UNIVERSITETET I AGDER

Visualization and diagnostic data for the Condition-Based Maintenance system

By
Yao Wenjun

Thesis submitted in Partial Fulfillment of the Requirements for the Degree Master of Technology in Information and Communication Technology

Faculty of Engineering and Science
University of Agder

Grimstad
May 2009

Abstract

Condition-Based Maintenance (CBM) system is getting quite popular in the industry and defense fields these days. A CBM system can efficiently predict the potential failures of the system and send the alarms. As a result, the lifetime of the system can be prolonged and the breakdown time can be decreased dramatically. The intelligent CBM system even can make the decision without any human intervention in the future.

This master thesis concentrates on the diagnosis part and the visualization part of the CBM system for a fan. The diagnosis is the core part of a CBM, which requires an efficient and optimum method for the different situations in reality. This method directly impacts the performance of the CBM. The visualization is an interface by which the personnel and the software system exchange information. The quality of this interface also can impact the feeling of customers and its availability.

In this master thesis, I developed a intelligent diagnosis solution for the CBM system for the fan. The system has the ability to learn from the experience and the known data, rather than just judges the values with the certain thresholds which is widely used as a tradition system state monitor method. I also implement the prototype.

This master thesis deals with a partial solution of an efficient CBM system for the fan in reality. The diagnosis solution can predict the potential failure in advance. The graphic user interface is designed to be easily understandable and neat. The personnel can read all related information from the interface at the first place. However, there is still much work to do. The method in this thesis and its prototype implementation can give the readers an overview and knowledge of a CBM system.

Preface

This thesis is the last project of my two-year's master program in Information and Communication Technology at the University of Agder, Faculty of Engineering and Science in Grimstad, Norway. This project is supported by Origo Engineering AS, which is a successful company in the system maintenance and support field, in Norway.

I wish to express my gratitude to my supervisor Trond Friisø (Origo Engineering AS) and professor Andreas Prinz (University of Agder) for their assistance throughout the whole thesis process. I am appreciated with the help from the other employees from Origo Engineering AS. They gave me many ideas about the CBM system and other technical supports.

Finally, I want to say, this master thesis not only gave me a chance to feel a practical project for a real problem, it also developed my project management and research abilities. I am really appreciated with the examiner to spend time and effect to evaluate my report.

Grimstad, 15 May, 2009

Yao Wenjun

Table of content

List of figures

1.Introduction

In this chapter, I introduce some basic information about this master thesis. This master thesis concentrates on a Condition-Based Maintenance (CBM) solution for a fan. The motivation is mentioned in 1.2. In section1.3, I present the problem description. At last, in 1.4, I briefly introduce the structure of this thesis report.

1.1. Introduction

CBM has become a famous topic in today's industry and defense fields. The reason is that with the development of the mechanical system, the machines are more and more complex. This new situation requires a monitoring system which can predict the specified the potential failures rather than after failures take place. An efficient CBM system can meet this new trend.

This master thesis concentrates on the design of the diagnosis part and the interface part of a CBM system for a fan. The diagnosis process is the core part in a CBM system. The analysis method's quality directly impacts the quality of the whole system. So how to find a stable and out-standing analysis method is the main job of the thesis. The visual interface is a tool via which the CBM system can communicate with the personnel. A good interface should be understandable and easy accessible.

The solution of the diagnosis and interface will be presented in detail in this report. After reading this report, you will have an overview of the CBM and how to build an intelligent diagnosis method.

1.2. Motivation

This project aims to solve a practical problem. With the development of the manufacturing industry system, it is impossible to let a person monitor everything and make decision when the failures happen. Just as everybody knows, a system breakdown causes huge lose of money or even lives. A good idea to handle this problem is to employ the CBM system to monitor the whole system and make the decision instead of people who may cause problems in some situations.

How to maintain a system's reliability and availability has been widely studied in the literatures because of their prevalence in industry systems. Maintaining a high or required level of reliability and availability is often an essential requisite.

A CBM system is a useful and also challengeable industrial IT solution. How to make such a system work well and how to optimize the system is the most important task of the project.

At last, this project is more practical, that means there is a chance the outcome of this project may benefit the reality. My interest also drives me to consider this thesis to be an attractive task to me. Besides that, this project applies me a view of the real industrial field which I may work in after I graduate. At last, but not least, in my point of view, the good attitude, interest and hard work can make common people to make out the uncommon results.

1.3. Problem description

This thesis is related to condition based maintenance. There are many philosophies for maintenance, but the trend is to go from following periodic preventive maintenance program (The histories of each machine type are analyzed and periodic overhauls are scheduled to occur before the statistically expected problems occur) to go towards doing maintenance when it is most optimal for the equipment and the process. To be able to do this it must be possible to determine the state of the equipment now and to predict how it will develop.

In a factory, or other industrial installations (e.g. onboard a ship), there are many systems collecting data. These normally collect data about a specific process or machine. Other systems are used for collecting other types of data like maintenance data and reports, results from laboratory analysis of samples etc. Normally these data is used more or less independently. For diagnosing with respect to maintenance optimization as much information as possible should be collected and analyzed.

The challenge in this thesis is to analyze and visualize these kinds of data. The research question is: How to visualize multi-dimensional time-series for use in diagnosis and prediction of the condition of process equipment and machinery?

The way the data are presented could be very context dependent. Visualizations depending on various situations and roles should be considered. It should also to get more detailed information if necessary.

After the discussion with the colleagues in Origo AS, the challenges of this thesis project are:

- Procedure of general CBM system development.
- Develop an optimum diagnosis method for a CBM system.
- Develop an optimum visualization strategy. Consider both generic and domain specific approaches.
- Implement prototypes/demos.

1.4. Scenario description

In the metallurgy industry, the manganese is the twelfth most abundant element in the world, and it causes the pollution both in air and soil. Manganese dioxide (MnO_2) is the normal state of the manganese in the nature. Manganese dioxide, shown in figure 1, is black or gray-black crystals or amorphous powder in the room temperature and does not dissolve in water.

Figure 1 Manganese dioxide in the room temperature [1]

In the process of the ferromanganese production, the coke is another important ingredient. Coke is a solid, dry and porous fuel which consists of 90-95% carbon (C), while 5-10% is non-combustible minerals (ash) [2].

In the process of the ferromanganese production, a lot of dusted air and water is produced. The polluted air is guided from the furnace in a smoke channel to a gas treatment department in where it is cleaned and cooled. This gas treatment guarantees the released gas is harmless to the natural environment. The gas is burned or sold to the market after all the strictly procedure.

As the gas treatment procedure, the polluted water, which is produced in the procedure of the water treatment, is also purified. Some purified water is reused in the gas treatment and the other is transferred into the final treatment plant (SRA). The reduction of the water is reinforced. The sludge is deposited.

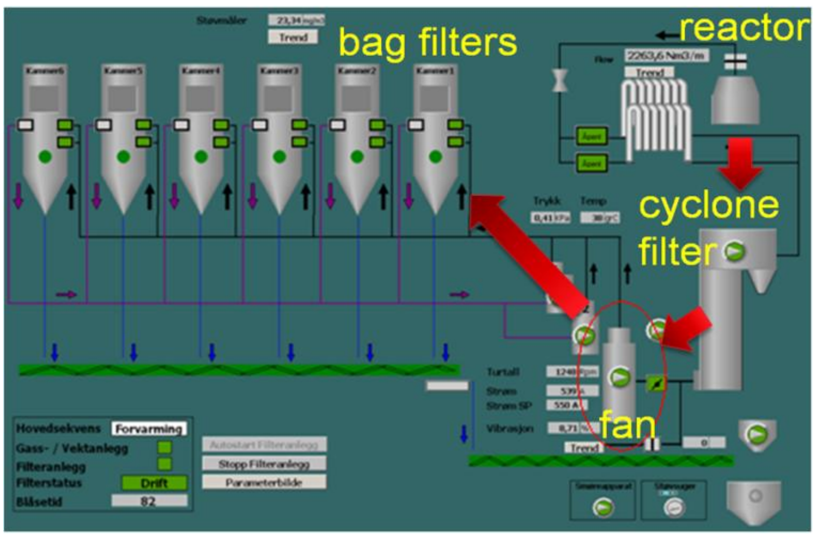


Figure 2 The off-gas purification process layout

In figure 2, we can see the layout of a MOR filter system. The MOR reactor, located in the right high corner of the figure, put out the high temperature extremely dusty off-gas. The off-gas is sent through the cyclone filter as the first step of the purification process. Then it is transported to bag filters for the further purification by the fan.

The fan which is focused on in this master thesis is the component in the red circle in the figure-2. The fan plays the role which transports the gas from the cyclone filter into the bag tanks.

1.5. Thesis structure

The thesis report contains seven chapters. The structure as follow:

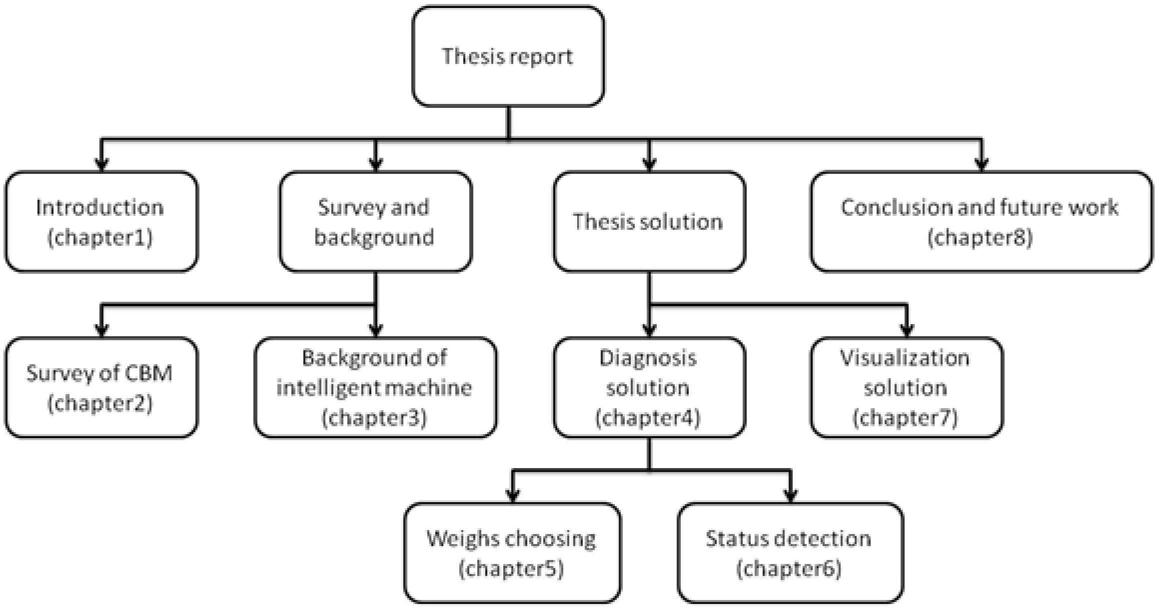


Figure 3 Outline of this thesis report

- Chapter 1 Introduction. This chapter gives the problem statement. It also applies the information about what is this thesis about.
- Chapter 2 Survey of CBM system. It releases the knowledge about the CBM system and its current state of art. This background gives an overview of the CBM system and its significance.
- Chapter 3 Background of intelligent machine. The intelligent machine constructs the diagnosis part of an intelligent CBM system. This background gives some basic background information about machine learning.
- Chapter 4 Diagnosis approach. This chapter is about the description of the main diagnosis solution. The diagnosis part is the key of an efficient CBM system.
- Chapter 5 Genetic algorithm for weights choosing. How to employ the genetic algorithm to solve the weights mapping problem is detailed depicted in this chapter.
- Chapter 6 Detection approach. This chapter describes the whole process of how the detection step works.
- Chapter 7 Graphic user interface design. It gives the knowledge of how to design to good GUI and the GUI solution for this thesis.
- Chapter 8 Conclusion and future work. This chapter discusses the conclusion or the thesis and the future work which is very important to make this system to be practical one in reality. The future work will make this fan CBM system to be much stronger.

2. Survey of Condition-Based Maintenance system

This master thesis focuses on how to build an optimum model of the Condition-Based Maintenance (CBM) system. So it is necessary to have a view of the CBM system in the industrial field. The content of this chapter will introduce you the common knowledge of CBM system.

2.1. History of CBM system

A system breakdown, some time just caused by a little component of a big system, in the commercial and defense markets can directly make significant negative impacts. During the last decades, engineers developed many approaches and implementation strategies tried to shorten the breakdown time as much as possible.

In the early days, the detections of the system were totally manual. The experienced engineers checked the systems or machines one by one and diagnosed the problems by experience. This kind of diagnosis approach was analog and comparably independent. When the system became too complex, such as a modern air plane, this approach was definitely no long suitable.

In the 1970s, the computers began to be used widely. This made the diagnosis of complex systems by huge quantity calculation possible. The outstanding digital calculation performance of computers could display different kinds of codes and identify each type of detected fault.

A CBM system is not a traditional fault detection system. It requires the system to be able to detect the faults in the early stage. In another word, the system must be capable to detect the potential failure some time before this causes the breakdown of the whole system.

Now CBM system is widely used in the industrial and national defense fields. For example, the U.S. navy uses the advanced CBM system widely on the planes and warships.

2.2. Significance of CBM system

The modern industry faces many challenges in the system maintenance field. With the introduction of the new concept of maintenance and technology development, the CBM has become the new leading approach.

The traditional maintenance approaches mainly contain: periodic time maintenance, corrective maintenance and urgent maintenance. The periodic time maintenance is the main approach which can prevent the potential faults in reality. Unfortunately, this approach has many disadvantages, such as, it wastes manpower, it may cause the unnecessary disassemblies which also decrease the health of the system, lack of pertinence, etc.

On the other hand, the CBM stinks to the state of the system, it can apply the personnel the real time or almost real time system status reports. The states of the systems are collected by the sensors and transferred via network.

An efficient CBM system has the following advantages:

- detect the potential fault in advance. The system is capable to diagnose the potential fault by the symptoms which are the related features extracted from the state data. The engineer can maintain the system base-on the report from the CBM, then fix the problems that cause the potential failures.
- Reduce the cost of maintenance, shorten the maintenance time or prevent the system from being over-maintained. The personnel only need to maintain the system when there are some potential faults can dramatically reduce the periodic maintenance cost. And also it keeps the system away from the too much unnecessary maintenance, over-maintenance also causes the high cost, decreases the system life and may increase the breakdown time.
- With strongly pertinence. Based on the states of the system, the potential faults can be specified. So the special engineer can be sent to the scene. This feature of CBM can accelerate the fast response speed and decreases the manpower cost.

In a word, the CBM is a revolution of the maintenance field. It applies the possibility of zero breakdown time system for the production industry which can make more benefit from it, and also for some system which cannot fail such as national defense system. Besides these current benefits, in the future the CBM will also contribute the out-space exploit, because CBM can work efficiently without any impact from human being.

2.3. System data and experience

An efficient CBM needs the support of knowledge from which the system can extract the features and do the diagnosis. The knowledge usually contains two main parts: data and experience. Figure 4 shows the general scheme for a CBM system.

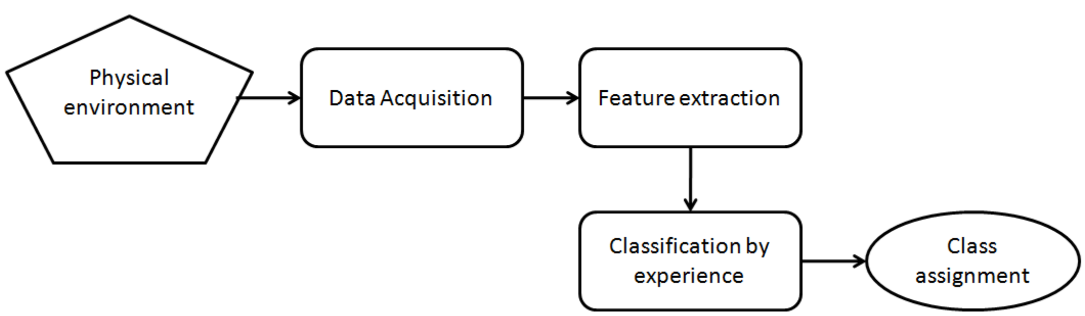


Figure 4 A general scheme for feature extraction and fault-mode classification [3]

2.3.1. Data acquisition

The collected data is the main source for the CBM diagnosis. This data mainly consist of sensor measurements and operator observations. The characteristics of data may change over time due to changes in the environment or changes in the system itself (due to wear and tear in the system, or component replacement). [4] The raw data from the sensor sometime does not stand for the parameter value directly. Usually some transformations or pre-calculations are done before the system receives it.

The data can also be analyzed then extracted some higher level information from it about the system. For example, the system needs the temperature of a component, and there are three temperature sensors with three different values. These values will be integrated into only one value which fits the format of the CBM's input set. These data represents a collective summary of the behavior of a system or a fleet of similar systems over a long period of time. [4]

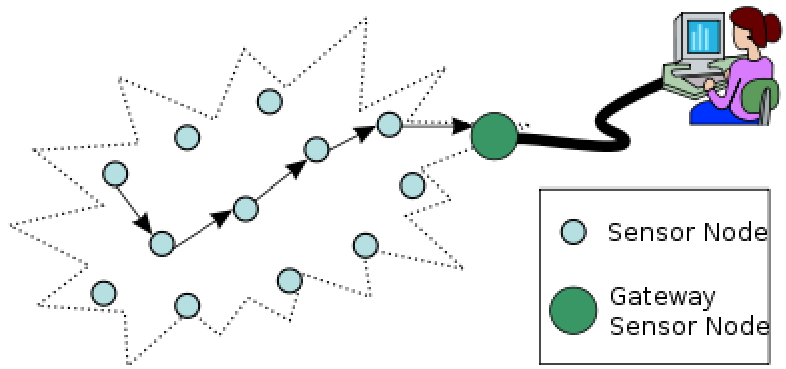


Figure 5 A typical Multi-hop Wireless Sensor Network Architecture [3]

2.3.2. Experience

Experience represents the knowledge deduced from data/information collected over a period of time from one system or a fleet of similar systems. [4] All the ideas of how to diagnosis the features of the system are based on the past experience of similar fault situations. The experience of the system either can be gathered by the expert designers or by the system learning process. For example, the system has to know what a failure is.

