A Fault Prediction Approach Based on Bayesian Network for System

Tianyu Si[a], Weiwei Hu*[b], Yuna Liu[c], and Jiamin Liu[d]

[a,b,c,d]School of Reliability and System Engineering, Beihang University, Beijing, China

Abstract: Due to the development of science and technology, the engineered products are getting more and more complexity, it is very important to assess their state and predict their fault and lifetime. Bayesian network (BN) is one of the best mathematical model that can handle uncertainty knowledge expression and reference especially suitable for complex systems. It has been successfully applied in fault diagnosis, data mining, reliability analysis, security analysis, and information fusion. Besides, BN also has great potential in fault prediction. In this paper, we will discuss a Bayesian network-based fault prediction approach combined with mean time to first failure (MTTF). Firstly, the system is modeled by BN, and then the failure prediction of the system is completed by using the probability of nodes and MTTF to calculate through BN. Furthermore, the lack of interconnection between different level in a system lead us to consider the system as three parts: component level, module level, system level. An active vehicle suspension (AVS) system will be used as an example to explain this method. This fault prediction approach will lay an important foundation for the development of fault prediction technology for complex system.

Keywords: Bayesian network, PHM, system, MTTF

1. INTRODUCTION

Due to the development of science and technology, the engineered products are getting more and more complexity, integration and intelligence have become the mainstream trend. From the aspects of efficiency and resource consumption, traditional methods of regular maintenance and after-service maintenance have gradually been eliminated by the times. In particular, ordinary maintenance and repair methods in complex systems are increasingly difficult to meet the ever-changing needs. In order to comprehensively improve the reliability of complex equipment, ensure its work efficiency and use safety, and greatly reduce resource consumption, the focus of research has shifted from traditional real-time supervision and management to prior failure prediction [1].

As a result, Prognostic and Health Management (PHM) was formally proposed. Its core strategy is not to focus on finding faults and dealing with faults, but instead to focus on when and where the faults are predicted to occur. Fault prediction is the core of PHM. Its implementation idea is to make use of as much as possible evidence and uncertain evidence information to predict the system's fault trend, degradation rules, and the possibility of fault occurrence within a period when the equipment works normally [2,3,4]. Bayesian network (BN) is a highly promising approach that can fuse various information from the perspective of probability [9,10].

In recent decades, Bayesian network has received more and more attention as a powerful model for knowledge expression and inference of uncertainty [5,11]. In the mid-1980s, research on uncertainty in the field of artificial intelligence gave birth to Bayesian network. The concept of the Bayesian network and its calculation method were summarized by Prof. Judea Pearl of the University of California in 1998. BN has been applied in various fields, from the initial expert system to success in fault diagnosis, data mining, reliability analysis, security analysis, and information fusion. Nowadays, BN has become a mature method for solving uncertain problems and precise logical inference. It has been successfully applied in many fields such as industry, agricultural, military, medicine, economy, fault diagnosis, etc. Besides, BN also has great potential in fault prediction [6,7].

*hww@buaa.edu.cn

The Bayesian network is a directed acyclic graph that can be used to describe the interrelationships between random variables then to perform probabilistic inference. There is a great difference between BN and other logical inference models. It visualizes the multi-dimensioned knowledge diagram and contains conditional correlations and causality between variables. Its most prominent advantage is that it can deal with uncertainty issues very well. To compare with fault tree (FT) approach, modeling with BN is more powerful, because BN can quantitative the uncertain relationship between each node through conditional probabilities, instead of deterministic "AND" and "OR". BN is also a highly inclusive network that can accommodate and diagnose various information related to decisions and handle them in a unified mathematical modeling approach. It can accurately express multi-source information and reasonably fuse information with related relationships [8].

In this paper, we will discuss a Bayesian network-based fault prediction approach combined with mean time to first failure (MTTF). Firstly, the system is modeled by BN, and then the failure prediction of the system is completed by using the probability of nodes and MTTF to calculate through BN. In next section presents some basic principles of BN. Our fault prediction approach will be discussed in Section 3. We use BN to model and layer the system, then locating the weak links and getting future state of the system in both static and dynamic. MTTF theory will be used to calculate the time of failure. We illustrate the proposed method with a case study in Section 4 and conclude the paper in Section 5.

