The Study of Fault Diagnosis

Based on Particle Swarm Optimization Algorithm

Qinming Liu & Wenyuan Lv

College of Management, University of Shanghai for Science and Technology

Shanghai 200093, China

E-mail: lqm0531@163.com, wy_lu@sina.com

Abstract

Diesel is a very complex system. Subsystems and components of diesel will be failure, between failure symptoms and causes have many uncertain factors. In view of this situation, fault diagnosis method based on PSO algorithm is presented. The effectiveness of this method is proved with an example of diesel system.

Keywords: Particle swarm optimization algorithm, Fault diagnosis, Probability casual model

1. Introduction

Monitored and diagnosed the machinery equipment by the fault diagnosis technology can find machine malfunctions in time and prevent the equipment's worst accident from occurring, so that it can avoid casualties, environmental pollution and enormous economic losses. Applied the fault diagnosis technology can find the potential causes in the process of producing equipment, so that eliminate potential causes of accidents through transforming to the machinery equipment and the craftwork.

In the paper, PSO algorithm is used to diagnose and research the faults and make the group including the possible solution to evolve based on the best group's theory of evolution. At the same time, it finds the most optical solution meeting requirements through searching the best individual in the optical group by the overall solution in parallel. Because of its simple ideas, easy to implement, as well as demonstrated the search ability can be used in many areas. Moreover, due to its parallel search mechanisms and characteristics of the global search, PSO algorithm can be used to diagnose faults.

2. The model of the fault diagnosis based on probability and causality

Applied the PSO algorithm to solve the problem of the fault diagnosis, firstly describe the problem of the fault diagnosis using the expression of the mathematical model. An assumption or solution is usually constituted by more than a fault, through a competitive mechanism to achieve fault diagnosis. A problem of the fault diagnosis can be defined as a model of probability and Causality.

Calculation of the diagnostic and reasoning methods gets form the formalization of probability and causality related with the shallow knowledge of solving the diagnostic problem. In the Reggia's saving and covering theory, a simple diagnostic problem can be defined

$$P = (D, M, C, M^+) \tag{1}$$

In the equation, $D=(d_1, d_2, ..., d_n)$ is a limited and non-empty set of the fault. $M=(m_1, m_2, ..., m_n)$ is a limited and non-empty set of the symptom. The relationship $C \in D \times M$ denotes the causal connection between fault and symptom. $(d_i, m_j) \in C$ shows that the fault di can cause the symptom m_j. $M^- = M - M^+$ is a collection of the known symptoms. $M^- = M - M^+$ is a collection of the known and non-existent symptoms. When the fault collection $D_l \in D$ is regarded as a hypothetical solution of the question, means assuming that all faults of $d_i \in D_l$ will occur and all faults of $d_i \notin D_l$ will not happen. The assuming collection D_l is called as "coverage" of a given M +. In other words, the assuming set D_l is a potential explanation of the symptom set, that is to say the explanation D_l of M^+ has not any subset to cover the M^+.

Each $d_i \in D_l$ is related with a number $p_i \in (0,1)$, pi is a priori probability for d_i. The causal strength C_{ij} between faults and symptoms gets value from (0,1), which is the probability that causes the symptoms under the condition of the known faults.

Accordingly, in the Peng's model of probability and causality, there is three assumptions

(1) Faults d_i independent of each other;

(2) Strength of cause and effect can not be changed, no matter when d_i occurred, it always gets m_j by the same probability C_{ij};

(3) Symptoms caused by non-faults do not exist, that is to say, all the symptoms are caused by the faults;

According to this three hypothesis, the likelihood function of a assuming set $D_l \in D$ based on the given M^+

$$L(D_l, M^+) = \prod_{m_j \in M^+} \left| 1 - \prod_{d_i \in D_l} (1 - C_{ij}) \right| * \prod_{m_i \in M^-} \prod_{d_i \in D_l} (1 - C_{il}) * \prod_{d_i \in D_l} \frac{p_i}{1 - p_i}$$

(2)

In the above equation, the first part can be seen as the weight response of symptoms existing in the given M^+ caused by the fault Dl. The second part can be seen as weight of non-existent symptoms M^-. The third part is the weigh of a fault's priori probability among D_l.