An important attribute of the experience accumulation is its temporal dependence. First, the data should be prepared and accumulated over a long period of time before a statistically meaningful set of experiences can be constructed. Next, various data mining and machine learning techniques can be used to consolidate the data and extract useful knowledge from it. Therefore, learning is another attribute that must be included in experience accumulation. [4]

2.4. P-F interval

There are many failure modes in the reality. Some are not age-related. And in most of these modes, the failures do not take place instantaneously, and can be detected by some evidences, which are also called the symptoms of the failures, in advance. When the failures are detected before they take place, it can be possible to take certain action to prevent the failure or avoid the consequences.

The curve in figure 5 is called the P-F curve. It illustrates a general process of a failure. In the figure we can see when the system's status deteriorates to some certain level, there are enough symptoms or evidences for a diagnosis system to detect the potential failure and predict is. The P point is the moment of prediction of the potential failure. If there is no any action to maintain the system, the functional failure will happen at some time called the F point (system failure). The deterioration process between P point and F point is faster than the normal process based on the observations.

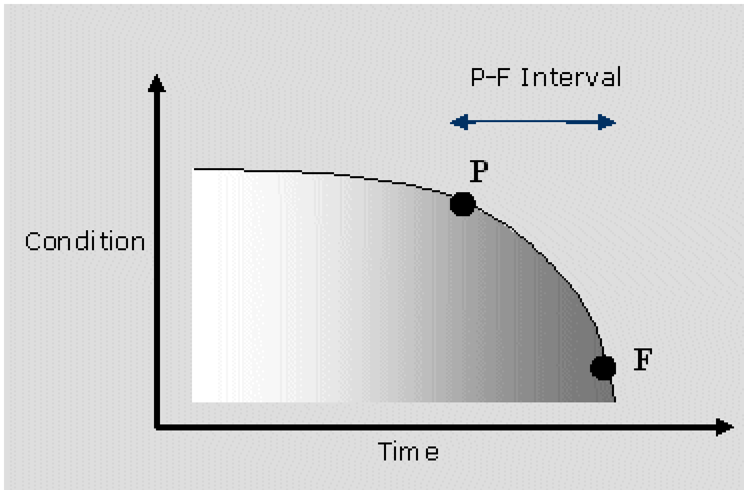


Figure 6 P-F curve [5]

The time between the P point and F point is the P-F interval. It is also the response time for the personnel to maintain the system. Technically, the P-F interval is the bigger the better, which means the system should detect the potential functional failure in the early stage. However, the designers must consider both the P-F interval and the negative alert possibility. When the system is too sensible, the possibility that the system send the alert to the personnel the fake alert is higher which will cause the high maintenance cost and personnel's psychological tiredness.

2.5. Intelligent CBM

Intelligent CBM means a CBM system is capable of understanding and making decisions based on the intelligent diagnosis method or even without human intervention. An intelligent CBM should include: the smart sensors which can be built in intelligence and be capable to collect rich and high grade information the system needs; and the intelligent diagnosis algorithm which can learn from the training and experience or recognize the trend from the history or other artificial intelligent approaches.

Usually, there are two typical ways to build the diagnosis method for a CBM system: model-based analysis and case-based analysis. The model-based analysis method is prone to build a mathematical model for the system. Then input the data and check out what will happen. The case-based analysis is a method concentrates on the certain types of failures (cases). And all the data and methods are about these certain types of failure.

2.5.1. Model-based Analysis

Model-based analysis is a very effective approach to detect the system potential functional faults in a complex system. The approach is to build a comparably complete model for the machine system. This model can simulate all the behaviors of the machine and then send the diagnosed result based-on the model. It can be used to delineate the true cause of an anomalous condition and its potential impact on the healthy system components, as well as the identity of the fault component.

The critical components that are prone to fail usually are constituent elements of a larger subsystem or system. [5] For example a helicopter, a car, etc. For these systems it is important to detect the potential failures of the detective components and specify the component which needs maintenance. The Model-base approach can apply the chains of the functions indicate functional flows and components link to the functions they support, whereas sensors link to the functions they monitor and conditions/constrains link to the function they control. [2]

In a word, the Model-based analysis approach is suitable to the complex system in which the components are strongly relative with each other, as the result, the system has to analyze the data together (as a model for every related components) rather than independently.

2.5.2. Case-based Analysis

Not as the model-based analysis based on the model for the machine, the case-based analysis approach concentrates on the specific cases. Compared with the model-based approach, the biggest advantage of the case-based approach is that, the system can learn the new case from the experience or knowledge field. In another word, the previous solution is likely to be in the proximity of the required solution for the current problem and hence and search space is considerably pruned

This approach offers a similar way in which people solve the problems usually. The following figure describes the steps of a typical case-based analysis approach process when a new case occurs:

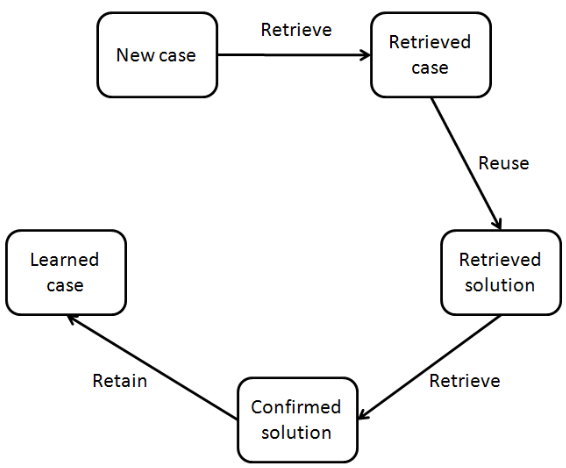


Figure 7 Case-based analysis circle

- Retrieve the most similar case(s). When the new kind of problem takes place, the system will find the most similar cases in the known domain.
- Reuse the case(s) to attempt to solve the problem. The old experience may be helpful to solve the new problem. The system will try to use the experienced knowledge into the new solution.
- Revise the proposed solution if necessary. In reality, there will be two exactly the same solutions for two different types of faults. The new solution should be adjusted by the artificial intelligent

techniques such as learning machine or data mining, etc. Currently, the system rarely does the adjustment without the human intervention.

- Retain the new solution as a part of a new case. The optimum solution for the new fault will be recorded in the knowledge base as its corresponding solution.

The case-based analysis approach is suitable to the comparable independent or isolated system or the component of the system. This approach does not need to build the complex model for every component, for instead, it just focuses on some specific faults.

3. Background of the intelligent machine

This chapter introduces the background of the intelligent machine. Generally speaking, the intelligent machines are the systems which with the learning ability and autonomy ability. The intelligent machine, especially the learning machine, provides the possibility to build a smart CBM system.

3.1. Genetic algorithm

Genetic algorithm provides a learning method which is illuminated by the biological evolution. It generate the potential solutions by repeatedly mutating and crossover the parts of the current solutions. After many repeating circles, the optimum solutions are hardly changed. When the optimum solution has been concluded, the algorithm can be stopped.

3.1.1 Brief introduction to Genetic Algorithm

In many science fields, people have to face to one problem: how to choose the optimum decision from the nearly infinite possible ones. There are many methods based on probability theory or other theories. These theories can lead scientists to assume a good solution according to the history statistic.

Unlike the other methods, the genetic algorithm doesn’t need refer to the experienced statistic. It implements the computer simulation for the possible solutions and their offspring, and then compares these solutions and picks out the best one. The process follows the principles of the survival of the fittest by Charles Darwin.

First pioneered by John Holland in the 60s, Genetic Algorithm (GA) has been widely studied, experimented and applied in many fields in engineering world. Not only does GA provide an alternative method to solving problem, it consistently outperforms other traditional methods in most of the problems link. Many of the real world problems involved finding optimal parameters, which might prove difficult for traditional methods but ideal for GA. [6]

3.1.2 Evolution in nature

Every day you can see a lot of people around: colleagues, friends or even yourself in the mirror. The configurations of human are quite familiar by everyone. But when the human being was in the beginning stage, everything looked strange and unfamiliar. The figure 8 illustrates a process which is quite familiar by everyone, the process of the human being’s evolution.

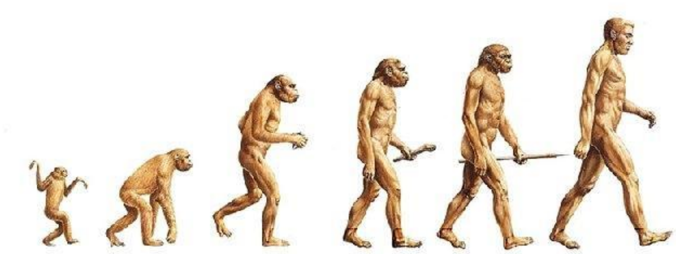


Figure 8 Evolution of human being

Millions years ago, the human were lower, hairy and just could use very simple tools. In the evolution process, the inherited traits of offspring of human had the change randomly, we call it gene drift. Genetic drift, an independent process, produces random changes in the frequency of traits in a population. It makes gene variants more or less common due to their causal effects on reproductive success. The changes due to genetic drift have no specific environmental cause, and may be beneficial, neutral, or detrimental to reproductive success. [6] These changes may subsequently change the gene and organisms. In this situation, the new species have a chance to dance on the stage of lives.

Most of these changes were the negative, and some were positive. The negative changes made the people not fit the nature, and they would die; on the other hand, the positive changes could make people to be better and they had more possibility to reproduce the offspring and continue these changes according to the natural selection. Natural selection, a process causing heritable traits which are helpful for survival and reproduction to become more common in a population, and harmful traits to become rarer. This occurs because individuals with advantageous traits are more likely to reproduce, so that more individuals in the next generation inherit these traits. [6]

The figure 9 shows the natural selection process. The darkness stands for the fitness of the nature. The circle stands for the one type of specie. The unfitness ones, the white ones and gray ones, disappear; the fitness ones are survival and their offspring keep the positive traits.

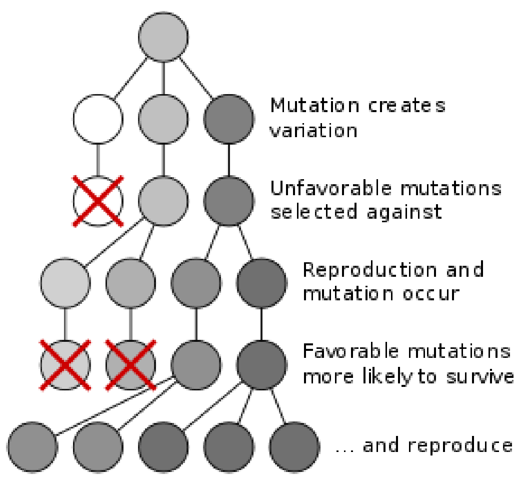


Figure 9 Natural selection of a population for dark coloration [7]

Another feature of the nature evolution is: as the other species, all the people have the same ancestor or the quite limited set of ancestor. Some theories consider the cradle of human, including all kinds of human who living in different continents, is the central part of Africa. And all the creatures on the earth are the offspring of the same ancestor. Some fossils of early stage of people, who lived about million years ago, are found in Kenya.

In the figure 10, there are four kinds of finches. They are different now, because the geographical isolation between them. This isolation can be considered as they have different fitness functions which will be discussed in the following content in this chapter. During the long time of evolution, each kind of finch finds the fittest way of the local environment. However, they share the same ancestor.

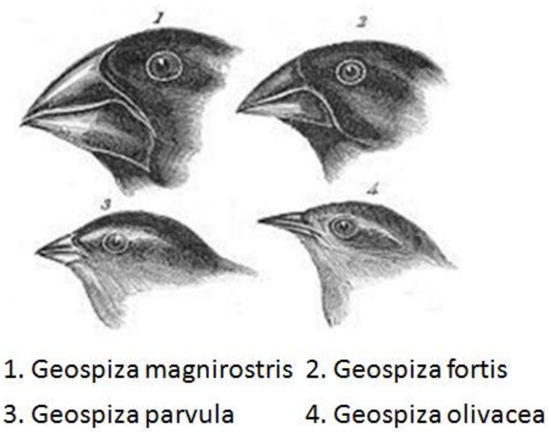


Figure 10 Geographical isolation of finches on the Galápagos Islands produced over a dozen new species [7]

3.1.3 Genetic algorithm

Genetic algorithm (GA) is a software problem solving technique which can simulate the biological evolution. This algorithm is widely used in the bioinformatics, computational science, engineering, economics, chemistry, manufacturing, mathematics, physics and other fields.

Some possible solutions are prepared as the initial set. The initial set should be encoded in a certain fashion, otherwise it is not easy to be evaluated. Usually, the initialization solutions should be set randomly as the ancestor set. This set we call it the solution domain. The solution domain is initialed by the designer on the first time of running. Then the selected offspring solutions will be set into the solution domain instead of the former ones for the next generation evaluation.

The evaluation part is dominated by the fitness function. This function contains the simulation of the solution domain. Generally speaking, how to simulation the known solutions is the most difficult part. Towards the different implementation field, the metric should be different.

In the figure 11, we can see how the GA works:

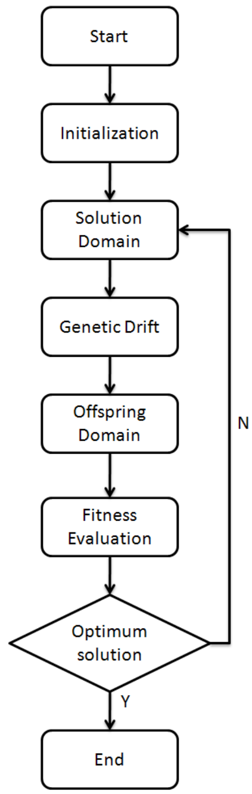


Figure 11 Genetic algorithm flow

- **Initialization.** The designers determine how many individual chromosomes. If the designers have an idea of the best solutions, the known chromosome can be set as those. However, most of the time the designers are probably better to initialize the chromosomes with random values.
- **Solution domain.** This is a collection of the known optimum chromosomes. The parents' chromosomes are stored in this domain. The size of this domain is decided by the designers. This size directly influences the quantity of the offspring's chromosome. Generally speaking, this size is larger the better.
- **Genetic drift. Usually,** there are two different ways of the genetic drift: mutation and crossover. Mutation is the parent chromosome or some part of the chromosome may mutate into a random value randomly. The mutation needn't any other chromosome involves. The following sample shows how the mutation happens.

  Parent Chromosome: 11-**63**-32-9-15-2-7-38-**59**-0 => 11-**47**-32-9-15-2-7-**38**-**22**-0

  The crossover is a process of two different parent's chromosomes exchanges their gene of the chromosome randomly. Every two parent-chromosomes can produce two offspring chromosomes.

  The figure below gives an example of the crossover process:

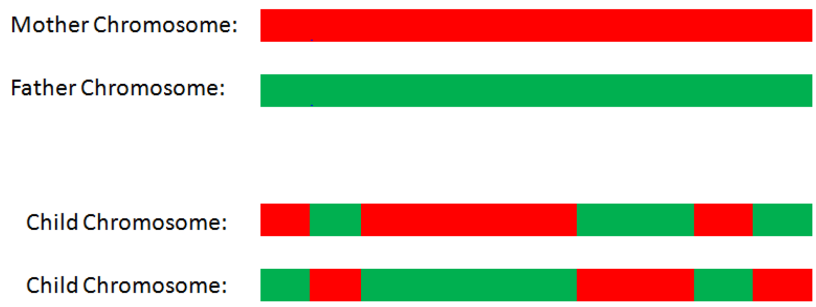


Figure 12 Crossover

- **Offspring domain.** All of the offspring, including the mutation offspring and crossover offspring, are stored in the offspring domain. The size of this domain is much larger than the solution domain. For example, if the solution domain contains 10 chromosomes, the solution domain may contain more than 3628800 chromosomes. Each chromosome of offspring domain will be evaluated and compared with each other and their parents. The best ones will be sent into the solution domain to reproduce the next generation and the others will be discarded.
- **Evaluation.** The evaluation step, in fact, is the simulation step. The designers use different metrics in different situation. But the purpose is always to find out the best ones. After certain generation cycle, the system will get the best one of the current solution domain as the final solution.

3.1.4 Advantages of Genetic Algorithm

There are many advantages of the GA. The most significant one is that the GA is intrinsically parallel. Most other algorithms are usually serial, that means they can only traverse the solution domain in one way, and if there is no optimum solution, the system have no choice except abandon all the previously work and start over. GA is quite different, it has multiple offspring. All the offspring are randomly mutate the traits, so the exploit direction of GA is multiple. When the choice in one direction turns out, it doesn't affect the final result, and the designers will be given an optimum solution.