2. BAYESIAN NETWORK

2.1. The Basic Principle of Bayesian Network

The Bayesian network is a graphical model that describes the relationship between data variables and probabilistic inference and can reflect the dependencies between variables. It provides a structural framework that can clearly express causation, complete inference on uncertainty issues in a clearer, more logical, and more understandable way.

BN is a directed acyclic graph (DAG). In the structure of BN, if there is a directed edge from node X to Y, X is called the parent node of Y and Y is the child node of X. The root node is the collective name of a node without a parent node, which in, and the leaf node is a collective name of a node without a child node. As a DAG, BN has four features:

1. There is a set of variables $V=\{X_i\}$, where $i = 1, 2, ..., n$, X represents a variable node, and set E contains a directed edge connecting the corresponding nodes.
2. The value of each variable can be either discrete or continuous.
3. A $DAG\ G=<V, E>$ is formed by the directed edges between the nodes and the nodes corresponding to the variables, where V is a set of nodes contained in the network and corresponds to the random variables in the domain, E is a directed edge set describing causal connections between variable nodes.
4. For every node X_i and its set of parent nodes $Pa(X_i)$, there is a conditional probability distribution table $P(X_i\ |Pa(X_i))$. It means the state of $Pa(Xi)$ can influence or determine the state of X_i. And satisfied:

$$\mathrm{P}(X_1, X_2, \cdots, X_N) = \prod_{i=1}^{n} P\big(X_i \big| Pa(X_i)\big) \tag{1}$$

It can be seen from the above definition, BN consists of a DAG and several conditional probability tables, where the DAG depicts qualitatively the dependence between variables, and conditional probability table quantitatively describes the dependence state between a variable node and its parent node. In addition, the realization of BN also needs to make a series of conditional independence assumptions, which are contained in the DAG of BN.

2.2. Bayesian Network Modeling

In the definition of BN, the nodes in the BN correspond to a random variable. The value of the node can be either discrete or continuous. Discrete value can be used in static and continuous value can be used in dynamic.

Discrete variables are generally divided into three types: Boolean variables, ordered variables, and integral variables.

1. Boolean variables: the value of the variable is divided into true and false or 0 and 1. In PHM, it could be called normal and faulty, we can use these two states to determine whether the system is faulty.
2. Ordered variables: the value of the variable has a certain order relationship. Two states are usually not sufficient to describe a system with a degenerate process. We can separate variables into several states according to their degeneration.
3. Integral variables: the value of the variable corresponds to a certain range.

Continuous variables mean that the value of a variable changes continuously according to a certain rule. In the prediction approach of this paper, the failure rate of each node in the system is a function related to time and will change continuously with time according to the distribution function.

In the structure of BN, one node can have multiple parent nodes and one node can also have multiple child nodes. A node and its parent node are called a family. By assigning a conditional probability table to each node and its family, BN quantitatively describes the strengths and weaknesses of the dependencies between variables. The size of the conditional probability table has a direct relationship with the number of parent nodes and the number of states. For a network where all nodes have two values, if one node has m parent nodes, there will be 4^n joint distribution. But in BN, after integrating them into family concept without considering probability normalization, there is just 2^{n+1} conditional probability values, significantly reduced data storage space and computing time.

Determining the conditional probability table is a very important part in constructing a BN. The accuracy and completeness of the conditional probability table directly determine the inference results. The conditional probabilities of each node in the traditional BN are determined by the expert system, which greatly depends on the expert experience and the knowledge of the system, and its accuracy cannot be judged. Today's research has developed many methods for parameter learning of BN. The determination of the conditional probability table based on the sample data does not require a deep understanding of the system and also improves its accuracy. There are two basic conditions for the parameter learning of BN: a complete data set and an incomplete data set. Maximum likelihood estimation and Bayesian estimation can be used for a complete data set, expectation-maximization algorithm (EM) can be used for an incomplete data set.

1. Maximum likelihood estimation

The maximum likelihood estimation method is based entirely on sample data and does not require prior probabilities. It regards the parameter to be evaluated as an unknown but fixed quantity without considering the influence of prior knowledge. Calculated as follows:

$$\mathrm{L} = \frac{1}{N}\sum_{i=1}^{n}\sum_{i=1}^{s}\log\big(P(X_i|pa(X_i), D_i)\big) \tag{2}$$

Which $pa(X_i)$ represents dependent variable of X_i, D_i represents the first observation, N represents the total number of observations.

