzky (1999, p.94) reiterated that new challengers using new business models have risen to take on almost every leading company in almost every industry with the new models producing cost advantages of 10% to 20% for the innovators. For the computer industry, Compaq reinvented the business model to the dismay of IBM. Then Dell reinvented the model again - to the dismay of Compaq. For the air carriers business, Southwest Airlines reinvented the business model to the dismay of American Airlines. For steelmakers, Nucor reinvented the business model to the dismay of U.S. Steel. To maintain competitiveness, technology and cost are key factors to today's companies.

This is further demonstrated by a case study by Miltenburg (2005). Rolls-Royce used to be the most famous engineering company in the world but eventually lost due to cost and technology. The company is known to the public as a producer of luxury automobiles. It promoted its aircraft engines products against established companies such as Pratt and Whitney and General Electric to win a major order from Lockheed in 1968.

After some time, Rolls-Royce began to realize that it lacked the capabilities required for the project. The new engine incorporated new, unproven technologies, which was difficult for Rolls-Royce because the company was a technology follower, not a leader. Unanticipated problems and delays eventually caused development costs, originally estimated at £65 million, doubled to £135 million by early 1970. It nearly doubled again to £220 million in 1971. In 1971, Rolls-Royce fell into bankruptcy, showing technology and cost is important for maintaining companies' viability.

How these companies can overcome these challenges? Foy (1996, p.23) stated, "The value of a corporation becomes its ability to generate and to effectively communicate needed knowledge throughout the system of suppliers, customers, employees and communities to which it operates." Key learning for companies to take into consideration is facilitating data exchange for efficient decision-making to drive cost and technology. Corporations should practice effective data flow across their supply chains to ensure that business information is communicated effectively and timely to the correct audience.

Changes are happening in the industry at a very rapid pace, especially for technology companies. Without strong data infrastructure, organizations are unable to effectively and timely maintain control their own processes, resources and inventory. In time, it renders those organizations to become less and less competitive, lose market

share and eventually driven out of the marketplace. As remarked by Werner and Hermansson (2002, p.130), companies have to utilize all their resources, including information and technology, and refine and combine them to show patterns and support conclusions that could be used to provide better service to the customers, gain market share and increase profit by reducing cost.

How a well-defined data infrastructure can improve cost? Bentlage et al. (2001, p.215) observed that many different processes in companies were 'islands of automation', and being controlled or monitored using unrelated and incompatible tools. It was not possible to share data between processes, merge data for statistical analyses, and gain access to the data from an individual process. Many processes have no controls in place, and those using paper-based SPC or Statistical Process Control charts, for example, makes acquiring data for further analysis difficult. These were regarded as limitations to an effective defect and yield detractors analysis and development of solutions. Without the right data collection methodology or accessibility to the required data, engineers were constrained in their efforts to improve their processes (Bentlage et al. 2001, p.215). With this in perspective, without well-thought data mechanisms in place, any changes desired to the manufacturing process either to reduce cost or to simplify the manufacturing process or to shorten the throughput time are too tedious and time-consuming to make them happen.

How to optimize the use of data infrastructure? According to Mena et al. (2002, p.225), in order for an organization to be competitive, it has to deliver value to customers by offering better products and services at reduced costs to be profitable and gain market share. To this end, the organization needs to organize its data in such a way so that it is easily accessible in a timely fashion to decision-making people in the company, so that the most optimized decisions can be made in the best interest of the

company to produce better products and services. In essence, effective data flow is essential for companies to be successful.

Moreover, Werner and Hermansson (2002, p.131) pointed out, the necessary information exists within the company, but is too fragmented and complex for a human mind to make efficient conclusions upon. Getting the raw data is one thing, applying them appropriately to get the job done is another, and these issues pose challenges in formulating intelligent business decision-making.

The diagram shown in Figure 2.2 can represent this decision flow. The flow starts with Data Collection Process, where data coming from a specific monitored process is being collected. The flow continues with Data Analysis Process where the collected data is processed and analyzed for patterns or trends. The analyzed data is then used as a base for making decisions. Foy (1996, p.25) noted an example, for an accountant to accomplish a particular audit recommendation, a person may need to extract company financial data, from which he may construct statistical analysis, followed by an executive summary for him or others to make appropriate decisions. In this example, extracting company financial data constitutes the Data Collection Process; constructing statistical analysis comprises the Data Analysis Process, whereas summarizing the findings to make appropriate decisions falls under Decision-making Process.