The second notable advantage of GA is that the GA is very suitable to the solution seeking in the very large solution domain situation. In some implementation, the designers can pre-set all the potential solutions in the domain for the system's evaluation. Usually, in some cases, the system is not easy to accomplish the exhaustive evaluation. Evolutionary algorithms, on the other hand, have proven to be effective at escaping local optima and discovering the global optimum in even a very rugged and complex fitness landscape. (It should be noted that, in reality, there is usually no way to tell whether a given solution to a problem is the one global optimum or just a very high local optimum. However, even if a GA does not always deliver a provably perfect solution to a problem, it can almost always deliver at least a very good solution.) All four of a GA's major components - parallelism, selection, mutation, and crossover - work together to accomplish this [8].For example, the typical case for GA is that there is a salesman who wants to travel several cities and finally arrive the starting city without pass any city twice. The potential solutions' quantity is n!, which n stands for the quantity of the cities. If the n is a small number, it is comparably easy to design a system to calculate each potential solution. But what about n is 500? The 500! is a astronomical figure. By employing the GA, you can see the result as follows:

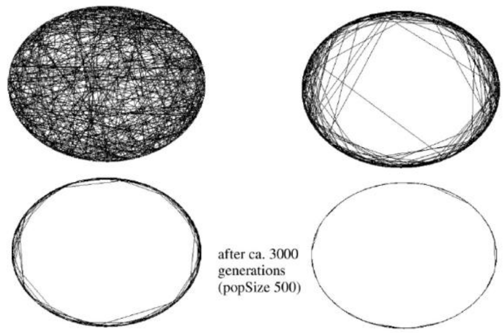


Figure 13 500 cities travel route after 3000 generations [9]

The last advantage of GA is that GA is competent to evaluate the unknown domain. Others algorithms need the designers must know all the details of the system, then the designers start to find the best solution. However, the GA doesn't need the designers to pre-know how to configure the values or parameters. And GA is capable to manipulate the values as the chromosomes automatically and simulate these chromosomes. This feature is suitable to the problems which aren't able to be stated clearly in advance or be specified the details. The figure 14 shows the designs of the satellite trusses for NASA. The first design is made by the experienced designer. And the second design is from the GA. You can see there is no one can design shape of the second truss. As the result, the GA's design is 20,000% better than the "normal one". The unmoral shape antenna in figure 15 is also a piece of work by GA.

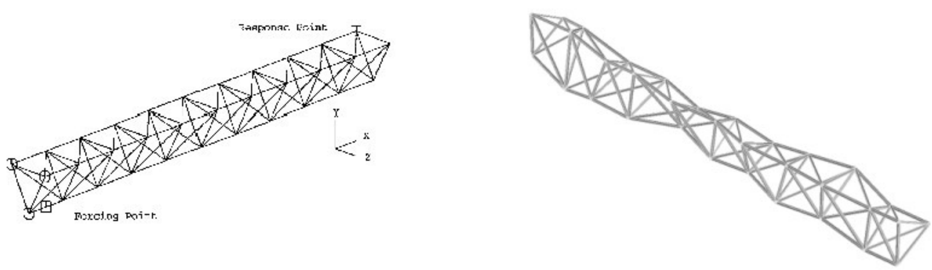


Figure 14 Trusses from designer and GA [9]

Figure 15 Satellite antenna designed by GA [9]

The virtue of GA technique is that it allows genetic algorithms to start out with an open mind, so to speak. Since its decisions are based on randomness, all possible search pathways are theoretically open to a GA; by contrast, any problem-solving strategy that relies on prior knowledge must inevitably begin by ruling out many pathways a priori, therefore missing any novel solutions that may exist there. [10]

3.2. Machine learning

The word learning, as the word intelligence, covers a very broad range of process and it is difficult to be defined precisely. All the human being and animals have this ability from the birth. Unfortunately, this ability is still not totally described and understood by the researchers in the computer science and mathematics field.

3.2.1 Brief introduction of machine learning

The machine learning usually refers to the chances in some systems which operate the assignments as artificial intelligence. These systems usually are related to recognition, diagnosis, planning, robot control, prediction, etc. In these system, the unexpected situation is unavoidable, or the completely status space is impossible to be totally completed. This is also the reason why the learning machine must be involved in some fields.

Work in machine learning is now converging from several sources. These different traditions each bring different methods and different vocabulary which are now being assimilated into a more unified discipline. [11]

3.2.2 Neighbor learning

The neighbor learning is a very classical instance-based learning approach. The most basic instance-based method is the k-nearest neighbor algorithm. This approach defines a distance is an n-dimension space. In this master thesis, the space is three-dimension, so I am not going to introduce the cases in other dimensions.

There must be at least one known data in the space, so that when the algorithm can consider them as the neighbors. Usually, the known data is initialized as the pre-known data from the training. All the data in the training step are clear and should contain all the potential status of the space.

The k-nearest learning procedure is not complicated. When all the neighbors are found in a certain distance, the average value of the neighbors will be assigned as the value of the point which is waiting for its status or value.

One obvious refinement to the k-nearest neighbor algorithm is to weight the contribution of each of the k neighbors. In many cases, the importance of the different types of data is different. For example, there are five points in a certain distance. One point may much nearer than others points. In this case, this very point should be considered to be more important than others. In another word, the weight of this point should be higher than others. How to choose the weight for different points should be closely related to the different cases.

4. Diagnosis Solution development

In this chapter, the solution for the fan is presented. First the research methodology is introduced which can make sure the solution design process can be repeat on other projects. Then the solution for this case is presented in detail.

4.1. Research methodology

The purpose of the methodology is let the readers know how the author collects or generates the data and how this data is analyzed. In another word, the methodology shows the others how the results are obtained. The different methods may produce the different results. An improper research method hardly causes the expected academic outcomes. And usually, there are more than one way of method can competent a project.

This master thesis is to build a software CBM system for a fan which has the features of the practical industry. So the method must be suitable to the reality. The best method to be adopted in this project is the engineering design. The following figure shows the steps of this master thesis.

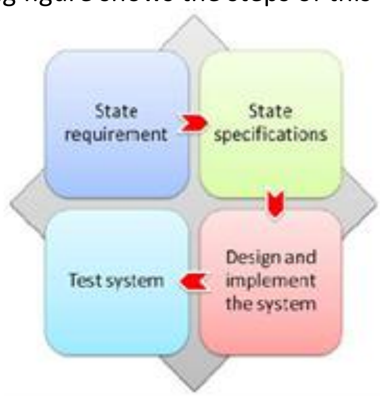


Figure 16 Development phrases of this project

- **State requirement:** The requirements of this project are that design a software solution of the CBM system which can monitor the fan in the real-time and announce the maintenances needs in advance. The result should be easy understandable to the personnel who are monitoring the system or going to fix the fan's problems. And all the features of the system must be shown. By using the system, the break down time of the fan system can be decreased dramatically.

- **State specifications:** The different situation needs different CBM solutions. The biggest difference is the data analysis method should be unique to the special environment. According to the different features, all the methods should be considered and compared. The result must meet all the requirements.

- **Design and implement the system:** The analysis of the methods should be based on the mathematic algorithms. Also the features of the case must be considered in the first place. Implementation of this step is not only to use a software language to realize the function of a certain mathematic algorithm. In this step, the efficiency of the system must be considered. How to build a simple, smart and neat system is also a research problem.

- **Test the system:** At last but not least, the testing part is a step which makes sure the system may work well in any situation. Usually the system may work well the normal situation. But this is far away with the end line. The testing work must make sure the system is available and steady all the time. All the states of the system must be test, and then find out the weak point of the system then fix it. Usually the test part will cost as much resource as design work.

4.2. Report Structure for solutions

This report presents the solution as two parts: diagnosis and visualization. In this chapter (chapter 4), the outline of the diagnosis solution is introduced. The development method is fully expressed. The chapter 5 and chapter 6 discuss the two problems in the solution in detail. Chapter 5 presents how to choose the optimum weights for the data mapping by genetic algorithm. Chapter 6 will releases how to detect the machine's current status. The visualization of the CBM system is presented in chapter 7. In chapter7, the guideline of graphic user interface design is listed and my design for this thesis is discussed.

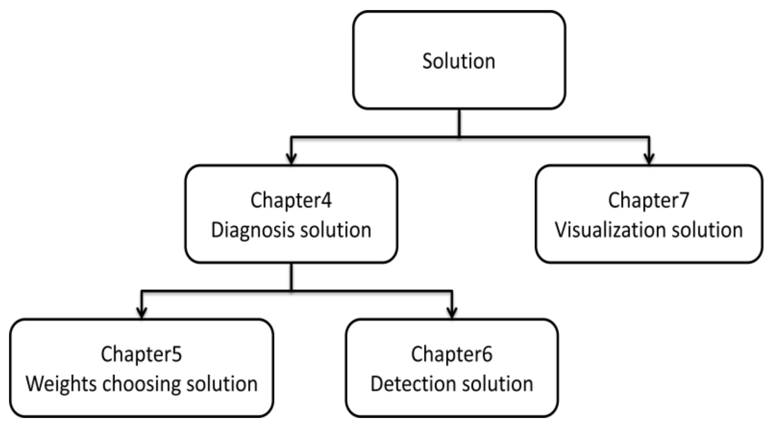


Figure 17 Report structure

4.3. General CBM system development method

The development of a CBM system is similar with other software systems. The first thing to do is to clear the requirement. Different CBM systems may face to the different situation. Some need intelligent diagnosis function; some may just need monitor the parameters and compare them with threshold.

When the requirements are clear, the specified data must be collect and integrated in certain format if it is necessary. Usually the data is real time.

The third step is to analyze the features hide behind the raw data. Usually the original data can not apply many symptoms of the system's current state. The feature choosing can directly impact the result of the system.

After the necessary data features are ready, the system diagnosis approach should be developed. This part is the core part of a CBM system. According to different requirements and situations, the optimum approach should be chosen. Whether the diagnosis approach can work as expectation is crucial.

The last step is to test whether the system can work well and need all requirements. If not, the solution must be improved until all needs are met. The testing is also very important. In reality, some time, the testing cost as much as the design.

The figure below shows a general CBM system development procedure:

Figure 18 General CBM system development

4.4. Solution for this thesis

The key part of this master thesis is how to design a mathematic solution to a fan system. This mathematic solution must focus on the symptoms of the data and detect the failure in future by diagnosis these related symptoms. This procedure is quite similar with the procedure a patient go to see a doctor. As what is discussed above, the solution should be specially designed to the case. The following content will state the details about this solution.

4.4.1. System requirements

The system requirement is the guideline of the software system design. It usually specify the input and the outcome of the system, in another word, it tells us what this system can do. Generally speaking, the requirements are assigned by the customers. The requirements contain two parts: minimum requirement and recommended requirements. The 'Minimum system requirements' must be satisfied for the software to be usable at all. And the recommended system requirements are often suggested by software vendors for optimal performance of software. [12]

After my pre-research the knowledge about CBM and discussed this thesis with the experts in Origo AS, where I do my master thesis in, I concluded the system requirement. In this master thesis, the minimum requirements are presented as follows:

Data:

- System receives real-time data.
- The data must be preprocessed.

Analysis functions:

- The analysis system should have an optimum method which can detect the equipment's failures in advance.
- The System should specify the different types of the failures.

- The model should be able to be evolved by the experience.

Interface features:

- The System should support several different multi-dimension visualization strategies.
- The interfaces should be easy to be understood.
- Both of the raw data and the analysis result should be shown.

4.4.2. Diagnosis solution structure

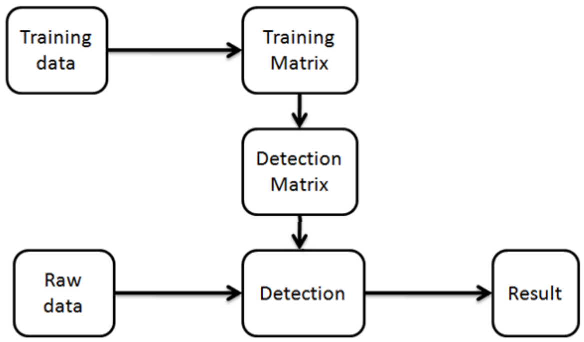


Figure 19 Diagnosis procedure

The figure above illustrates the solution for diagnosis. There are two main parts: training and detection. The training is a process that the system learns some known knowledge which the extended solution based on. The detection step is based on the learning machine. It can provide the diagnosis result even there is no corresponding system status in the detection matrix which contains all the system states and statuses.

4.4.3. Overview of raw data

The raw data is collected by the sensors around the fan, and the data is transferred from sensors to the computer by the network. The data is real-time data and preprocessed.

There are three types of the raw data: vibration, temperature and electronic current. They are the input of the system. The diagnosis is based on these three types' parameters.

The following figure is a sample of the electronic current data. The data in the first column is the time of the sampling. The period of sampling is ten minutes. The vibration and temperature data are also sampled in every ten minutes. In the second column, it is the data itself which will be analyzed and predict the failure. The last column contains the status of the system. In this master thesis, there are three statuses: normal, error1 and error2. In this master thesis, only the data in the second column is concerned.

11/7/2008 16:19	696.189819	262336
11/7/2008 16:30	698.551636	262336
11/7/2008 16:40	698.232483	262336
11/7/2008 16:49	700.750244	262336
11/7/2008 17:00	702.157898	262336
11/7/2008 17:10	699.677307	262336
11/7/2008 17:20	700.228394	262336
11/7/2008 17:29	705.537476	262336
11/7/2008 17:39	724.085327	262336
11/7/2008 17:50	738.322021	262336
11/7/2008 18:00	582.114868	262336

Figure 20 Sample of the current data

The figure 21 and figure 22 show the profiles of temperature and electronic current data on November 17th:

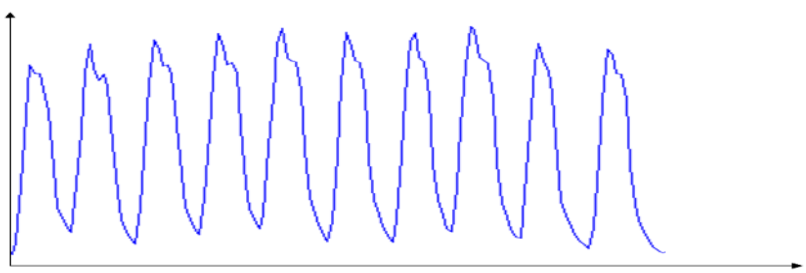


Figure 21 Temperature of a day

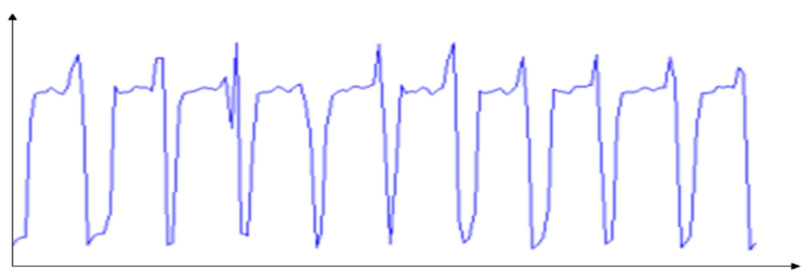


Figure 22 Current of a day

4.4.4. Diagnosis unit

In the problem scenario description, it is presented that there are ten times of the air cleaning processes in each day. So the data also profiles the strongly periodical character. In the figures above you can see the wave-shape profiles.

In order to diagnosis the data, the whole data sequence must be split into units by time. In each unit, the data must contain potential symptom. That means there must be no missed symptom after the data split into units. Based on the apparently periodical feature, split one day's data into ten pieces is wise. We call each one of this time piece a "Time Window". In each time window, there are about fourteen points. There points will be analyzed together as one unit. The figure 23 is an example of the time window on the temperature data. The width of the window is one of tenth of a day.

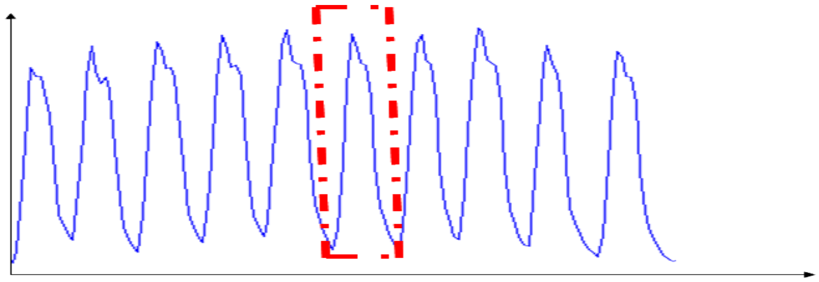


Figure 23 Temperature data with the time window

4.4.5. Features of data

After the observation, we conclude that, before the failure, the value of the wave is much higher than the normal area and the profile of the wave is more flat. With these two characters of the symptom, the mathematical expectation and variance are employed to help the system to extract the useful information before the failures.

In probability theory and statistics, mathematic expectation of a random variable is the integral of the random variable with respect to its probability measure. For discrete random variables this is equivalent to the probability-weighted sum of the possible values, and for continuous random variables with a density function it is the probability density -weighted integral of the possible values. [12]

In this master thesis, the mathematical expectation is to calculate the average value of the parameter vibration, current and temperature within a time window.

In probability theory and statistics, the variance of a random variable, probability distribution, or sample is a measure of statistical dispersion, averaging the squared distance of its possible values from the expected value (mean). Whereas the mean is a way to describe the location of a distribution, the variance is a way to capture its scale or degree of being spread out. [14] Variance here is used to analysis the profile of the wave. When the wave is as sharp as usual, the variance is high. And in the condition of the wave is flat, the variance is small number.

4.4.6. Data mapping

There are three different types of the monitored parameters: vibration, temperature and electronic current. The raw data is first filtered by the time window. Then the values in the each time window are calculated, and get their mathematic expectation and variance. After the calculation step, there are six parameters now: the MSE of the vibration, the variance of the vibration, the MSE of the temperature, the variance of the temperature, the MSE of the current and the variance of the current.

These six new parameters must be integrated by the way in which the symptoms can be extracted easily. The integration process in this master thesis is to map the data into a three-dimension space. The coordinate's axises of this space are: vibration, temperature and current.