2. Bayesian estimation

Bayesian estimation is more suitable for learning small sample data. Under a given topology S, find all possible values of an assumed constant unknown parameter θ, and then use the Bayesian formula to calculate the posterior probability of the parameter using the prior knowledge in the topology S and the training sample set D. The maximum value is the final value of the parameter:

$$\mathrm{P}(\theta|D,S) = \frac{P(D|\theta,S)P(\theta|S)}{P(D|S)} \tag{3}$$

Which $P(\theta|S)$ is the prior probability of parameter θ under topological structure S, $P(\theta|D,S)$ is posterior probability. There are two methods for estimating the parameter values: one is the maximum posterior distribution method, uses the value of the parameter θ at the maximum of the posterior distribution mentioned before as the estimated value; the other is named conditional expectation estimation, which is the mean of the posterior probability values of all the parameters θ, then take this as the estimated value.

3. EM algorithm

EM algorithm is an iterative learning algorithm. When the a priori data obtained is incomplete use this method to calculate the maximum likelihood of the network. The main idea of the EM algorithm is to arbitrarily give an initial value of the pre-estimation parameter, then compare and analyze other observations and this initial value to figure out the conditional mean value of other unobserved nodes. Use the estimated observations to supplement the incomplete data with the complete data. Substitute this complete data set into the model's estimation formula to maximize the formula and figure out the most probable value for this parameter. Repeat the above steps until the given cycle ends or the parameters converge.

3. BAYESIAN NETWORK-BASED FAULT PREDICTION

Since BN has a very strong ability to deal with probability problems, the fault prediction methods proposed in this paper are based on the failure rate. By constructing a complete BN, it is possible to infer the failure rate of any other node based on the failure rate of its parent node. However, we should layer the system first. BN is also a powerful layering and fusion tool, it can handle different levels and different kinds of data in a mathematical way. In a system, different levels have different characteristics. We consider the system as three parts: component level, module level, system level, as shown in **Figure 1**. For a product, many components make up many modules, modules make up the whole product. We always care about the lifetime of the entire product, but we can only get the data of the components. Most of the monitoring work can only be done at the module level. Evidence information comes from different levels, and the predictions we want to obtain are at other levels. BN can perfectly integrate the relationship between them and do the calculation, as shown in Figure 2.

The structure and conditional probability table of BN is shown in **Figure 2**, an 8-node system is used as an example. *T* represents the node is in normal state, and *F* represents the node is in a fault state. For example, when R is faulty, but B is normal, we can see from the conditional probability table that the probability of D failure is 0.37, which is influenced by R and B (its parent nodes). If we set the failure rate of A is 0.2 and S is 0.15, the failure rate of X will be 0.2681 and D will be 0.1974, according to **Formula 1.** So, when we set the failure rate of root nodes or any node, the state of every node will be estimate. There are many kinds of inference engine, we use connection tree algorithm. This is an accurate inference algorithm that is widely used in Bayesian networks. It can handle multi-access networks.

Figure 1: Schematic of Layer

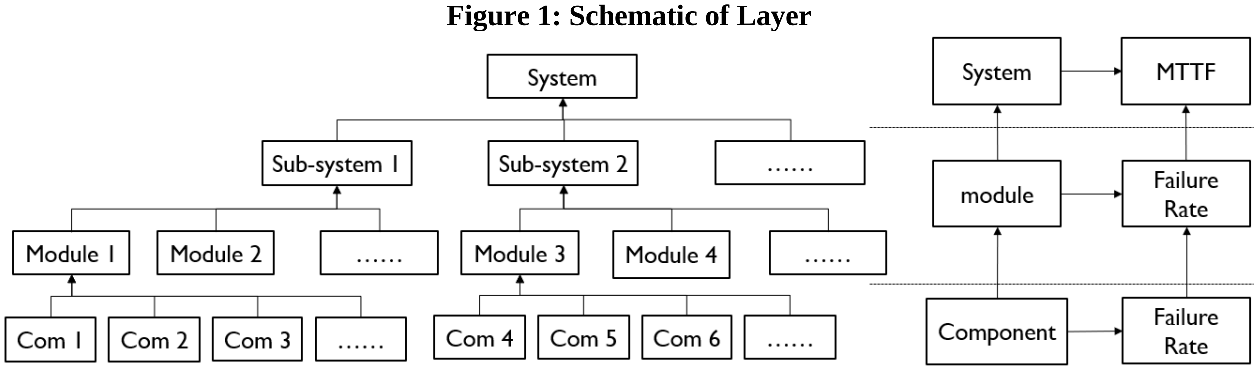
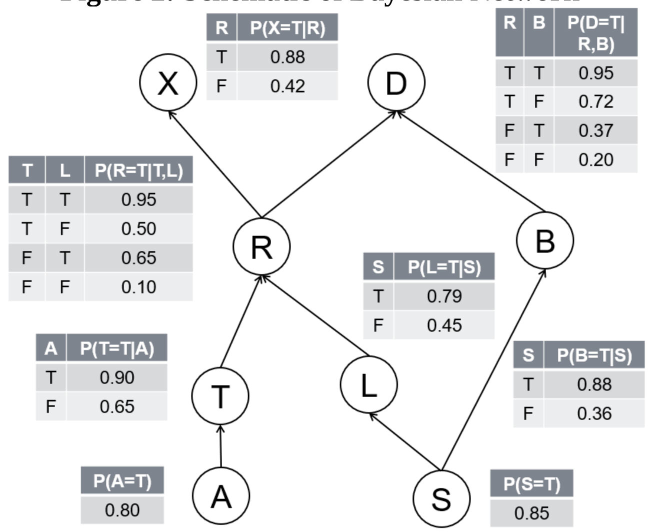