Obviously, as a given M^+, any one of the diagnostic assumptions can solve the likelihood value based on the equation (1) and the size of likelihood value represents the D_l possibility of the occurrence under the condition of the known M^+, so is regarded as the value of fitness function of the PSO algorithm that is reasonable. Moreover, because the existent faults and non-existent faults are independence of each other, when the number of faults is large, the collection of faults will be larger, making the calculation of the posterior probability of the fault collection become a large workload and using mathematical analysis and test methods to resolve such problems is very complex, even some of the problems is not solved. In view of this, in the paper using the PSO algorithm solves the fault diagnostic problem of diesel engine based on computer.

3. The fault diagnosis based on PSO algorithm

3.1 The principle

PSO is that the Kennedy and Eberhart are inspired by the foraging behavior of bird populations a technology evolution developed in 1995. The swarm originates from the particle swarm, in line with 5 basic principles of the group intelligent that is put forward by the Millonas in the process of developing model of artificial life. However the particle is a compromise choice, because the members in the group will be described as non-quality, non-volume, as well as need to describe velocity and accelerate situation of the members. Due to the simple concept and easy to achieve of PSO algorithm, just during a few years, PSO algorithm will acquire great development and get some application in many areas.

At first PSO algorithm is used to graphically imitate the beautiful and unpredictable movements of the bird's group. Through observing animal social behavior, finding the social share to information in the group will help acquire superior in the evolution, and becomes the development basis of PSO algorithm. PSO algorithm is similar to other evolutionary algorithm, also based on the groups. According to the fitness to environment, individuals in the groups will be moved to the good region. PSO algorithms does not like other evolutionary algorithms that applied the evolution operator to the individual, this will regard each individual as a non- volume particle point in N-dimensional search space and fly by a certain speed in the search space.

PSO algorithm is based on the groups, and according to the environmental fitness, individual in groups will be moved to the good region. But it does not use the evolution operator to the individuals. Each individual is regarded as a non-volume particle in D-dimensional space, flies by a certain speed in the search space. The speed is dynamically adjusted in accordance with flight experience of itself and fellows. The algorithm evaluates the optimal result by using evolutionary fitness function of group, and each particle in the algorithm has a fitness value determined by the fitness function, two properties of position and speed that are used to show the location and moving speed of the current particles in the solving space, at last by the fitness function value corresponding to particle position coordinate determines the performance of particles.

In the D-dimensional search space assumed that m-particles form a particle group, the i particle's space location is

$$X_i = [x_{i1,}x_{i2},...,x_{iD}], (i = 1,2,...,m)$$

It is a potential solution of the optimization problem, and will be to the optimization objective function and can calculate the corresponding fitness value, measure the X_i merits according to fitness value; the best location that the i particle's experienced is known as the best history location of individual.

Define

$$P_i = [p_{i1,}p_{i2},...,p_{iD}], (i =,1,2,...,m)$$

At the same time, each particle also has itself flight speed.

Define

$$V_i = [v_{i1}, v_{i2}, ..., v_{iD}], (i =, 1, 2, ..., m)$$

The best position among all positions that particles experienced is regarded as the best history position of overall situation.

Define

$$P_g = [p_{g1}, p_{g2}, ..., p_{gD}]$$

The fitness value is regarded as the best history fitness value for overall situation, regarded as Fg. To each generation of particles, its d-dimension ($1 \leq d \leq D$) alternates based on the following equation

$$V_i(t+1) = w \times V_i(t) + c_1 \times rand() \times [p_{sb}{}^i(t) - P_i(t)] + c_2 \times rand() \times [P_{gb}(t) - P_i(t)] \qquad (3)$$

$$P_i(t+1) = P_i(t) + V_i(t+1) \qquad (4)$$

In the above equation, vi(t) is the speed vector of the i particle executed to the t time, pi(t) is the location of the i particle executed to the *t* time, c_1 and c_2 are the study factor and get the values from (0,2), *rand()* is the random function changed within [0,1], the inertia weight of w is a dynamic variable and controls the search capabilities of particles. When the value of w is larger, the particles have better global search capability; when the value of w is smaller, the particles have strong part search capabilities.