The method of mapping is to use the weights for the different parameters, to measure the importance of the parameter. The following functions are the mapping method. The a,b,c,d,e,f are the weights which will be concluded by the genetic algorithm. The reason why I used the genetic algorithm and how to use it in this thesis will be released in the following chapter.

$$\begin{cases} X = a \times MSE_{vibration} + b \times Var_{vibration} \\ Y = c \times MSE_{temperature} + d \times Var_{temperature} \\ Z = e \times MSE_{current} + f \times Var_{current} \end{cases}$$

After the mapping, all the raw data are considered as a point in a space. In order to get a better view, we can introduce a matrix to present the different points. In the following figure, there are twenty-seven points in the space as the matrix view.

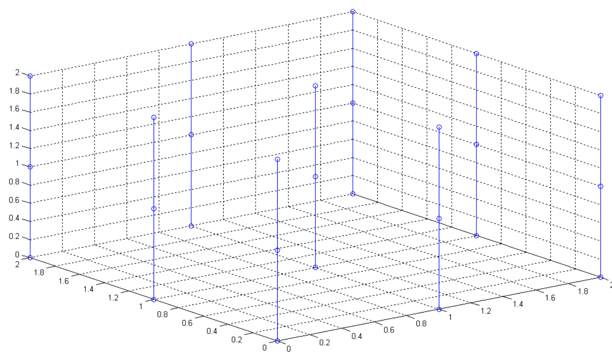


Figure 24 Twenty-seven points in the space as the matrix view

This view does not change the relative position between each two of data, this is just a better view for the designers to understand the data. In the figure, the points in the matrix stand for the positions of the data, not the values. The values do not change two. This approach of view is related to the self-organizing map approach which is discussed in the next section.

For the whole system diagnosis, the size of the system space is fixed. There are boundaries of the data space. Any data of out the boundaries must present some known error. In the master thesis, the space in the boundaries is divided into ten parts on each direction. So there are one thousand points or states in the system space. Figure 25 shows this space in the matrix view.

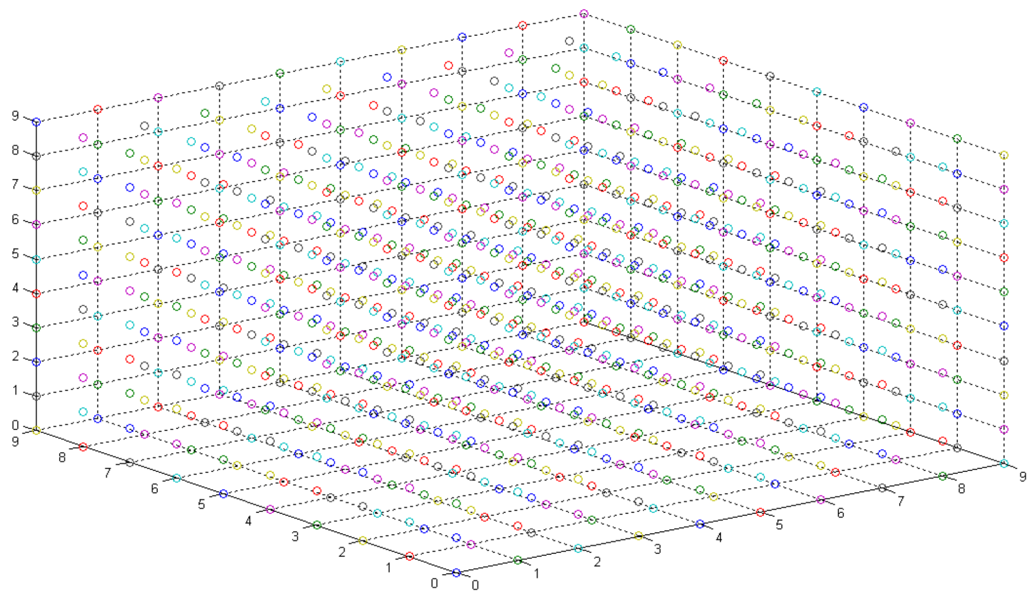


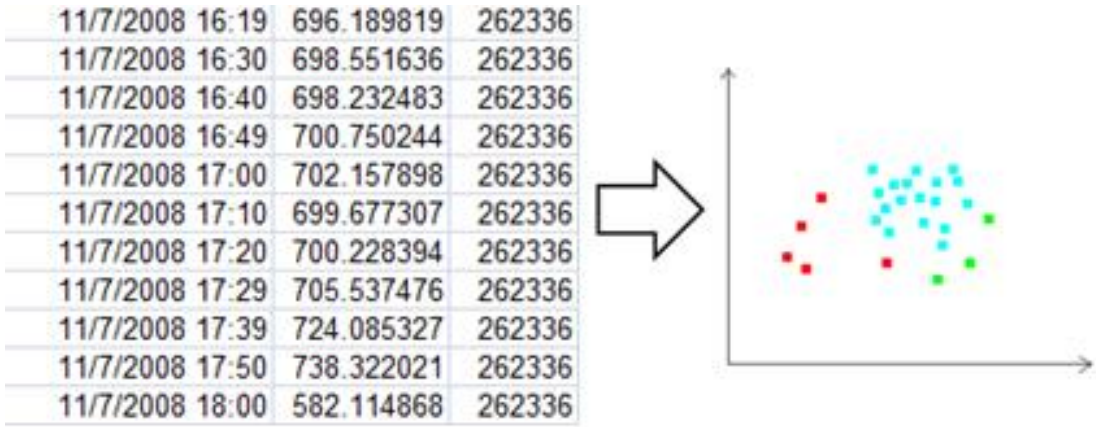
Figure 25 System space matrix

The significance of the introduction of the matrix is that, in the reality there are infinite points in the space. As a result, it will be very big workload for the computers to calculate and record each single data in the space. However, the space matrix only introduces some certain number of the points in the space instead. In the master thesis, we consider the whole space only has one thousand states. And by this way, we can much easier to detect which state the system is in right now.

4.4.7. Self-Organizing Map

The Self-Organizing Map (SOM) is an algorithm used to visualize and interpret large high-dimensional data sets by projecting them to a low dimensional space that has typically one or two dimensions. The SOM theory and algorithm is presented thoroughly by Kohonen. [12]

The mapping method in this master thesis is the SOM method. The figure 26 shows an example of the data maps into a two-dimension space. As the discussion above, the weights are introduced here to calculate the mapping function. The figure 27 shows how to change the normal view of the data space into a matrix view. In the figure, it is the training data which will be discussed later. The training data have twenty-seven points in the space. We use three colors: red, blue and green to present the different status of the points. Blue means normal, red means error1 and green means error2.



11/7/2008 16:19	696.189819	262336
11/7/2008 16:30	698.551636	262336
11/7/2008 16:40	698.232483	262336
11/7/2008 16:49	700.750244	262336
11/7/2008 17:00	702.157898	262336
11/7/2008 17:10	699.677307	262336
11/7/2008 17:20	700.228394	262336
11/7/2008 17:29	705.537476	262336
11/7/2008 17:39	724.085327	262336
11/7/2008 17:50	738.322021	262336
11/7/2008 18:00	582.114868	262336

Figure 26 Mapping from data into space

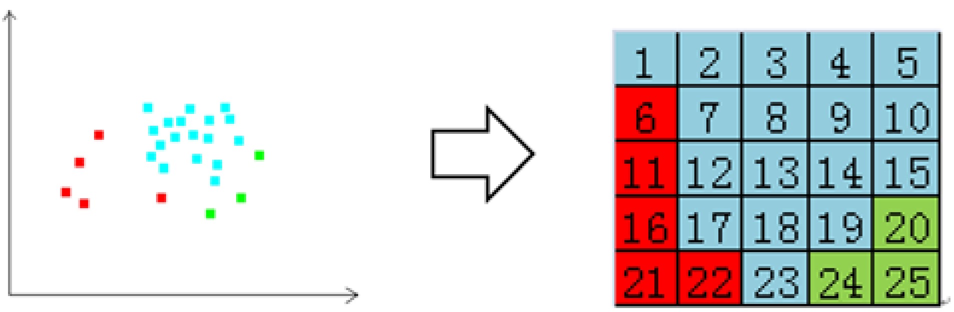


Figure 27 Matrix view of the space

According to the theory of SOM, the points do not locate in the space randomly. The same color points are close to each other. In the figure 27, the red ones are in the left corner and the green ones are in the right corner. This means the raw data is organized by its own features. The significance of SOM in this thesis is that, after using SOM the data is organized. And the organized data is much easier to do the distance calculation in genetic algorithm which is mentioned in the chapter 5 and to implement the machine learning approach which is mentioned in chapter 5.

By using the SOM method, the organized data is much easier for us to analyze. However, in this master thesis, this is not enough. The view, that the points in the space, is not good enough to be unified. In the figure 25, the view is changed into the matrix view. This change is just for the view. All the relations between he points in the space and relative positions are not changed at all. The benefit of the matrix view is to let the designers analyze the data in an easy understandable way, and let the implementation part by the computer language to be much easier.

4.4.8. Machine learning

Learning is the most important attribute of intelligence and an evolving knowledgebase. Machine learning is concerned with the design and development of algorithms that allow computers to improve their performance over time based on data. [13] The system should learn from experience about how the environment behaves and should create a model of this environment. [14]

As discussed above, in this master thesis the whole state space contains one thousand states. At each single time, the system's status responding into one and only one state which is also considered as the system's status.

The learning machine helps the system conclude the whole space states from limited known data which is name the training data. The first step of the learning is the supervised learning. Under supervised learning the agent is provided with a target or a purely instructive feedback, for example, the environment tells the learner about what exactly its output should be. The agent then compares its response with the target and adjusts its internal memory in such a way that it produces a more appropriate response the next time it receives the same input. [14]

The training data in this master thesis contains twenty six points in the system space. The training data is first used to calculate the weights which are used to map the raw data into the state space. In this procedure, the genetic algorithm is introduced. Genetic algorithm is quite suitable to find out the optimum result in a certain large area. The whole procedure will be presented in the chapter of the genetic algorithm.

Then the training data is mapped into the space and the system remembers these states and responding location as what is shown in the figure 24. And we call this matrix as the training matrix. In the following content, there will be a detection matrix which can present the whole state space. And the training matrix in face is a subset of the detection matrix.

4.4.9. Detection

After the training is completed, the detection step is going to start. The first thing of the detection step is to make the system map the trained matrix into the detection matrix, as introduced above, the training matrix is a subset of the whole space and the detection matrix presents the whole system space.

In this step, the system boundaries should be made. For example, the temperature of the fan system cannot be upper than 200℃ or less than 0℃. The boundaries can give us an impression of how big the detection matrix should be.

Then we remap the training data into the detection matrix by using the same weight we get from the genetic algorithm. Figure 28 shows an example in the two-dimension view for easy understanding.

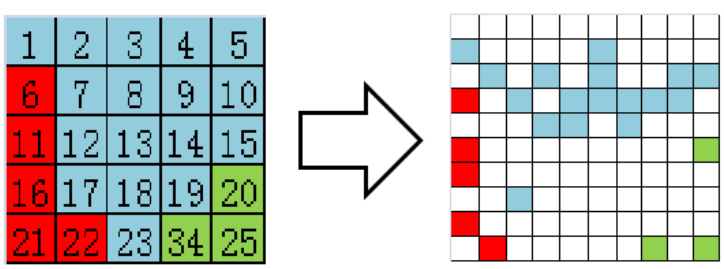


Figure 28 Mapping from the training matrix into the detection matrix

The last step is to map the current monitored state data into the detection map. If the corresponding lattice has a state, then the current state is this state; if not, the system will detect the neighbor lattices which have the same lattice distance, and find out the majority state as the fan system's current state.

The lattice distance in here means how many lattices between two lattices. For example, the lattice distance between the yellow lattice and purple lattices is 1; the lattice distance between the brown lattice and the grey lattices is 2.

When the current state is in an unknown state lattice, the system first search the neighbor lattices with lattice distance is 1. For example, in the figure 29, the yellow lattice will search the purple lattices, and find out the there is one blue lattice. So the blue is the majority, and the yellow lattice will be marked itself as blue. The system will do the same procedure for the brown lattice. The number of the distance 1 neighbors is 0.Then go to search distance 2 neighbors, find out the majority states is green. So the brown lattice will be marked as green. If there are two states have the same number of lattices with the same the distance, the state will choose a random state for the current state.

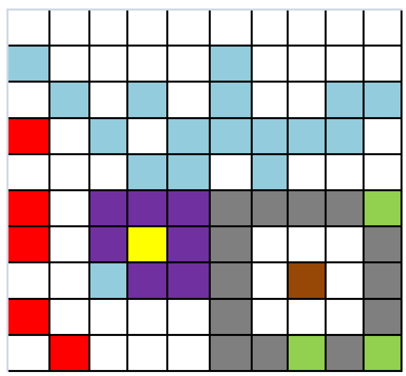


Figure 29 Detection

Finally, all the lattices will be marked as unknown state. And the system will just need to map the data into the detection map and find out the corresponding lattice's state. The finally detection matrix in two-dimension may like figure 30. And the whole procedure of the detection is illustrated by figure 31.

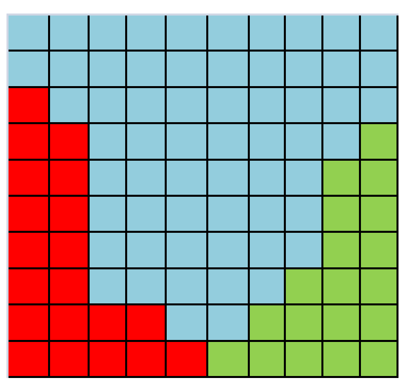


Figure 30 Example of the final detection matrix

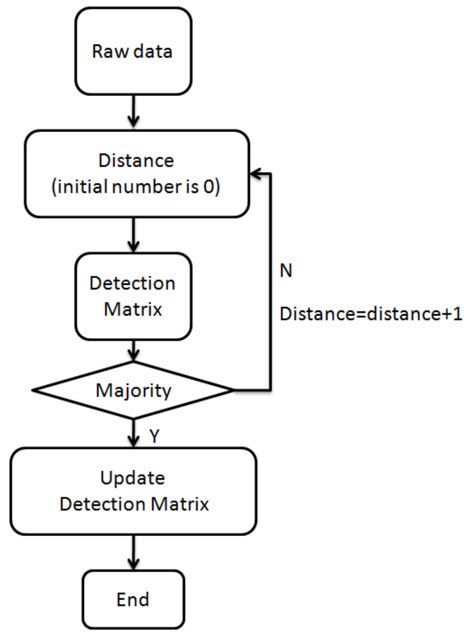


Figure 31 Detection procedure

4.5. Summary of diagnosis solution

The solution design steps of this thesis can be concluded as follows:

- Raw data (three dimension). The data stands for the states of the machine. There are three different types of parameters: vibration, temperature and electronic current.
- Extract features (six dimension). Based on the raw data, the features, from which the symptoms can be concluded, must be extracted. The features usually are more important than the data itself. In this thesis, there are two features in each type of parameter.
- Rebuild data (three dimension). According to different situations, the data should be reorganized. An organized data set can be analyzed much easier.
- Mapping. Mapping is an important step. The mapping approach is related with the diagnosis approach or other approaches, such as the fitness function of GA in this thesis.
- Training. The learning system must own some pre-known knowledge as the initial knowledge domain. After it, the system will learn more from its own experience.
- Detection. The target of a CBM system is how to detect the potential failures. So the detection step is crucial. The core part of an intelligent CBM system is how to let the detection part “smart” enough.

5.Genetic algorithm for weights choosing

This chapter introduces the background of the genetic algorithm (GA) in the computer science. And how to employ this method for choosing an optimum set of weights which are used to mapping the data into the certain format (detection or system space).

5.1. General genetic algorithm approach

The background of the genetic algorithm has been introduced in the chapter 3. The GA is a computer science method motivated by the biological evolution. The biological gene drifts take place to all direction and the nature works as the fitness function. As the time goes, only the optimum drifts are kept, others disappear.

For this thesis, GA is used as the approach of weights choosing because that, GA can search spaces of hypotheses containing complex interacting parts, where the impact of each part on overall hypothesis fitness may be difficult to model. [15] How to choose the optimum set of weights is such a problem. The problem is mentioned in the last chapter's "data mapping" section.

The implementation of genetic algorithm usually has five steps:

- Create initial domain.
- Generating the offspring.
- Evaluation by fitness function.
- Updating the solution domain.

5.2. Genetic Algorithm on weight choosing

According to the method of this master thesis, the features of the data must be extracted. The features are the symptoms from which the CBM system can diagnose the current state of the fan and predict the failure.

After the observation of the data, the mathematical expectation (MSE) and variance (Var) of the vibration, temperature and electronic current of each time window are the features should be extracted. These features perform unmorally before the failure in the near future, and by detecting this symptom some alerts will be sent at least one time window before, about 2.4 hours.

Before we using the training matrix and the detection matrix to do the machine learning and system status detection, we need to find out an optimum way by which the different types of data can map into the matrix and become an restrict one-to-one relation. The mapping is a very important step which will influence the sensibility of the system and even whether this system can work.

The concept weight is employed before the mapping. The weight can combine the MSE and Var into one value. Using weight to merge the related values into one is a common method to integrate the data. So the there will be one value of vibration, one value of temperature and one for electronic current. Then these three values can be considered as the coordinates in a three-dimension matrix. With this mapping method, the machine learning and detection are capable to be developed.

It is very difficult to define the optimum weights for parameters. The potential domain is too large for the normal value-decision methods. Toward this situation, the GA is introduced to find the optimum weights by genetic random variance.

5.2.1. Chromosome set

The chromosome set is the collection of the chromosomes. The chromosome in the GA stands for the different values the GA tries to seek. In this master thesis, the chromosomes denote the six weights we are seeking. The length of each chromosome is seven. The six chromosomes denote the weights and the last one stands for the key number of chromosome which is used to distinguish the different chromosomes.