Figure 2: Schematic of Bayesian Network



In the static state, BN can be used in combination with the evidence node to analyze the system faults, locate the weak links according to the fault threshold, and maintain the equipment or improve the design. For the example above, if we set the fault threshold is 0.25, then we know X should be repaired or replaced, and D could be considered normal. At the design stage, we can test the robustness of the system by simulating the state of any multi-node in the system. When weaknesses are discovered, tolerance designs can be added to improve product reliability. In operation stage, BN can also analyze the current state of the system, any evidence observed can be entered into the Bayesian network to update information to improve the accuracy of the forecast. This method is not only suitable for fault prediction, but also fault diagnosis. BN inference process can be done both directions, if we observed the system or any sub-system is broken, we can use BN to infer where is the highest probability of failure.

Under dynamic conditions, the failure rate of each node at different times can be obtained according to the failure rate of the input node. The time-related failure rate function of system can be obtained by getting the change of the failure rate in several time points. When we get the failure rate function, we can calculate MTTF by **Formula 4-6**.

$$\mathrm{R}(t) = e^{-\int_0^t \lambda(t)dt} \tag{4}$$

$$\mathrm{MTTF} = \int_0^{\infty} R(t)\, dt \tag{5}$$

Therefore,

$$\mathrm{MTTF} = \int_0^{\infty} e^{-\int_0^t \lambda(\theta)\, d\theta}\, dt \tag{6}$$

Where *R* is the reliability, λ is the failure rate. In this way, we can predict the life of the system.

To summarize the above, we propose a Bayesian network-based fault prediction approach framework as shown in **Figure 3**, detailed process shown in **Figure 4**.

4. CASE STUDY

In this section, we demonstrate our methodology on a hypothetical mechatronic system: an active vehicle suspension (AVS) [13,14]. The BN structure of AVS system is shown in **Figure 5**, it's a 13-node network. Suppose that the reliability structure of the passive device reliability becomes uncertain. Therefore, we should describe the relationship between node 2 and its parent nodes, 6 and 7, through conditional probabilities. The condition probability table is shown in **Table 1**, the conditional probabilities are set by expert experience.