In the equation (3), the first part is the particles' previous speed and the second is the cognitive part, which shows particles own thinking, the third is the social part, which indicates to share information and cooperate with each other among particles.

The steps of the Standard PSO algorithm

a. Initializing a group of particles, including the random location and speed;

b. Evaluating the fittness value of each particle;

c. To each particle, its adaptive value will compare with the best position $P_{sb}(t)$ experienced, if the adaptive value is better, will be regarded as the best present position $P_{sb}(t)$.

d. To ach particle, its adaptive value will compare with the best position $P_s(t)$ overall situation experienced, if better, then reset the index of $P_{gb}(t)$.

e. According to the equation (3) and (4) change the speed and the location of particles;

f. Do not meet the conditions of the end (usually the adaptive value is good enough or achieve the predictive largest iterative count G_{max}), then return to the b.

3.2 The solving steps based on PSO

The PSO algorithm begins form a possible group that represents a possible potential solution set of the problems, and a species is constituted by a certain number of individuals encoded. Each individual is actually kind of physical breakdown. PSO algorithm through simulating evolution changes the parameters of the model. A fault diagnostic problem can be defined as a network of the probability and causality, including the D, M, C, p_i, c_{ij} and entered M^+. The solution of the problem is the likely occurred assumption D_l under the condition of the known symptoms set M^+, that is to say, it is a set of faults with the highest priori probability among all possible solutions.

The solving process of the PSO algorithm based on the probability and causality model

(1) Generated initial population

In the paper selecting four typical faults, each individual location is defined as the length of 4-dimensional and each dimension is constituted by 0 or 1. Randomly generated four individuals constitute a group, which the position of the *i* individual is d_{i1}, d_{i2}, ..., d_{i4} and d_{ij}=1 represents the existent faults, d_{ij}=0 shows the non-existent faults. In each of the individual the faults corresponded by the value 1 constitute the diagnostic assumption D_l and each individual shows a possible solution of the diagnostic problem.

(2) Initialization

According to the size M of particle swarm generates inertia weight ω, accelerating factor c1 and c2, the maximum allowable number of iterative G, the initial location and speed of the particles, and so on.

(3) The objective function

The likelihood function $L(D_l, M^+)$ gotten by the probability and causality model regards as the size of individual fitness. If the $L(D_l, M^+)$ is bigger, the posterior probability of the fault collection is greater, under the condition of the known symptoms collection, the posterior probability and fault collection will be greater associated with it. So choose $L(D_l, M^+)$ as the individual's fitness, which provides the basis for the selection and evolution of species.

(4) Comparison with the current optimal value:

To each particle, its current fitness value compares with the best history value of individual, if

$L(D_l, M^+)$ is larger than *pBest*, then *pBest* equals $L(D_l, M^+)$; Group current fitness value of all particles compare with the best history value of overall situation, if $L(D_l, M^+)$ is larger than gBest, then gBest equals $L(D_l, M^+)$.

(5) Calculating the new speed and location of particles based on the equations (3) and (4).

(6) Stopped criteria: error of adaptive value reaches the limit of adaptation error or iteration number exceeds the maximum allowed number, stop to search and output results. Otherwise, return the equation (3) to continue.

4. Application Example

PSO algorithm has the ability to achieve global optimization, according to the probability and causality model of solving fault, at first we must first determine the symptoms sets, the faults sets and the strength of cause and effect of the fault diagnostic problem, in the table 1, the faults sets is regarded as the output sample, and the symptoms sets as the input sample. Each output has respectively two states, 0 on behalf of non-fault state and 1 on behalf of fault state.

In the paper, using the diesel system as the diagnostic object, due to it is the entire system, so the whole system is divided into 3 symptoms of fault, define the vibration (F_1), pressure (F_2) and temperature (F_3). Main fault is divided into 4 categories, namely, crank-link fault (S_1), valve fault (S_2), fuel system fault (S_3) and other system faults (S_4). After a long experience to know a priori probability: P = [0.05 0.30 0.50 0.05]. The so-called prior probability is that after a long-term accumulation of experience determines the frequency of a variety of faults. After dealing with fuzzy and experts' experience to identify and carry out, the relationship between the symptoms and causes of fault is shown in the table 1.