The chromosome is stored as the array. And the type of the each element of chromosomes is integer. In this master thesis, we assume all the weights are the integer number from zero to one hundred. For example, one chromosome may like this: [12, 3, 98, 71, 21, 40, 45, 2]. The last number "2" stands for this is the third chromosome in the set, we start to count from zero. The other numbers are the weights we will use to calculate the distance between each type of status; we will explain this in the content below.

The size of the set is defined to five. So there are five chromosomes in the solution domain. That also means there are exactly five parent chromosomes for each generation reproduction. Technically, the size of the set is the larger the better. A large size of solution domain is capable to production more offspring in each reproduction, in another word, there is a higher possibility that the benign genetic variance takes place. However, from the whole of view, the more generation -recycle times also can provide the optimum solution. In this master thesis, the chromosome set size is five, and the generation recycles two thousand times.

This is an example of the chromosome set. There are five rows and seven columns. Each row is a chromosome. The column one stands for weight of the MSE of vibration; column two stands for weight of the Mathematic Variance of vibration; column three stands for weight of the MSE of temperature; column four stands for weight of the mathematic variance of temperature; column five stands for weight of the MSE of electronic current; column six stands for weight of the mathematic of electronic variance; column seven stands for key number of the chromosome.

$$\begin{bmatrix} 86 & 64 & 00 & 41 & 23 & 03 & 21 & 0 \\ 01 & 86 & 73 & 71 & 47 & 38 & 45 & 1 \\ 12 & 03 & 98 & 71 & 21 & 40 & 45 & 2 \\ 61 & 30 & 11 & 44 & 37 & 43 & 51 & 3 \\ 27 & 80 & 21 & 56 & 63 & 78 & 09 & 4 \end{bmatrix}$$

This chromosome set is also the initial set. The initial values are set by randomly. In C#, the random class provides the function of producing a random number is a certain area: "Random r = new Random();

chromesomeSet[i, j] = (int)r.Next(100);". The i and j are the coordinate of each elecoment of the set. The random class gives each weight a random number from zero to one hundred.

The initialization just work once when the program is running. In each round, the finest five chromosomes will be set into the chromosome set and eraze the former chromosomes. The key does not change. And this new chromosome set will work as the parent chromosomes for the next round reproduction.

5.2.2. Chromosome Mutation

Mutation is a genetic operator that alters one or more gene values in a chromosome from its initial state. This can result in entirely new gene values being added to the gene pool. With these new gene values, the genetic algorithm may be able to arrive at better solution than was previously possible. Mutation is an important part of the genetic search as help helps to prevent the population from stagnating at any local optima. Mutation occurs during evolution according to a user-definable mutation probability. This probability should usually be set fairly low. If it is set to high, the search will turn into a primitive random search. [22]

In a general GA, there are several types of mutation:

- Inversion. This mutation is suitable for the binary chromosomes. The chromosome may change the polarity to the opposite one by a possibility. For example: 1 **0** 1 **1** 0 ->1 **1** 1 **0** 0.
- Substitution. Substitution is the two chromosomes exchange their positions in order to produce the new chromosome. For example: 30 **16** 3 **2** 10 ->30 **2** 3 **16** 10.
- Update. Another kind of mutation is update. In the process of update, each element of chromosome randomly update to another value with a low possibility. For example: 30 16 3 **2** 10 ->30 16 3 **17** 10.
- Insertion. The last one is insertion. There are two steps of this process: first one is choose an element in a chromosome randomly; then insert this chosen element into a random position, all the elements after this position shift. For example: Choose the "2" of the chromosome 30 16 3 **2** 10. And insert this "2" into a random position: 30 **2** 16 3 10.

In this master thesis, the substation mutation is employed. In each generation cycle, all the genes in the five parent chromosome have a 1% chance to be substituted by another random integer value from zero to one hundred. In the program, there is an instance of the Random class of C# called "mutationPossibility". It is a random value from 0 to 100. And the threshold of the mutation is set as 99. When the mutationPossibility is bigger than 99, the logical judgment will run the mutation process. Otherwise, the mutation is in sleep.

The outcome chromosome set of the mutation process will be evaluated. Compared with the original chromosomes, keep the finest five chromosomes as the parent chromosome set for the crossover process and discard the others. The whole chromosome set will be reorganized by the rank of the evaluation. The best one will be stored in row one, with the key is zero. The worst one will be in row five, with the key is four. When the new solution domain is decided, the mutation process is over in this production cycle.

5.2.3. Chromosome crossover

Crossover is a significant part in the genetic algorithm. It plays a key role in GA. The main part of the genetic gene drift is from the crossover, and only a small part of the mutation process. The traditional crossover method contains the one-point crossover, two-point crossover and uniform crossover. The following figure shows these three types of crossover:

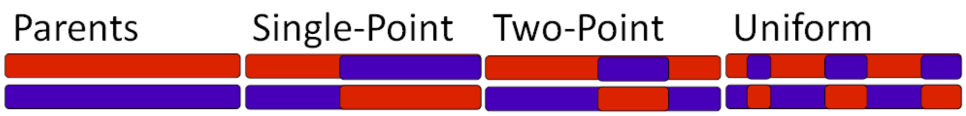


Figure 32 Three kinds of crossover

In this master thesis, the uniform crossover is selected. Because the uniform crossover can more delegate the gene's random variant in the offspring domain.

In the implementation, each chromosome has to mate with other chromosomes and produce two child chromosomes. These two child chromosomes will be evaluated and compared with the solution domain's chromosomes, all five parent chromosome rather than two. If the child chromosome is better than the ones in the solution domain, it will replace the poor chromosome in the solution domain.

However, some parent chromosomes are good. Maybe not as good as the offspring chromosomes, but we still keep the best one in the parent chromosome set. In another word, there always is at least one finest chromosome from the last generation's solution domain is kept.

When the crossover process is finished, the chromosome set will be also reorganized by the rank of the evaluation as the final process of mutation. As a result, in every generation, the optimum chromosome is always stored in row one of the matrix.

5.2.4. Offspring optimization

When the offspring are produced, they are judged by the evaluation method. The evaluation method is specified by different cases. The simulation should be capable to diagnosis the features and orient to the target.

In this master thesis, the target of the GA is to find out the appropriate weights for the matrix mapping. As the discussion in content above about the diagnosis method, the distance, between each two different types of status, is the larger the better. Because what we concern is the suspicious statuses, the statuses near the borders. With this purpose, the evaluation is to calculate the distances. The function of the distance calculation between point A and point B in a three dimension space is:

$$\text{Distance} = \sqrt{(X_a - X_b)^2 + (Y_a - Y_b)^2 + (Z_a - Z_b)^2}$$

The training data is the data we have known the values and the status. When the system is built, it's necessary to use some known data to let the system to learn the pattern or set the some parameters of the system. Our training data has the all the values and status of twenty seven points. There three different types of statuses: normal, error1 and error2.

In order to calculate the distance easily, we can get the average values of each type's MSE and Var as the status points. As the result, there are three points in the space now: normal, error1 and error2.

The following set of functions is how we map the values into the three dimension coordinates:

$$\begin{cases} X = a \times MSE_{vibration} + b \times Var_{vibration} \\ Y = c \times MSE_{temperature} + d \times Var_{temperature} \\ Z = e \times MSE_{current} + f \times Var_{current} \end{cases}$$

The a,b,c,d,e,f are the weights in the chromosome, and MSEs and Vars are known as the statuses points. So we can easily to calculate the distance between statuses.

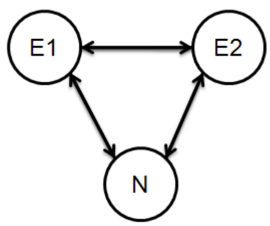


Figure 33 Distances between different statuses

In the figure 33 you can see, there are three distances: distance from Normal (N) to Error1 (E1), the distance from Normal (N) to Error2 (E2), the distance from Error1 (E1) to Error2 (E2). We rather concern the distance between the N to E1 (distanceN-E1) and E2 (distanceN-E2) not the one from E1 to E2. And we want the lengths of distanceN-E1 and distanceN-E2 are similar and also larger than the distances between the same status such as the distance between two normal points (distanceNN, which is discussed in the following content).

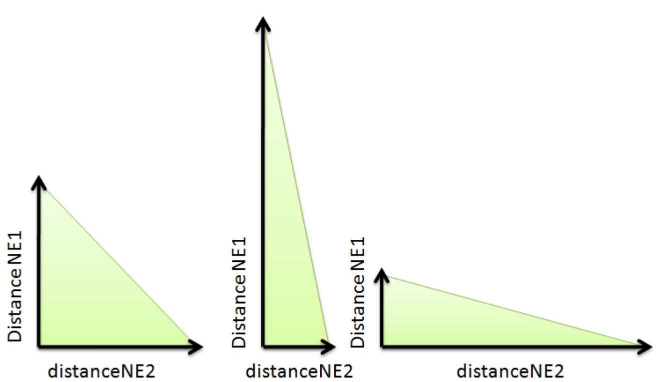


Figure 34 Relation between distanceNE1 and distanceNE2

When we only consider the distanceNE1 and distanceNE2, the figure 34 shows the relation between them. In some certain area, when the green triangle's area is largest, the two right-angle sides are the same. With the same theory, the evaluation should concern the distanceNE1×distanceNE2 rather than just the value of each distance.

Another factor has to be considered is the distanceNN. All the distances in this master thesis are relative values. They are meaningless except to be compared with each other.

The distanceNN is the average distance of all each two normal points. The distanceNN works as an adapter of the fitness function, to let the fitness function concentrate on the comparable distance rather than the exactly value of distances. Without it, the weights will all be ninety-nine, and that will make the GA fail.

With the consideration of both factors of the different types of distance, the final evaluation function in GA's implementation is:

$$\text{Evaluation Factor} = \frac{\text{distanceNE1} \times \text{distanceNE2}}{(\text{distanceNN})^2}$$

The evaluation factor is what we concern. Our target is go find the optimum set of weights to let this evaluation factor to be larger. All the parents and offspring' chromosomes are evaluated by this function. And the best five chromosomes are kept, the others are discarded.

5.2.5. Discussion of the fitness function

In this thesis, the fitness function is to evaluate the comparable distances between normal with error1 and error2. As mentioned above, in order to make the solution not to be too complicated, I just use the average values of normal, error1 and error2 to calculate.

The advantage of this decision is to make the algorithm to be simple. In this thesis, the data is comparably not random.

The disadvantage is, in reality, there is no guarantee that all the data has no overlap. In many cases, many normal points may locate in the same area as error, or the same type points do not locate closely. If these situations happen, it is no long to be suitable to just use the average value to simplify the fitness function.

Also in this thesis, the domain for the chromosomes is small. Technically, the size of this domain is the bigger the better. In this thesis, the size is five, that because the GA I used is not complication, and the area for solution is not wide. When the number of generation circle is big enough, the optimum result still will come out.

5.3. Implementation result

The figure below shows the initialization process of the chromosomes. There are five chromosomes in the solution domain. When the programming is running, the program starts from the generation zero. The values are set by the random values:

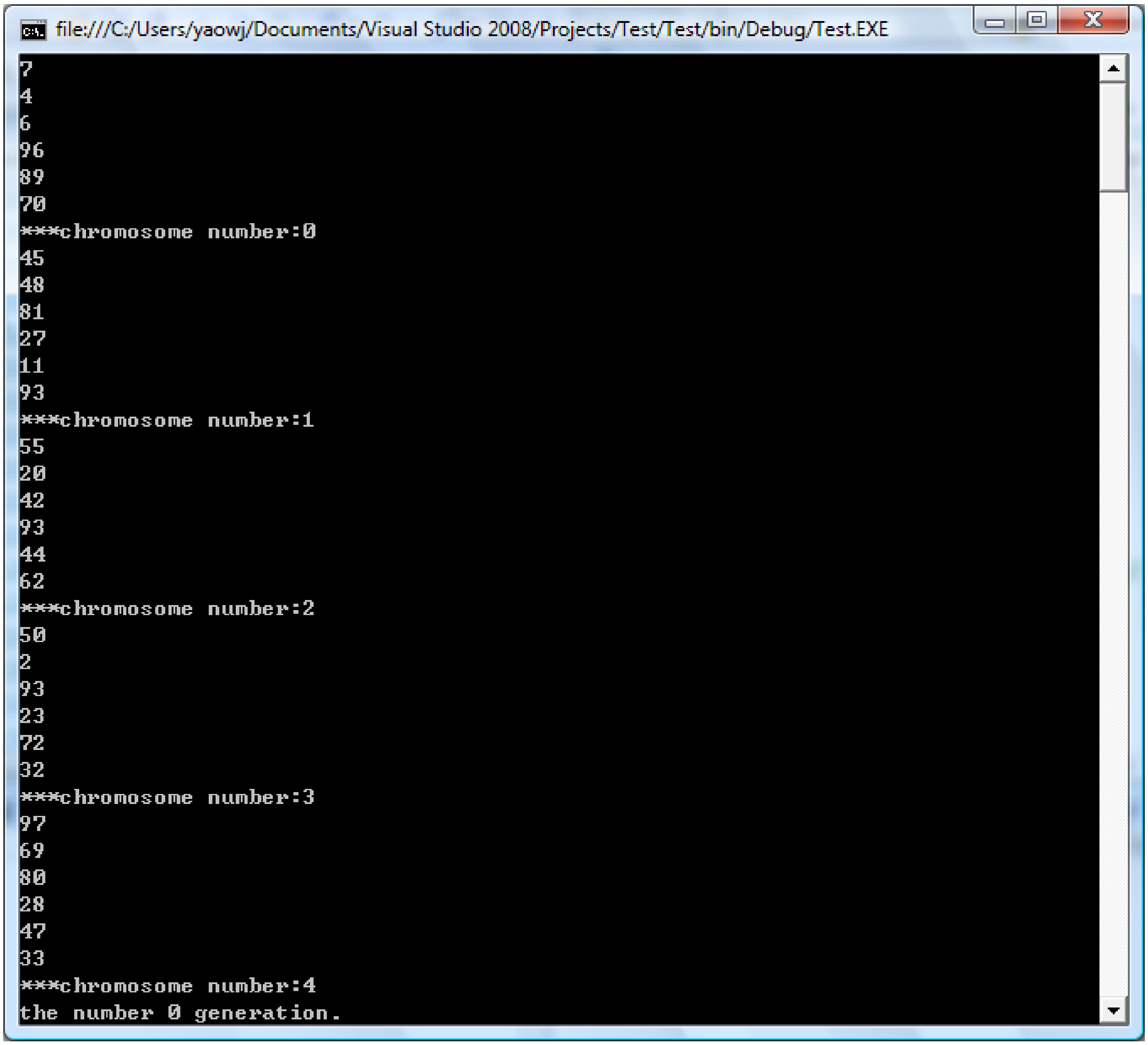


Figure 35 Chromosome initialization

The mutation is one gene drift approach in GA. The figure below shows the mutation. One of the chromosome's value changes from 68 into 60.

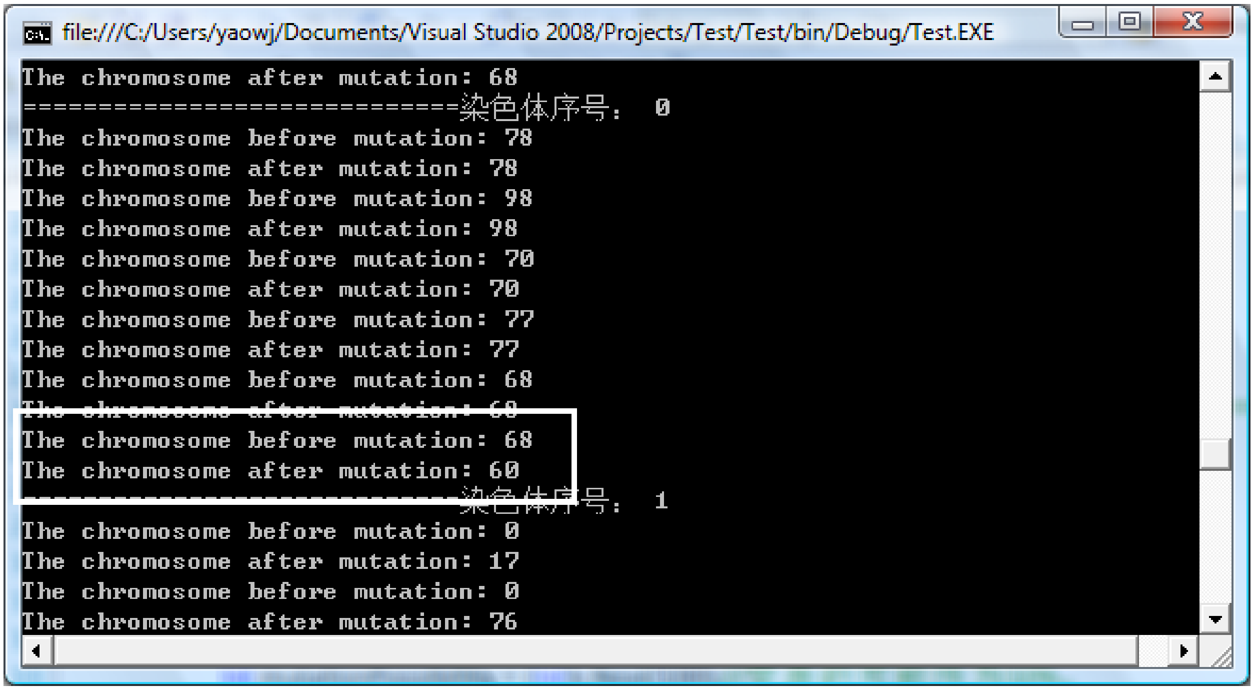


Figure 36 Mutation

After the mutation, the distances are recalculated and compared with the distances of the parents chromosomes in the solution domain. The distance calculation acts as the fitness function.