Figure 3: Framework of Bayesian Network-based Fault Prediction Approach

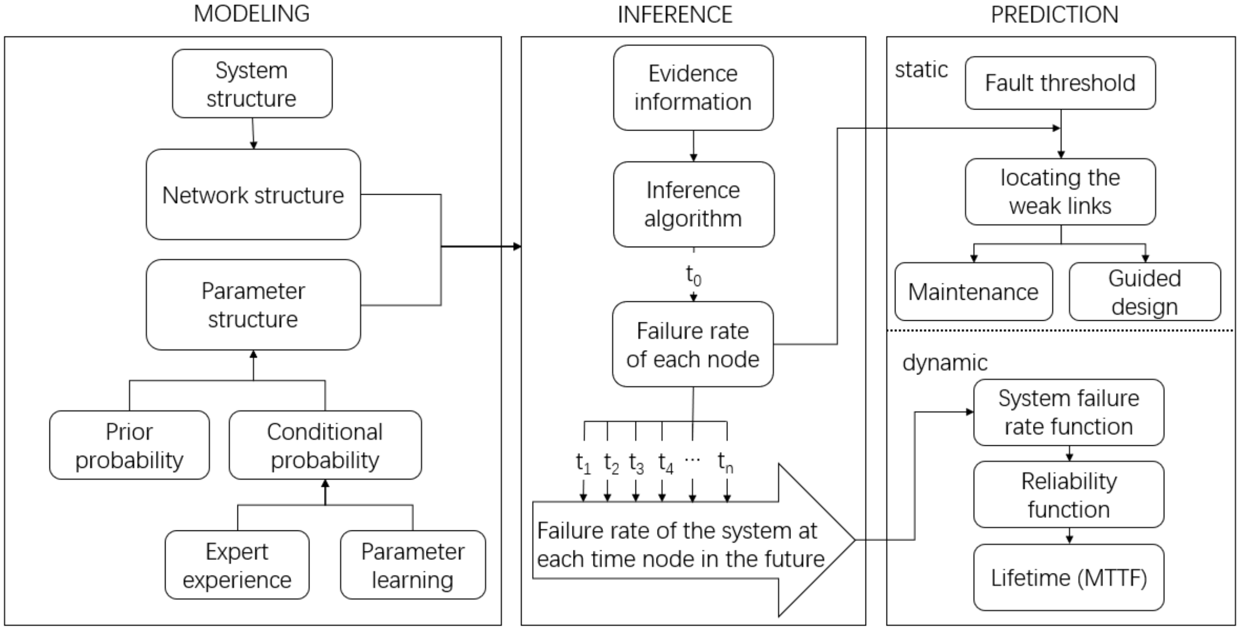
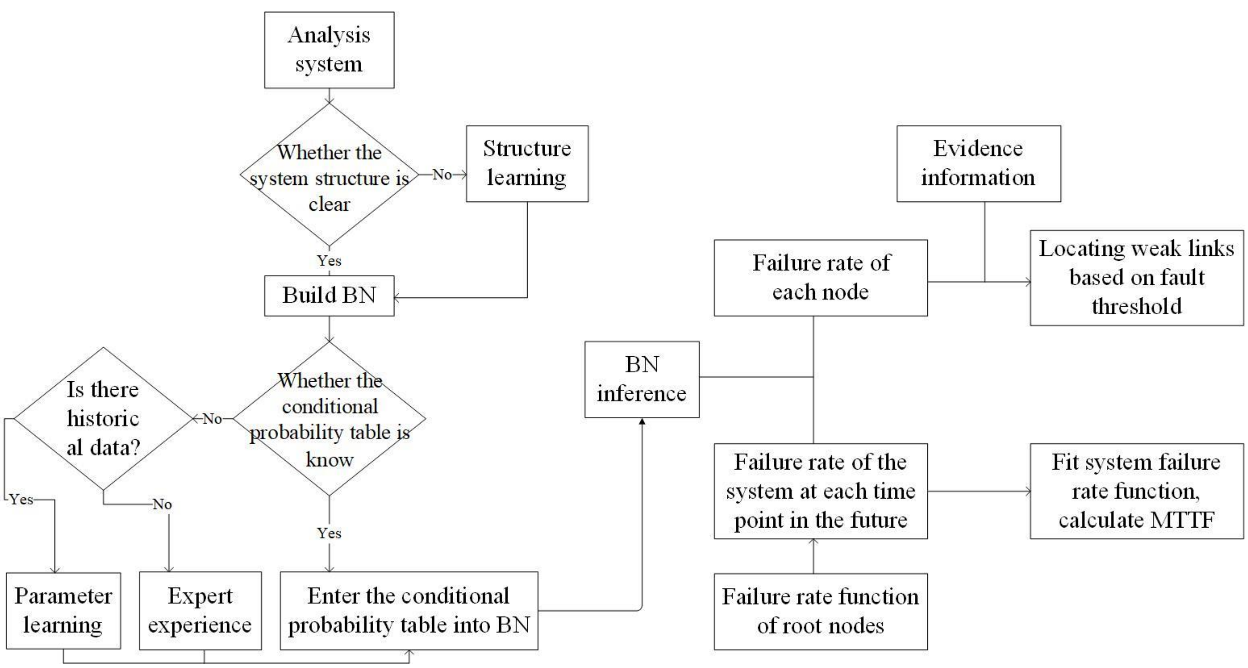


Figure 4: Detailed Process of Bayesian Network-based Fault Prediction Approach



And failure rate of the nodes is given in **Table 2**. *T* represents the node is in normal state, and *F* represents the node is in a fault state. Because AVS is used in car, so we use kilometers to represent lifetime of this system. All calculations are implemented in MATLAB.