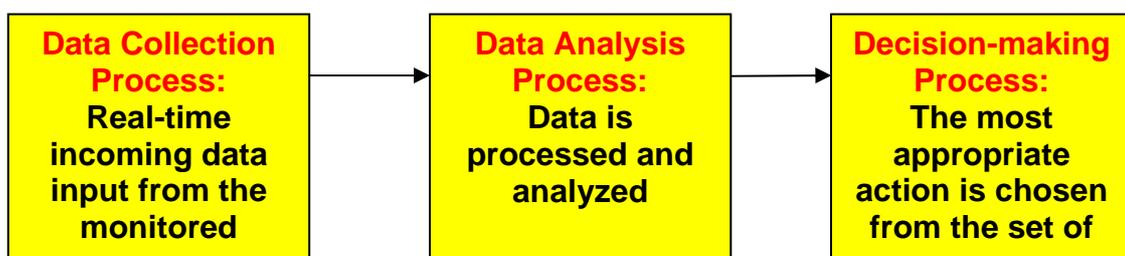


Figure 2.2 Block Diagram of Decision Flow

Going down to second level detail, one important piece of data to a company is in-depth equipment troubleshooting information, which can be in the form of paper-based file ledgers or records or an equivalent electronic database, used by technicians (in Intel Corporation for this research) to fix equipment problems. This repository or database of equipment troubleshooting information taps on the expertise of experienced people, acquired over the years on top of generic equipment training. This type of information enables experienced people to make good troubleshooting decisions. As for inexperienced people, for unfamiliar tasks, they will tend to perform trial and error judgment before arriving at the correct decisions, unless working under supervision of experienced people. When experienced people move out or transition to different positions or in the event that their skills decay after periods of skill disuse (Hall et al 1998, p.184), this valuable information goes with them. This is considered a loss to the company.

Luxhoj et al. (1997, p.437) stated that the changing needs of modern manufacturing necessitate a reexamination of the role that improved maintenance management plays in achieving key cost and service advantages and maintenance improvements to the organization.

From this point, to understand current issues in maintenance industry, Luxhoj et al. (1997, p.437) and Tu and Yeung (1997, p.453) reviewed two major maintenance benchmarking studies from Scandinavia and the United States. In February 1992, a EUREKA (European Benchmark Study on Maintenance, 1993) project was initiated that attempted to benchmark maintenance in Scandinavian countries, i.e. Denmark, Norway, Sweden, and Finland. The purpose of the benchmarking study was to establish a trade-by-trade overview of maintenance methods to assist companies in identifying current issues and general maintenance trends for improving maintenance.

Actual interpretation of the study's results among countries will largely depend on dynamic factors such as varying age and quality of machinery and buildings, interpretation or use of maintenance concepts, varying environmental conditions, differing forms of production operations (due to number of shifts and production technology).

Luxhoj et al. cited that the benchmarking study, beginning with Denmark, was based on an analysis of questionnaire responses from 43 industrial companies. The companies accounted for approximately 12% of the total revenues in Danish industry and approximately 8% of industrial employment. The industrial sectors of chemical and petroleum, nonmetallic mineral products, and manufacturers of food, beverages, and tobacco accounted for approximately 64% of the industry sector turnover in the sample.

On average, approximately 4.9% of the companies' turnover in 1991 was spent on maintenance, which was similar in percentage as 10 years earlier. It is interesting to note that from 1981 to 1991, there were increases in maintenance costs (expressed as a percentage of capital value) for the overall survey average (0.6%), for production, transport, and storage equipment (0.9%), and for spare parts (0.4%).

The "average" Danish company represented in the survey spent 32% of its maintenance budget on spare parts, 32% on salary and wages, and 31% on external services. In the average company, 23.8% of the maintenance costs were attributed to unforeseen repairs, 28.7% to preventive maintenance, and 45.5% to planned repairs.

Approximately 39% of the time spent on maintenance is used for unforeseen repairs, 20% for preventive maintenance, and 37% for planned repairs. Planning and control of preventive maintenance is performed in 45% of the companies. Use of the computer to control spare parts increased from 10% to 50% from 1981 to 1991, and

computer usage to control preventive maintenance increased from 9% to 60% in the corresponding period. However, 25% of the companies do not have any inventory control procedures in place for spare parts.

In Finland, the benchmarking survey was based on responses from 80 companies, which accounted for approximately 12% of the total Finnish revenues and approximately 14% of industrial employment. On average, approximately 4.8% of the companies' turnover in 1991 was spent on maintenance.

The Swedish maintenance survey was based on responses from 71 of 200 large and medium-sized companies from varied industries, such as chemical, paper, and pulp, steel and metal works, machine and transport equipment, electromechanical, and food. The Swedish survey illustrates that despite discussions of decentralization of maintenance resources, in the participating organizations, the majority of maintenance resources used (approximately 70%) are centrally organized.

The companies in the survey identified the highest priorities for improvement as the maintenance skills of the production staff, involvement of the production staff in maintenance work, continuous use of key figures, knowledge of maintenance throughout the organization, and control of the effects of maintenance on production volume. In addition, the survey indicated that the companies with the fewest number of shifts, or the shortest production time, reported a greater need for improvement.

Norway received 194 responses to its maintenance benchmarking study; approximately 60% of the respondents were from the food, engineering, and chemical industries. Seventy percent (70%) of the companies were small and medium-sized enterprises (SMEs). About 56% of the companies had no clear maintenance and availability objectives. Most of the companies had a centralized maintenance function.