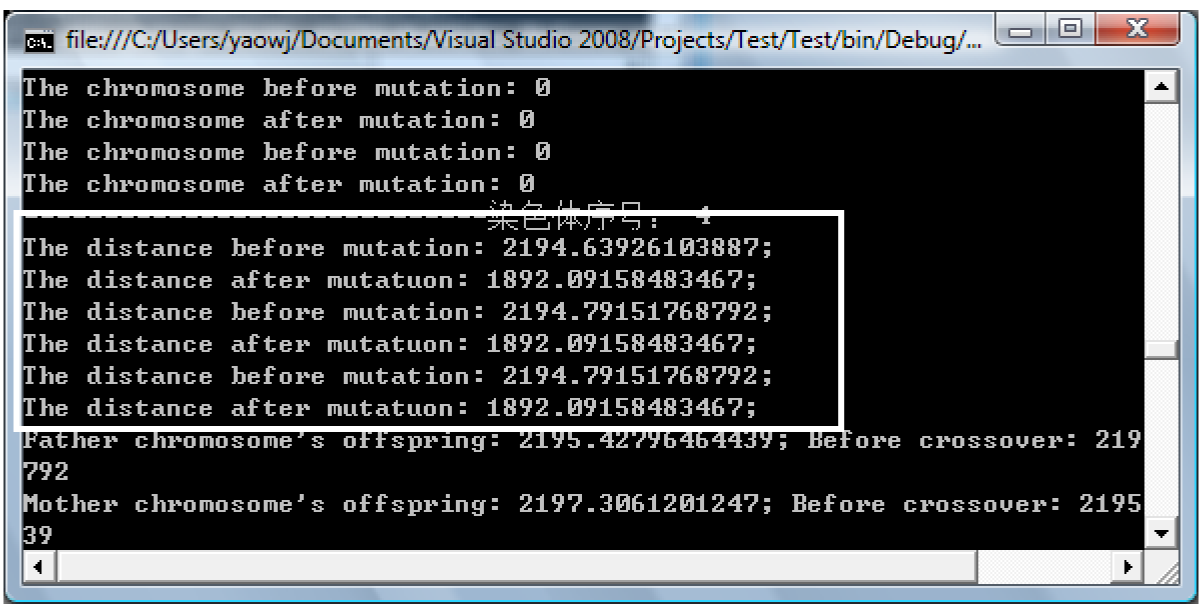


Figure 37 Evaluation for the offspring of mutation

The crossover is the main gene drift approach. The offspring's chromosomes also processed by the fitness function. The finest ones will be kept and the others are washed out:

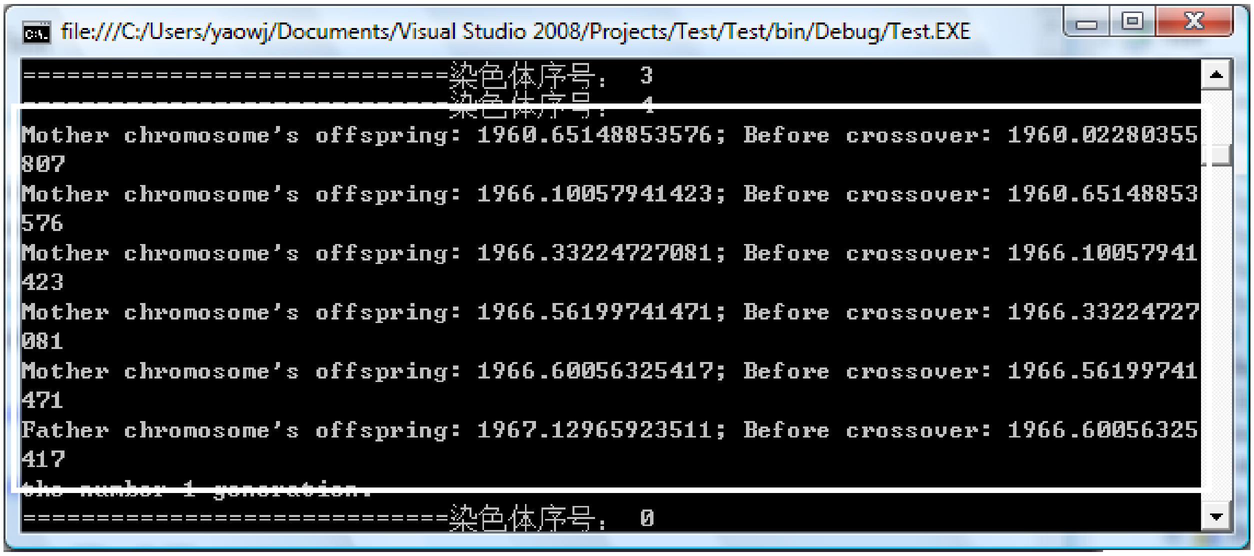


Figure 38 Evaluation for the offspring of crossover

After 3000 generation's refine process, the final result of the weights choosing program is:

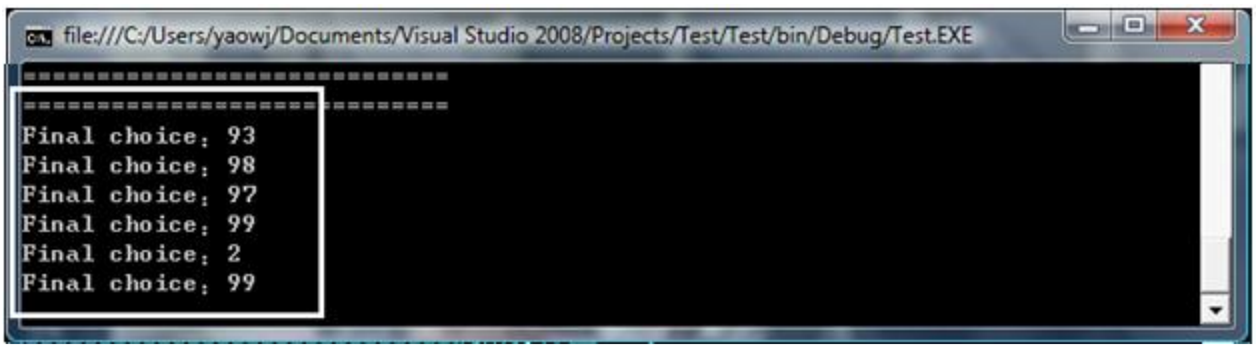


Figure 39 Optimum result from GA after 3000 generations

5.4. Result discussion

The final result of the GA implementation shows us that all the parameters are almost the same important, except the average value of current which is nearly unnecessary.

The function of mapping should be rewritten into:

$$\begin{cases} X = 93 \times MSE_{vibration} + 98 \times Var_{vibration} \\ Y = 97 \times MSE_{temperature} + 99 \times Var_{temperature} \\ Z = 2 \times MSE_{current} + 99 \times Var_{current} \end{cases}$$

6. System status detection approach

In this chapter, the procedure of the system state detection and the software implementation will be introduced in detail. The detection is the core part of this system. It will send the current system status to the GUI which can show it to the personnel.

6.1. Overview of the detection

When the training and initialization steps are over, the detection is simple. Firstly, the pre-processed data is sent in the system as the input. Then the system finds the corresponding state point (the responding location) in the detection space or detection matrix. At least, the state of the corresponding lattice will be assigned to the current state as the diagnosis result. Before the detection matrix is finished, this system uses the learning from neighbors as the intelligent method to learn from the known experience. The following figure shows the process when the detection matrix is completed:

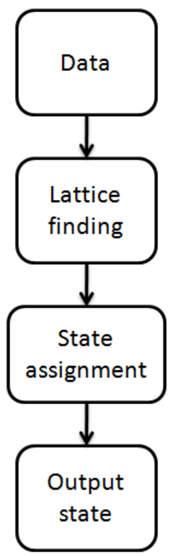


Figure 40 Simplified process of detection process

The whole process of the detection, which contains more steps about the detection matrix, is shown as figure 41:

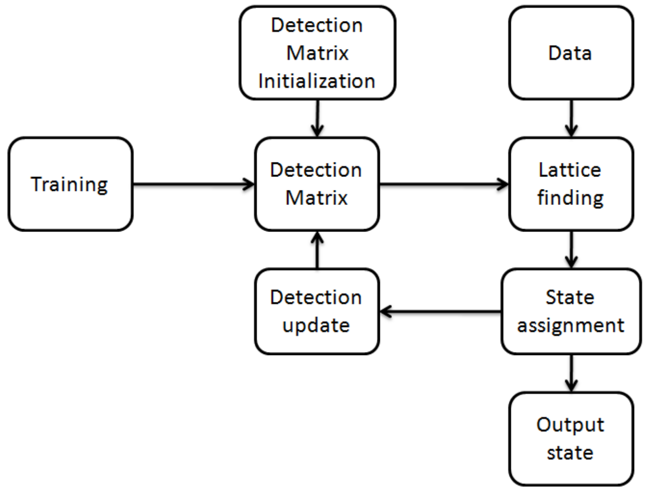


Figure 41 Complete process of detection

6.2. Detection matrix initialization

As discussed in the chapter 3, all the logical calculations of this system are not in an infinite value space as in the reality. There are some system boundaries, which are pre-set by the system designers, to limit the state area and the quantity of the system states. Without these necessary limitations, it is hardly to calculate so huge quantity values by the computer.

The system boundaries are set based-on the system's history. In this master thesis, the span of the vibration value (mse*A+var*B, A and B are the weights) is from 35 to 155, the span of temperature is from 137100 to 318000 and the span of electronic current is from 133000 to 243000.

These values of the boundaries are chosen by the system's history. The designers have to observe the data and find out the area of all the normal states are in. Some states in this space maybe not the normal state, but it is must be made sure that all the states outside this space are definitely the unmoral states.

And also the values of the boundaries are not from the raw data directly. Some necessary transforms of the data are done. The vibration's mse times 100; the vibration's var times 40,000; the temperature's mse times 10 and the temperature's var times 2. These values transformations aim to make the values after calculated with weights are in the similar level of magnitude. The distance calculation, which is introduced in genetic algorithm calculation to find out the optimum weights, only is meaningful when all the values are in the same level of magnitude.

According to the method, the whole cube space is divided into 1000 small lattices (10*10*10) evenly. After this division, the detection matrix is initialized as the figure 25.

Each point in the figure 25 stands for a lattice with a unique coordinate, which from (0,0,0) to (9,9,9). The system store these lattices as a matrix this is the reason this space cube is also called space matrix or detection matrix. The status of each lattice could be 0, 1, 2 or 3. "0" means the status is unknown; "1" means this status is a normal state; "2" means this status is error1 and "3" means this status is error2.

So, each point has an ID such as (3, 5, 2, 0), which means the lattice is (3, 5, 2) and its status is “Unknown”.

In this master thesis, the states or lattices near the normal and fault status are concerned. This is also the evaluation guideline of the GA for choosing weights. The lattices beside the system boundaries contain much larger value area of reality. That means these lattices have high possibilities to have some fault status. However, these lattices are not concerned by the system. So the system’s detection results, which only focus on the lattices which can predict the fault rather than after the failure takes place, will not be affected by these lattices.

6.3. Training

Training in this detection process is a step of learning from experience or the training data. The learning process of a system is a kind of subfield of the artificial intelligence in the computer science. It is an approach to improve the system’s performance over time based on the known data, for example, the training data or the history of system.

The training part in this master thesis, which loads the known training data into the system, is the first step of the system learning. The reason we need the training is that usually the designers cannot not to set all the states for the system before something happens. That means, some time the designers are not capable to exhaust exploit all the states. So the optimum way is to let the system behavior as a human who only know the limited knowledge but when the new situation takes place the new solution or idea can be illuminated by the experience and history.

In this master thesis, the training data is composed by 27 points. These self-organized 27-points data in a 3*3*3 space is as figure 24.

The data is mapped into the detection matrix according to the mapping method with the weights from GA:

$$\begin{cases} X = a \times MSE_{vibration} + b \times Var_{vibration} \\ Y = c \times MSE_{temperature} + d \times Var_{temperature} \\ Z = e \times MSE_{current} + f \times Var_{current} \end{cases}$$

Before the training, all the 1000 statuses in the detection matrix are mark as “0”-state unknown. After the training process, there are 27 points are marked as “1”,”2” and “3”. The training data must contain all kinds of known states. In this master thesis, if there is one state is not shown in the training data, the system will never detect this type of fault unless the personnel adjust the detection matrix manually.

It is very important that there must be some known training data for the machine learning or pattern resignation. There is no way to let the system to detect the state without any history or experience. So before the system starts to work, the training data must be loaded.

6.4. Status seeking

The current system data is mapped into the detection matrix firstly; then the system will return the state of this lattice as the status of the current system. There is one and only one lattice can be mapped into by a set of system data.

6.4.1. Lattice with known status

After the detection matrix evolves into the matrix that all the lattices with some known status, the process is quite simple: the system just needs to find out the lattice by the coordinates and return its status. The coordinates and statuses are stored in the same array in the system. So this process is also very fast. This process is exactly as the process in the first figure of this chapter.

Before the detection matrix is totally completed, only a part of the matrix has the known statuses. This partial completed matrix treats these known lattices as the lattices what in a totally completed detection matrix.

6.4.2. Lattice without known status

Before the detection evolves completely by the rule of machine learning, many lattices in the matrix have no known status. As a result, when the data is mapped into these lattices, the system has to search the neighbor lattices to try to find out an appropriate status for the current lattice by the rule of machine learning.

One concept must be introduced here- lattice distance. The lattice distance is used to indicate in what area the system is looking for the neighbors. The next figure is an example in one-dimension:

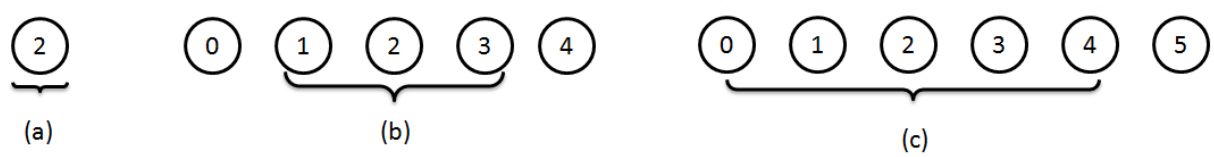


Figure 42 Distances in one-dimension array

When the data is mapped into the second element of the array in figure, the system searches the neighbor with distance is 0, that means itself as shown in figure 42 (a). If there is no known status, the system searches neighbors with the distance are 1, as shown in figure 42 (b). All the statuses of these neighbors are recorded. Then the system calculates what the majority status is. The majority status will be assigned as the status of the second element which has a known status now. If there is no majority known status in the situation of distance is 1, the distance increases as figure 42 (c), until one known status is assigned. This process is illustrated by figure 43:

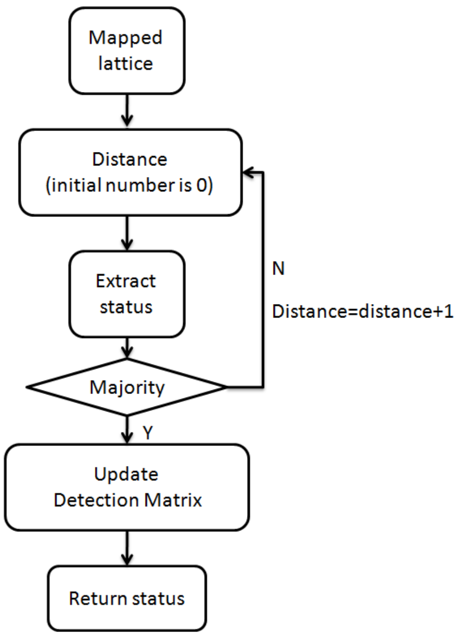


Figure 43 Status finding approach

In this master thesis, there is no guarantee that there is always a majority status. The existence of that all the neighbors, no matter with what lattice distance, have the same quantity of statuses is possible. In another word, there are two known statuses always have the same quantity of lattices in any distance, so there is no majority status. In this situation, if one status is "normal" status, then the other status will be assigned to the uncertain lattice; if the statuses are "fault" types, one of them will be assigned randomly to the uncertain lattice. This decision is shown as below:

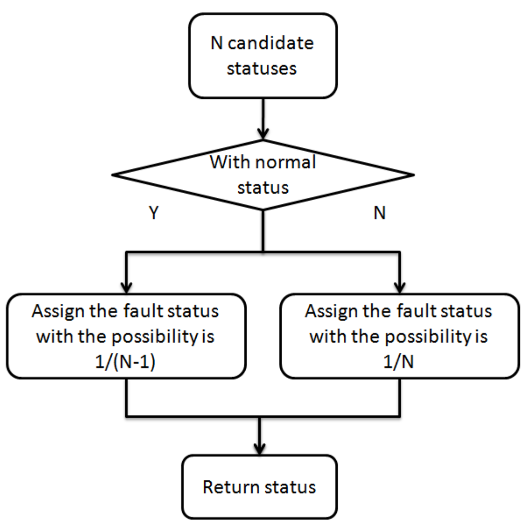


Figure 44 Status decision process

The quantity of the neighbors with the same distance varies for different lattices some time. To the lattices near the system boundaries, the quantity of neighbors is dramatically less than the ones in the middle of space.

6.5. Alarm detection

The detection method introduced above can apply an effective and possible way to detect the normal status and the fault status. However, this method cannot work alone if we the want to distinguish an occurred fault and a fault in near future. The following figure shows the status the method can distinguish, but cannot sense the difference between the prediction and happened fault.