After building BN, we can do the analysis. At design stage, when the vehicle is not in use, we can use BN to figure out the system's weak links. We select the time when the vehicle has been running for 10^5km to do static analysis. At that time, the failure rate of AVS is 0.0252. Suppose that AVS fails at 10^5, node *X1* will be *F*, the failure rate of other nodes is shown in **Figure 6**.

Figure 5: The Corresponding BN Model

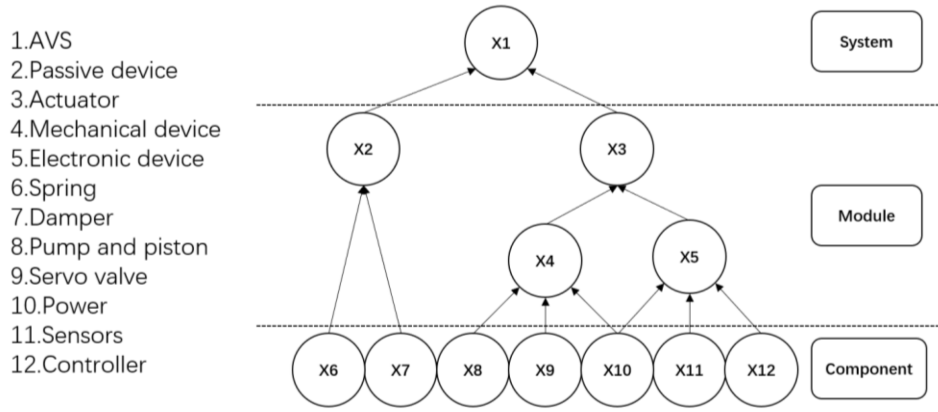


Table 1: Conditional Probability Table

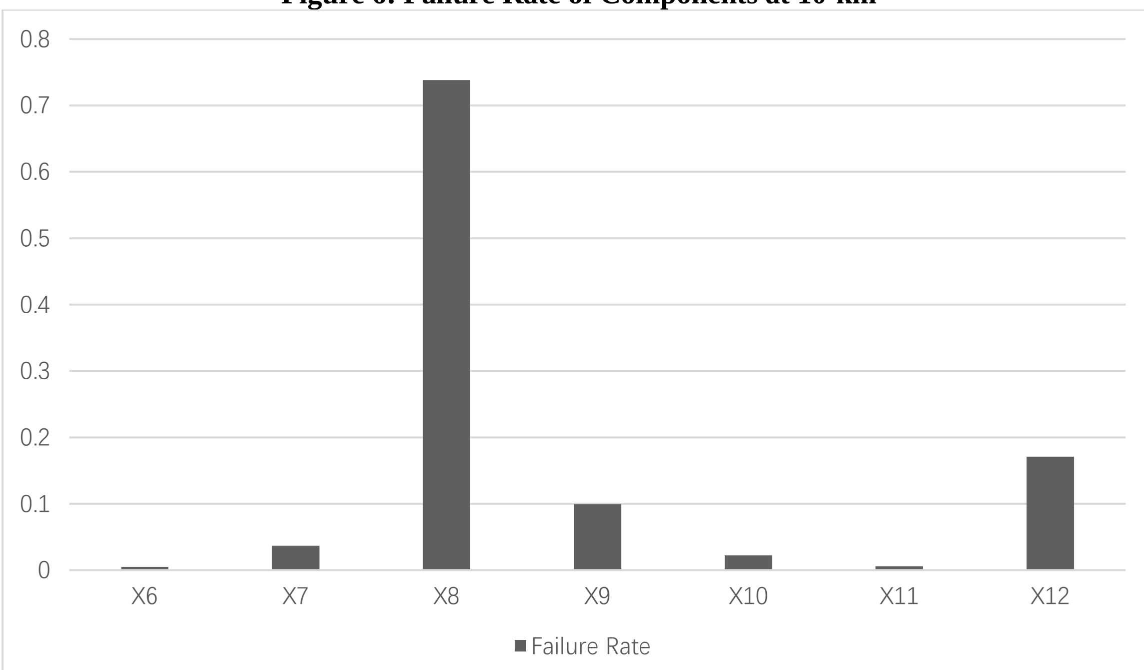
X2				
X6	X7	P(X2=T|X6,X7)	P(X2=F|X6,X7)	
T	T	0.84	0.16	
F	T	0.23	0.77	
T	F	0.11	0.89	
F	F	0.12	0.88	
X4				
X8	X9	X10	P(X4=T|X3,X4,X5)	P(X4=F|X3,X4,X5)
T	T	T	1	0
F	T	T	0	1
T	F	T	0	1
T	T	F	0	1
F	F	T	0	1
F	T	F	0	1
T	F	F	0	1
F	F	F	0	1
X5				
X10	X11	X12	P(X5=T|X10,X11,X12)	P(X5=F|X10,X11,X12)
T	T	T	1	0
F	T	T	0	1
T	F	T	0	1
T	T	F	0	1
F	F	T	0	1
F	T	F	0	1
T	F	F	0	1
F	F	F	0	1
X3				
X4	X5	P(X3=T|X4,X5)	P(X3=F|X4,X5)	
T	T	1	0	
F	T	0	1	
T	F	0	1	
F	F	0	1	
X1				
X2	X3	P(X1=T|X2,X3)	P(X1=F|X2,X3)	
T	T	1	0	
F	T	1	0	