At the same time, in the paper, encoded by *VC++*, and PSO parameters

In the equation (3), the inertia weight of w is an important parameter bearing on search capabilities of the particle swarm optimization algorithm, the greater w has the ability of global search, and the smaller w has the ability of local search, As a result, considering the inertia weight will be linearly reduced from the larger value to a smaller value, so that the beginning of the algorithm has a strong global search capability, and near the end of the algorithm has the better local search capabilities, the specific calculating equation is

$$\omega=1.2-(T/T_{max})*0.8 \tag{5}$$

In the equation, T is the current iterative number. T_{max} is the largest stopped iterative number.

In the example, based on the experience the weight of inertia takes the initial value of 1.2, then according to the equation (5) linearly reduced to 0.4. In the equation (3), the accelerating coefficients c_1 and c_2 take 2.0 in the light of experience.

In addition, through a series of tests when the species group iterates to 100, the gained solution has to be the most optimal solution, the maximum allowable number of iterations of the species group takes $G=100$, the size of the species group takes $M=10$.

After calculating 100 times, the result is shown in the table 2.

Through determining by experienced machinery staff, the fault corresponded by the symptoms F_1 and F_3 is the fuel system fault (S_3). This shows that the PSO algorithm is absolutely correct and solving the fault diagnostic problem using the PSO algorithm is feasible.

5. Conclusion

Diagnosis can be defined as a network of the probability and causality, the network is made up of a sign-and-fault sets and strength of the probability and causality. In the paper, combined with the probability and causality model and the PSO algorithm diagnose the fault to diesel engine system. Experiments have proved that using the PSO algorithm to diagnose fault can be shortened the average time to diagnosis, improve the efficiency of the diagnosis and reduce the amount of computing, in the actual training, these are very significant. The system can effectively identify system fault, and the approach is feasible and practical. Using the PSO algorithm acquire knowledge in the complex fault diagnostic problem under the condition of lack experience and inadequate knowledge, there will be a wide range of applications.

References

BERHARDR and KENNEDYJ. (1995). 'A new optimizer using particle swarm theory'. *Proceeding of sixth international symposium on micro machine and human science*', NJ, USA: IEEE Service Center,1995,39-43.

Chen, Changzheng and Liu, Qiang. (2001). 'Probability of cause and effect in the steam network fault diagnosis'. *Chinese Journal of Electrical Engineering*, 2001, 21 (3): 782-811.

ENNEDYJ and EBERHARDR. (1995). 'Particle swarm optimization'. *Proceeding of IEEE Int'l Conference on Neural*

Networks, NJ, USA: IEEE Service Center, 1995, 1982-1948.

Peng, Y and Reggia J. (1987). 'A probabilistic causal model for diagnostic problem solving-part II diagnostic strategy'. *IEEE Trans. SMC*, 1987, 17(3): 395-406.

Qin, Xiaojian. (2006). 'Particle swarm optimization approach for semiconductor furnace bath scheduling problem'. Shanghai: *University of Shanghai science and technology* 2006:28 (5): 500.

Wang,Jiangping. (2001). *Mechanical fault diagnosis technology and application*. Xi'an: Northwest Industry University Press, 2001:8.

Wu, Jinpei and Xiao Jianhua. (1997). *Intelligent fault diagnosis and expert system*. Beijing: Science Press, 1997: 4.

Wang, Daoping and Zhang, Yizhong. *Theory and method of fault diagnosis system*. Beijing: Metallurgical Industry Press, 200: 3-8.

Zhong, Binglin and Gu, Yanyan. (1995). 'Genetic algorithm based on the probability of diagnosis of cause and effect model'. *Journal of Mechanical Engineering*, 1995, 31(6).

Table 1.The relationship between the symptoms and causes of fault

Symptoms /Faults	F1	F2	F3
S1	0.4	0.9	0.3
S2	0.6	0.33	0.5
S3	0.8	0.45	0.98
S4	0.3	0.88	0.54

Table 2. The diagnostic results

Symptoms	Faults	Running Time(s)
F1, F3	S3	3