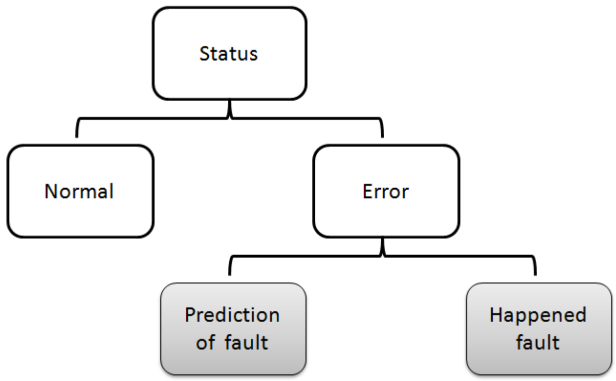


Figure 45 Statuses the method can sense

As the aim of a CBM system, we focus on the fault prediction rather than the fault which has already happened. The traditional detection system is capable to monitor the normal and fault status. So the best for the system is to employ the detection method and the traditional approach at the same time, let them work together. After this improvement, there are three statuses for the system: normal, alarm and failure. The alarm status is the significant shinning point of the method. It can buy the time for the engineers to maintenance the system which is monitored and specify the potential fault type which also can reduce the cost for the company.

6.6. Disadvantage of this method

The whole CBM detection solution can work well in some situation. Unfortunately there are still some flaws. These flaws may decrease the sensitivity of the system and increase the negative alarm.

The first disadvantage is that the training step and the detection or space matrix detection step are based on the machine learning. And there is no guarantee that this learning process will become to a correct way. The learning process can be affected by the sequence of the learning which can directly impact the final result. So there is still some need of the manual work involved which can adjust the matrix's status.

The second disadvantage is this system only can detect the current state of the monitored system. In another word, the whole method assumes the system's next state is only related with the current state, no former states are involved. In reality, this assumption some time is not true. The history of the system also can release some abnormality even the current status is normal. In the chapter seven, the pattern reorganization method will be introduced as an assistant method works with this approach.

The third disadvantage is that this method is not adjustable except the training process. The detection or space matrix should be flexible. This is not only because the learning procedure is not always correct, but also because the system's normal statuses are not always the same. For example, this software system may be used by twenty machines. These machines must have tiny difference. Besides that, the system's status also may change from aging or component changes.

After all, a successful CBM system which can work alone without any human impact is very complex. This master thesis just develops a comparable easy case and method to research. There is still a long way to go if this solution can be used in reality.

6.7. Implementation result

When the detection programming is running, the first step is to load the training data. This step is the initialization of the detection matrix. In the following figure, the first three columns stand for the location of the training data, and the last column stands for the status of the state.

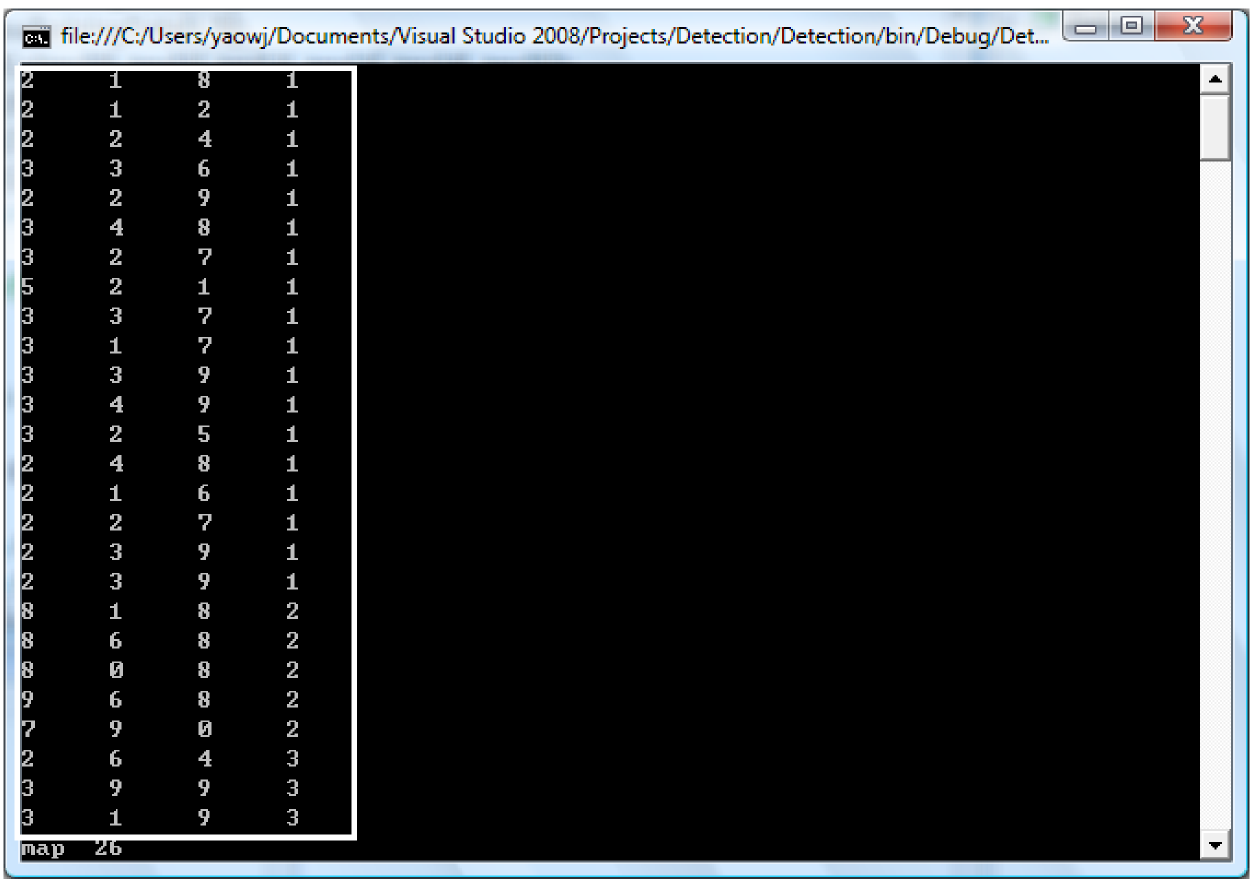


Figure 46 Load the training data

In this master thesis, there are eight normal states, five error1 states and three error2 states. As I introduced before, there are twenty-seven training data, and some of the data will mapped into a same detection matrix lattice. This is the result the number of the trained detection matrix lattices is less than twenty-seven, the number of the original training data.

When the data of the current system's state is input, the system finds out the corresponding location in the detection matrix.

Figure 47 Coordinates of the sample data

The next figure shows the final result of the detection:

Figure 48 Final result

The final result is "1", which means the system is normal.

7. Graphical interface solution

This chapter introduces some background of how to design a good interface. And the interface solution for this thesis is presented at 6.4. For software, the functional part and the interface part are both important.

7.1. What is GUI

A graphical user interface (GUI) is a type of user interface which allows people to interact with electronic devices such as computers; hand-held devices such as MP3 Players, Portable Media Players or Gaming devices; household appliances and office equipment with images rather than text commands. A GUI offers graphical icons, and visual indicators, as opposed to text-based interfaces, typed command labels or text navigation to fully represent the information and actions available to a user. [24] The actions are usually performed through direct manipulation of the graphical elements. [25]

7.2. Guild lines for a good GUI strategy design

The feeling of GUI is the first thing the producer passes to the customers. So a successful visual design of course can give the user a nice impression. When designing a GUI, the designer should make it to be [26]:

- Clear. The job of a GUI is to let the customer can communicate with the software. If the GUI makes the customer be confused, definitely it is not a nice art of design.
- Concise. If you can meet the need by one function then do not use three functions. To be concise also means make the complex or advanced functions behind of the usually used functions. A very good example is the volume adjustment function in the down right corner of the Microsoft windows operation system.
- Familiar. Do not try to design a strange GUI to the common functions. If the function is almost the same, there is no one like to try the totally different style of GUI. Generally speaking, challenging the patience of customers is not a good choice.
- Responsive. The customers should know what is going on with the software, or at lease let the customers assume they know what is going on. When the software needs some time to make the response, the GUI should show some “loading…” or “Please wait…” information rather than nothing.
- Consistent. Some productions are related or in a suite. In this situation, the GUIs should be as similar as possible.
- Attractive. The word “attractive” means different to different persons. However, there is no doubt that, if there is no big functional difference, the customers prefer to choose a good looking production. This element may improve the impact of a brand inconspicuously.
- Effective. The GUI should make the customers can do they want to do as easily as possible. So it is a good idea to let the customers can find the high using ratio functions in the comparable fixed position and just need one click.

- Forgiving. This is both the GUI and function design. Anyone can make mistakes when operating the software. The system should let the customer have a chance to "redo" or "confirm" when something important is about to happen or it has a high mistaken ratio.

In a word, there are no hard rules for a good GUI, but when we compare two GUI designs, anyone have a favorite one. The designers should make the GUI design meet the most of the potential customers. Some time it is also necessary to collect the responses from customers then make the improvement. Anyway, the GUI is the only thing the customers can see and feel. So try everything to let the customers be infatuated with your design is necessary.

7.3. Characters of a good GUI details

In the nowadays society, every best-selling production has good visual design. A good visual design not only can distinguish the productions from different companies, but also can upgrade the fame popularity of the company. The clients should feel the production is fully designed in each detail from the GUI. And also the clients may consider their feelings are cared about by the thoughtful production. The following content introduces the characters of a good GUI design in detail parts in Microsoft software productions.

7.3.1. GUI design of initialization

When the software system is active, the first GUI is the initialization GUI which only is used to be shown during the time that computer responses the client's initialization set and launching time. This GUI should give the clients some basic information of the software clearly such as the name of the producer, the logo, the version, the copyright announce and the website of the producer. This is very important for the producer to establish the impressive figure which can stimulate the sell most directly.

This GUI also must consider the screen size of different monitors and the different color format for different operation system. So generally this form of the initialization GUI's size should be 1/6 of the whole screen and the colors should be less than 256 colors.

7.3.2. Architecture design

The architecture of the GUI refers to the functions of use. The designers must know all the functions of this software so that let the clients can find all the functions and settings easily. The architecture should be concise and neat, less the unnecessary elements and colors. It also must consider the screen's zoom. Usually the logo of the producer is in the upside of the window or the top left corner, the main menu should on the top left and the scroll bar on the right side. The designers should meet the software consuming.

7.3.3. Button design

The software buttons should have the interactivity. Usually there are six statuses: click, mouse on it but no click, mouse not on it and no click, after click the mouse on it, cannot click (mute) and auto-change status. The icon of a button needs to be concise and neat too. The style of one set of buttons should be the same.

7.3.4. Layout design

The layout should have the zoom in and zoom out function. When the client is zooming the form or window, the layout shouldn't have big change and no button should be missing. The layout of different functions should be clear. The dialog forms and pop-out forms should be used in order to save the space and facilitate the window switching.

7.3.5. Menu design

The menu should on the top of the form. It should have the chosen and un-chosen statuses. The name of the function should be on the left, the responding hot key should on the right side. If there are some sub menus under one function, the arrow should be involved. Between the different function parts, there should be a line which can clearly divide zones.

7.3.6. Icon design

The icons in a software system should have the simple color with no more than 64 colors, and have the size of 16x16 or32x32. The icons are the small size of visual art. A good icon should have strong visual impact and reference to the function.

7.4. Thesis solution

The next figure is the main form of the CBM system's GUI solution. This form is clear and neat.

Figure 49 Main form of the CBM system

In the middle of the window form, the statuses of the system (the system's whole status) and potential errors are very clear. The status presentation is designed as a traffic light: the red light is the failure, the yellow light is an alarm and green light means everything is ok. There are two errors: error1 and error2. In the GUI, the error should have an easy understandable and neat name, especially for the junior engineers. In this thesis, the name of error1 is overheating, and the name of error2 is too much dust. So the second traffic light shows the status of the fan's heating. When the fan works well, the green light is on; when the fan is going to be overheating, the yellow light is on; and after the failure takes place because of overheating, the red light is on. The person, who is monitoring the system, should send the engineer to fix the potential failure or temporally reduce the workload of the fan. When the red light is on, the system should shut down the machine automatically in order to avoid more loss.

But there is a big problem: some people are not sensible to the red and green! In order to let the interface work well to everyone, some changes are made:

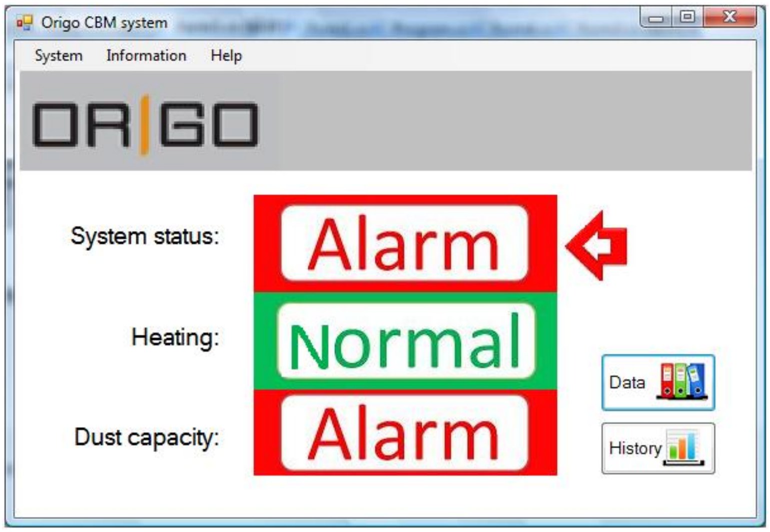


Figure 50 Improved GUI

In the figure 50, the traffic lights are replaced by the “Failure”, “Alarm” and “Normal” signs in black, red and green respectively. And also when the system status is not “Normal”, there is an arrow on the right side of the system status sign to notice the potential danger.

Figure 51 gives an example that the fan may be caused a failure in the near future because of too much dust. In this situation, the engineer should receive the alarm sent by CBM system and clean the dust in the fan as soon as possible. If the dust is not cleaned in time, a failure will happen. The figure below shows the GUI when the failure happens and the fan stop working:

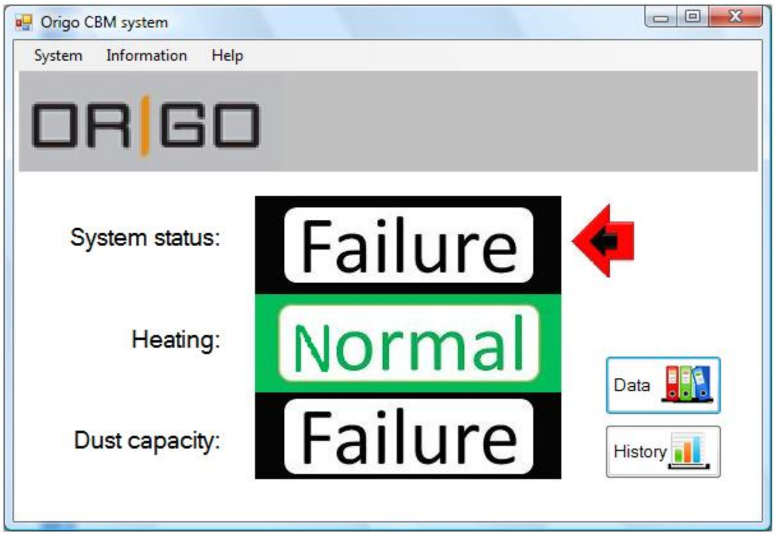


Figure 51 When failure happens

On the top of the form, it is the menu bar. The options in menu bar can control all the functions. It makes it is possible that the client can still control the GUI without mouse of PC of scroll wheel of mobile phone. Under the "System" option, there are two options: "start the CBM system" and "close the CBM system". These two functions can start and close the CBM system's work status. Under the "Information" option, there are "data" and "history" options, the same functions can be operated by the buttons on the left. The last one is the "help". In this option, the version information and the instruction of the software platform will be found.

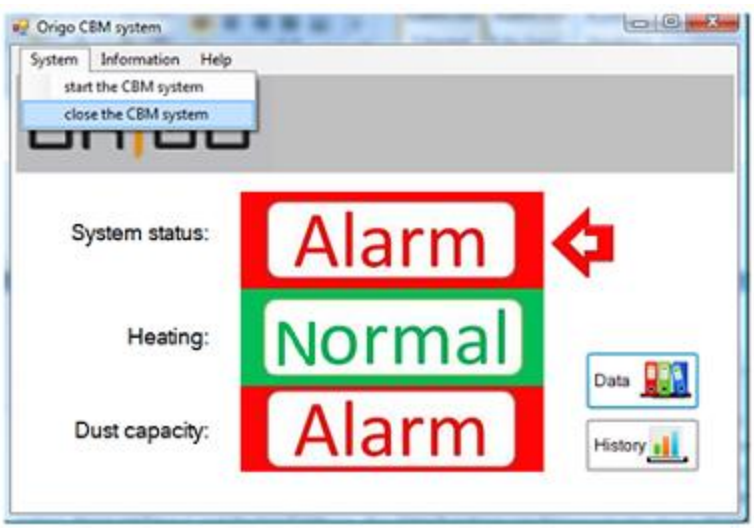


Figure 52 Menu bar

By clicking the button "data" or the option "data" in "information" in the menu bar, the current data will be presented in a pop-out window.