T	F	1	0
F	F	0	1

Table 2: Failure Rate of Components

Item name	Failure Rate (10^{-7}km)	Failure Rate when 10^5km (km^{-1})
X6. Spring	0.105	0.00105
X7. Damper	0.68	0.0068
X8. Pump and piston	11.23	0.1123
X9. Servo value	1.51	0.0151
X10.Power	0.34	0.0034
X11. Sensors	0.091	0.00091
X12. Controller	2.6	0.026

Figure 6: Failure Rate of Components at 10^5km



From **Figure 6** we can easily know *X8* is the most likely location of failure, and X12 is also a bit. So, when the vehicle has been running for 10^5km, we should pay more attention to the state of X8, X12 and overhaul it.

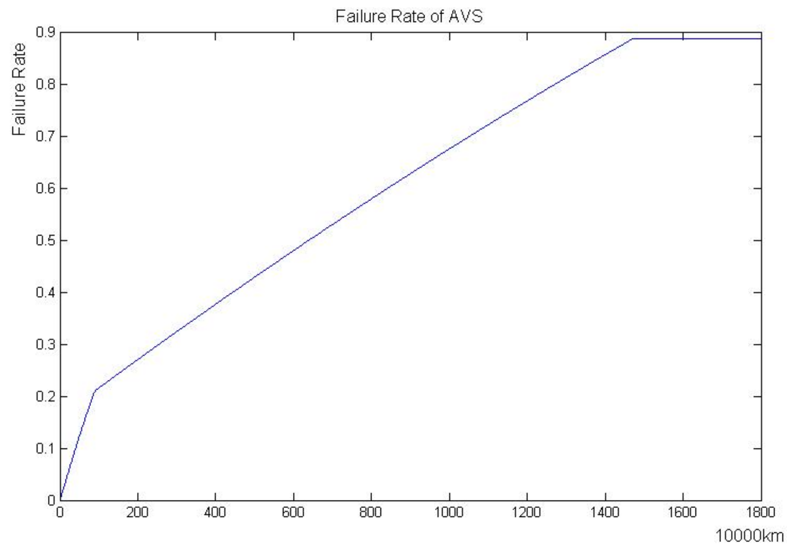
Under dynamic condition, we take a value every 10,000 km and take a total of 1800 data sets. Failure rate of AVS as a function of time is shown in **Figure 7**. The results show that this is a piecewise function and the fitting results are as follows, which $x=10^4km$. As we can see from the result, the failure rate function of AVS is a piecewise function with 3 segments. Its mathematical expression is shown in **Formula 7-9**. By substituting these three failure rate functions into **Formula 4**, the reliability of the AVS system can be calculated as a function of mileage. Then use **Formula 5** to calculate MTTF of AVS, which is 7.78434×10^6 km. The result means the AVS system will not be able to continue working after 7784340km.