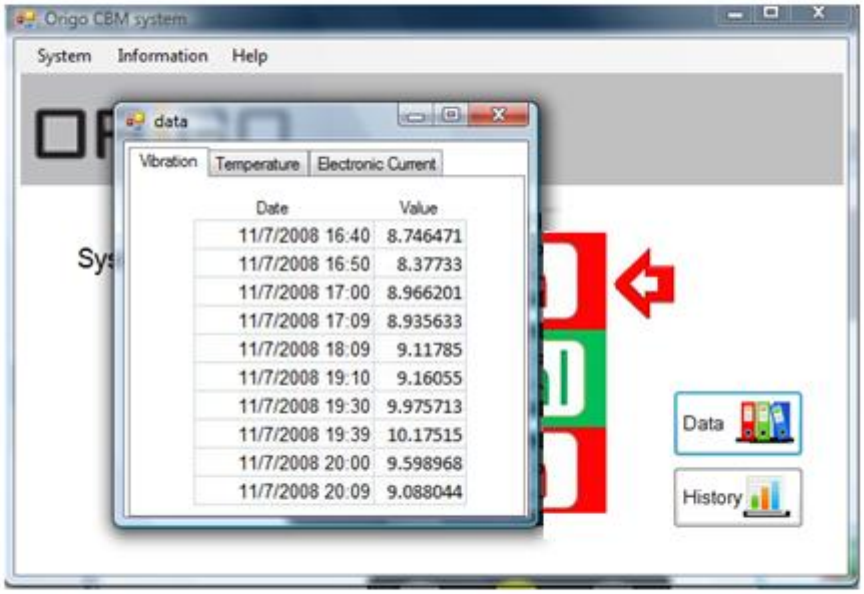


Figure 53 System data

In the window of data, there are three tabs. In each tab you can see the date and values of the system in detail.

By clicking the button "history" or the "history" function in the menu bar, the client can check out the data's history trend in chart. This chart may offer the client a much easier way of what happened to the system in certain time ago until now. The tabs help the client to change the view between the vibration, temperature and electronic current.

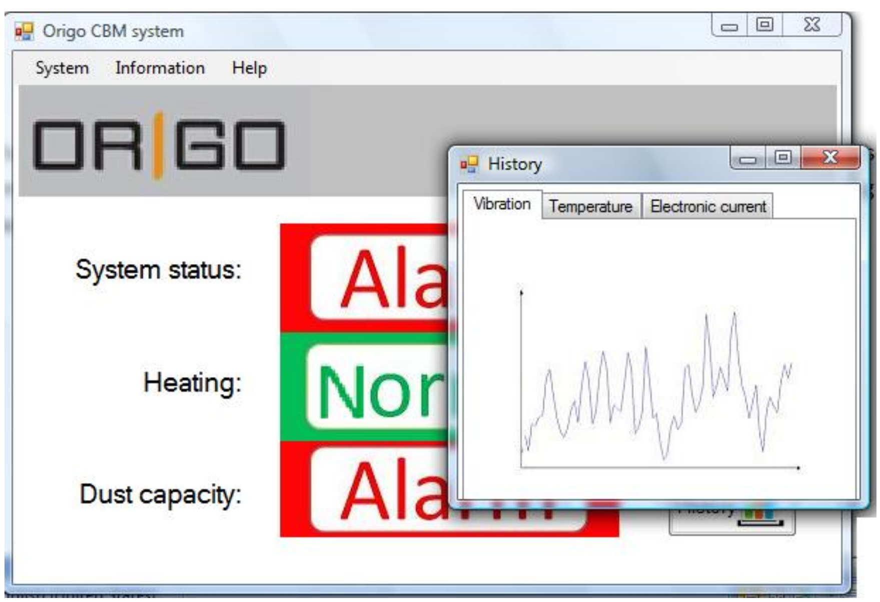


Figure 54 System data's history

The additional information such as the version information and the instruction of this software system are presented as below. These windows can be activated by clicking the corresponding function under the “help” option in the menu bar.

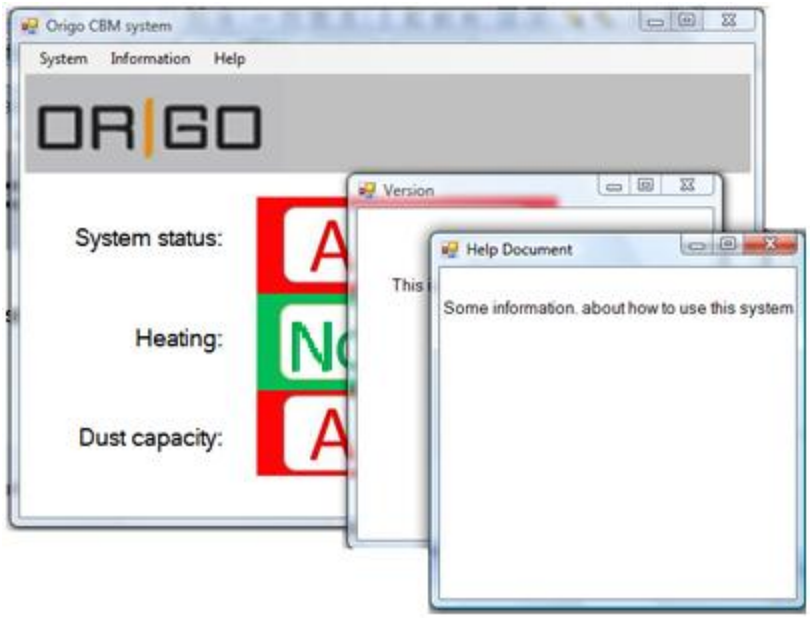


Figure 55 Additional information

At last, but not least, by clicking the logo of Origo company, the client may access to the website of Origo which can apply the client more information about the producer.

8. Conclusion and future work

This chapter concludes the research outcome and the benefit I have received during this challengeable engineering research project. The solution of this case still has many disadvantages. The future work part of this chapter presents many other methods to improve the current solution.

8.1. Concluding remarks

In this master thesis, I have developed a partial solution of a CBM software system for a fan which is used to transport the dusty air from two filters in industry. The target of this master thesis is to build a system which can predict the potential failure in future.

I have studied and researched several intelligent system methods, and I integrated some methods to build an optimum diagnosis approach to the fan. The software implementation made me to understand the software architecture and programming skills even more clearly. This engineering development methodology has given me the experience of development and analysis of projects. This significant progress definitely will benefit me in my future career.

The graphical user interface in this master thesis can apply the personnel, who are monitoring the system, a real-time and clear statement of the current state of the fan. It makes the personnel just needn’t to have too much experience about this field or senior skills. A good software system not only demands the good analysis ability and powerful functions, but also needs easy to be understood and operated.

All in all, the objectives of this master thesis are had been achieved. Of course, the system I built still need to be tested and improved by the real fan in reality. Anyway, this engineering developing process has taught me how to build an efficient system in reality.

8.2. Future work

The partial solution CBM system works well when I simulated on my computer. However, a real system in reality still needs more improvement. What I implemented is just a prototype for testing my analysis method. I have more ideas about how to expand and improve the applicability of this CBM system. Suggestions for future research are discussed in the following context.

8.2.1. Distributed CBM system

The CBM system is an applicable software system which is employed to monitor the state of complex systems such as production line or airplane. So the best architecture for this system is distributed architecture. The figure bellow shows the details:

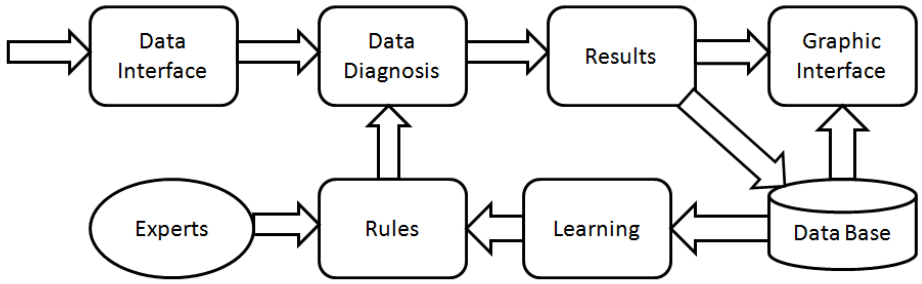


Figure 56 Architecture of a distributed CBM system

The data, no matter the raw data or pre-processed data, for the diagnosis is not from the sensors directly. The system should only accept the data in the strictly format. This format needs some data format standards, for example the company data standard or an industrial standard. This is very important. Because we need to make sure this system can be used not only for a specific machine. This system must be able to deployed in different machines or accessible by different sensors made by different companies. So maybe the CBM system producer should also develop some Application Programming Interface (API) suite to let the system to be much flexible to be deployed.

The data diagnosis method should be more than only one as the diagnosis method in this master thesis. Each feasible method has its advantages and disadvantages. Some methods work together can conclude more stable and exact predictions. In the content below, another method will be introduced as an assistant method which I believe can improve the performance of the CBM system.

The diagnosis results are presented to the personnel by the GUI and also are stored in a database. All the results should be stored. Because the result data is about the history of this system. The system adjusts and evolves only based on its own history. After some time of adjustment and evolution, different CBM systems for different machines in the same type will be different. The unique characters will be remembered.

The evaluation rules for the diagnosis may changed from both the experts and learning process. The experts have the right to adjust the rules. The experts can change them manually or by updating the rules. For example, the CBM systems are deployed on one hundred helicopters. When one system of a helicopter detected a new type of error, this error is not only learned by its own system, but also updated to other ninety-nine systems. The rules can be changed by its learning process as well.

The significant advantages for disturbing are:

- System update. It is important that software can be updated. The CBM system also needs to be attached by the producers. The new type of errors should be diagnosed by the developers at the first time then update this new error to all CBM system. This is kind of like the firewall update.
- Share database. The database is a place stores all the raw data, processed data and diagnosis results. The data is usually kept for years. So maybe some machines do not have room to install the big and safe database. So maybe a remote database is the optimum option for some cases.

Sharing the same database also make the developers data mining work comparable possible and easier.

- Remote control. For some cases, it is not convenient or possible to monitor the machine or watch the screen of GUI all the time. The GUI should be deployed on the network which is easier to be accessed via PC, PDA or even cell phone. The remote-accessible GUI also let the senior personnel can monitor several machines at the same time.

Of course, there are some disadvantages with the distribution as well. The network security should be the problem number one. In the reality, there is more work to do to make sure the network is secure.

8.2.2. GA for mathematic features extraction

In this master thesis, the data from the fan has the strong periodical character. So I decided to introduce the average value and the variance of certain time (2.4 hours) to be the features which were analyzed in the diagnosis step.

In the reality, what features are the optimum ones to be extracted should be calculated and compared by the genetic algorithm, not by the observation by the designers. There should be a list of known mathematical features for each kind of parameters. Then the GA involves checking out all these features. As the result, some certain features will be chosen.

This features extraction process is more academic and strict to the mathematics. An efficient and widely used system should always have at least a piece of academic theories in each design idea.

8.2.3. Lattice breath

In this master thesis, the space in the system boundaries are divided into one thousand lattices evenly. This idea can dramatically reduce the number of state so that the computer can calculate the states easier.

However, both the designers and the clients do not concern the states inside of the same status area. They all concern the states between two different statuses. For example, the states between “normal” and “error1” have more possibility with the status of “alarm”.

In order to stand out the concerning states, the space or detection matrix cannot be divided evenly. The lattice breath is such an idea that is to expand the size of the concerning lattices. The next figure shows an example of the lattice breath. The L1 is more sensible than L2 and L3. So L1 expands its size and become to new lattices. The L2 and L3 are less important, so they merge into one lattice.

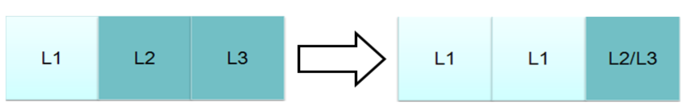


Figure 57 Lattice breath in 2-dimension

The technique is employed when the sensible lattices, the “alarm” lattices, area is too small. The technique can expand the prediction time, gives the personnel more time to maintenance the machine, Also improvement the sensibility of the CBM system.

The whole process of lattice breath is:

- Complete the detection matrix. The lattice breath needs a completed detection matrix. It means the lattice breath cannot work at the beginning of the CBM running which with a uncompleted detection matrix.

- Find out the “alarm” lattices. In this master thesis, the statuses of detection matrix are “normal”, “error1” and “error2”. The “alarm” status is a result of working with the traditional detection method which only with “normal” and “failure”.

  However, in lattice breath, all the “alarm” lattices must be found out. So the detection matrix will contain two more status: “alarm1” and “alarm2”. This is different with the matrix in this master thesis’s detection process.

- Set threshold. The threshold should be a system requirement of the sensibility. The sensibility is not the higher the better. When the system‘s sensibility is lower than the requirement, the lattice breath will work, until meet the threshold; When then system’s sensibility is higher than the requirement, the lattice breath will work reversely, until meet the threshold.

  The threshold choosing should be related with how often the system stays in the lattice. According to the information theory, the lattices with high possibility have low information content; the lattices with low possibility have high information content. This must be considered when the designers try to build up a function to judge the lattices’ importance and system’s sensibility.

- Lattice expansion. The lattice starts to expand the size. Each step of expansion need a threshold check which is guild line for how far this expansion goes.

8.2.4. Bayesians Formula for CBM

In this master thesis, I assume all different types of errors are independent. This assumption usually is not established for the complex system in reality. Usually, in a system one small failure may cause a big failure, and a big failure may cause the breakdown of the whole system. Also maybe there are some lattices have two different statuses. For example, one lattice has status “error1” and also “alarm2”.

In order to solve this problem, the Bayesians formula must be involved. By using the formula, the diagnosis approach can calculate the current state dynamically, rather than just mapping the system state into the detection matrix and find out the corresponding status.

The difficulty is how to find out the dynamic relations between statuses. The statistic should be another weapon to against this problem. The designers must have sufficient data from which all the relations can be deduced.

8.2.5. Pattern recognition

Pattern recognition is a sub-topic of machine learning. Pattern recognition aims to classify data (patterns) based either on a priori knowledge or on statistical information extracted from the patterns. The patterns to be classified are usually groups of measurements or observations, defining points in an appropriate multidimensional space. [27]

Pattern recognition is another very important method to solve the CBM problems. It is not suitable to this master thesis, if the pattern recognition works alone. But if this method works with the method adopted in this master thesis, the CBM system definitely will become much stronger.

The pattern recognition needs a different but similar detection matrix. Not like the detection matrix divided evenly or expand the lattices which with high information content introduces above. The lattices of the matrix for pattern recognition should have the same or similar system staying possibility. Only this kind of matrix can depict the history of the system state as detail as possible.

Generally speaking, this method also needs the supervised or unsupervised training. There should be some known patterns, then base on them the system starts to work and learn from experience. Yet in this master thesis, the pattern recognition does not need any training. All the patterns come from the original detection method. Each time when the system sends “alarm”, the pattern recognition algorithm starts to trace the history of the system. When the same type of “alarm” is sent many times, the genetic algorithm involves and calculates the most likely pattern for this type of “alarm”.

As a result, this type of failure can be predicted by both detection method, and the system will choose the result of the one sent earlier. These two methods works together also can increase the prediction time.

The pattern recognition approach is only suitable to the situations in which the behaviors before failure have comparable long time regular track. So there is no guarantee that this method fits all the cases, this is also the reason this method works as an assistant method.

Reference

[1] www.se.gtk.fi/index, Geologiska Forskningscentralen, nationellt organ för geologisk vetenskap i Finland, accessed 2007-09-10.

[2] www.ne.se, Nationalencyklopedin 2007, accessed 2007-09-10.

[3] Intelligent Fault Diagnosis and Prognosis for Engineering Systems, by George Vachtsevamps, 2006

[4] knowledge-based architecture for integrated condition based maintenance of engineering systems, by Abhinav Saxena, 2007.

[5] http://www.weibull.com/hotwire/issue76/hottopics76.htm, April 7th, 2009

[6] Holland, J. H. Adaptation in Natural and Artificial Systems, University of Michigan Press, 1975.

[7] http://en.wikipedia.org/wiki/Evolution#Reference-Gould-2002, Mar. 28th, 2009

[8] Adam Marczyk, Genetic Algorithms and Evolutionary Computation, 2004

[9] Aleksander M. Stensby, An Introduction of Genetic Algorithms and Evolutionary Computation, 2008

[10] Koza, John, Forest Bennett, David Andre and Martin Keane. Genetic Programming III: Darwinian Invention and Problem Solving. Morgan Kaufmann Publishers, 1999.

[11] Introduction to machine learning, Nils J. Nilsson, Dec. 1996

[12] http://en.wikipedia.org/wiki/System_requirements, April 9th, 2009

[13] http://en.wikipedia.org/wiki/Expected_value, April 9th, 2009

[14] http://en.wikipedia.org/wiki/Variance, April 9th, 2009

[15] Machine learning, Tom M. Mitchell, 1997

[16] Kohonen T. (1997) Self-organizing maps. Springer-Verlag, Berlin, Germany, 426 s.

[17] http://en.wikipedia.org/wiki/Machine_learning, April 11th, 2009

[18] Abhinav Saxena,(2007) knowledge-base architecture for integrated condition based maintenance of engineering system

[19] Holland, J. H. Adaptation in Natural and Artificial Systems, University of Michigan Press, 1975.

[20] Avers, Charlotte (1989), Process and Pattern in Evolution, Oxford University Press

[21] http://en.wikipedia.org/wiki/Evolution#Reference-Gould-2002, Mar. 28th, 2009

[22] Adam Marczyk, Genetic Algorithms and Evolutionary Computation, 2004

[23] Koza, John, Forest Bennett, David Andre and Martin Keane. Genetic Programming III: Darwinian Invention and Problem Solving. Morgan Kaufmann Publishers, 1999.

[24] http://en.wikipedia.org/wiki/Graphic_user_interface, May 1st, 2009.

[25] window manager Definition PC Magazine, Ziff Davis Publishing Holdings Inc.

[26] http://www.sawin.cn/doc/SD/GUI/TheEdge188.htm (in Chinese), May 13th, 2009.

[27] http://en.wikipedia.org/wiki/Pattern_recognition, May 10th, 2009.