$$\lambda_1(x) = 0.002364x \tag{7}$$
$$\lambda_2(x) = 0.0004913x + 0.18 \tag{8}$$
$$\lambda_3(x) = 0.88 \tag{9}$$

5. CONCLUSION

The fault prediction approach based on Bayesian network proposed in this paper is feasible. By BN modeling, learning, analysis, and inference of the system, the failure states and future failure rate information can be obtained. Then the subsequent failure prediction is completed. Compared with the fault tree, BN can handle more data and more complex logical relationships. Its conditional probability can

Figure 7: Failure Rate Function of AVS



express both the determined logical relationships and the uncertain logical relationships. It is more flexible and practical. Learning and bidirectional inference ability are also one of the advantages of BN, and BN can reflect the connection relationship between units of the same level and different level more accurately. The state of the underlying units determines the state of the upper unit, and its prediction results take full account of changes in the underlying unit and the internal conditions of the system. However, there are many deficiencies in this study. Future work will start from the following points:

1. The case in this paper approximates a linear system, but most of the actual systems are nonlinear. The application of this approach in nonlinear systems still requires practical testing.
2. We consider three parts of the system, but for the lifetime of products, environment condition is very important. So, we will add the environment information as the lowest level nodes to the network, take full account of the impact of environmental changes on the products.
3. Conditional probability is not always changeless in a complex system. BN can also realize conditional probability changes over time.
4. BN as a powerful fusion tool should also incorporate more information from different levels.

However, the potential of BN in PHM is far from being noticed, we should continue to study it.

References

[1] Sheppard, J. W., and M. A. Kaufman. "*A Bayesian approach to diagnosis and prognosis using built-in test.*" IEEE Transactions on Instrumentation & Measurement 54.3(2005):1003-1018.
[2] Hess, A., and L. Fila. "*The Joint Strike Fighter (JSF) PHM concept: Potential impact on aging aircraft problems*." Aerospace Conference Proceedings IEEE, 2003:6-3021-6-3026 vol.6.
[3] Qiu, Hai, et al. "*Robust performance degradation assessment methods for enhanced rolling element bearing prognostics*." Advanced Engineering Informatics 17.3(2003):127-140.

[4] Zhang, Xiaodong, et al. "*An integrated approach to bearing fault diagnostics and prognostics*." American Control Conference, 2005. Proceedings of the IEEE, 2005:2750-2755 vol. 4.
[5] Suisse, Martigny Valais. "*An Introduction to Bayesian Network Theory and Usage.*" Idiap (2000).
[6] Langseth, Helge, and F. V. Jensen. "*Bayesian Networks in Reliability.*" Reliability Engineering & System Safety 92.1(2007):92-108.
[7] Boudali, H., and J. B. Dugan. "*A discrete-time Bayesian network reliability modeling and analysis framework.*" IEEE Transactions on Reliability87.1(2005):86-97.
[8] Bobbio, A., et al. "*Improving the analysis of dependable systems by mapping fault trees into Bayesian networks.*" Reliability Engineering & System Safety 71.3(2001):249-260.
[9] Mann, Jordan. "*Fault detection and diagnosis using a probabilistic modeling approach.*" Dissertations & Theses - Gradworks (2011).
[10] Ma, Dezhong, et al. "*Constructing Bayesian Network by Integrating FMEA with FTA.*" Fourth International Conference on Instrumentation and Measurement, Computer, Communication and Control IEEE, 2014:696-700.
[11] Eddine, Abbass Zein, et al. "*Improving Fault Isolation in DC/DC Converters Based with Bayesian Belief Networks.*" IFAC-PapersOnLine49.5(2016):303-308.
[12] Al Tamimi, Abdallah Moh'd AR. Sh. M. "*IMPROVED PROBABILISTIC LIFE ESTIMATION IN ENGINEERING STRUCTURES: MODELING MULTI-SITE FATIGUE CRACKING.*" Dissertations & Theses - Gradworks(2014).
[13] Yontay P, Pan R. "*A computational Bayesian approach to dependency assessment in system reliability*." Reliability Engineering & System Safety, 2016, 152:104-114.
[14] Zhong X, Ichchou M, Saidi A. "*Reliability assessment of complex mechatronic systems using a modified nonparametric belief propagation algorithm.*" Reliability Engineering & System Safety, 2010, 95(11):1174-1185.