Analysis on Dimensional Reduction Method for Reliability Assessment with Few Data by Means of Multiple Factors

Tao Zhang* Member

Masatoshi Nakamura** Member

Lack of data is an obstacle to realize accurate assessment of reliability. Dimensional reduction method was successfully used for reliability assessment with few data by means of multiple factors. In order to generalize the dimensional reduction method for reliability assessment and expand its application area widely, this paper presents a general definition of the concept, algorithm and application area of dimensional reduction method for reliability assessment and analyzes the properties of this method, such as selection of factors, uniqueness of dimensional reduction results, determination of dimensional reduction sequence, etc. Some simulations have been made based on the actual data to verify the effectiveness of dimensional reduction method for reliability assessment.

Keywords: dimensional reduction method, reliability assessment, few data, multiple factors

1. Introduction

In reliability engineering, reliability assessment of devices and systems is essential for reliability analysis, design and optimization. Since an accurate assessment of reliability is definitely required in industry, how to realize it is an important problem all along. Generally, event probability is regarded as one of the effective approaches to assess reliability⁽¹⁾. A vast number of algorithms based on the probability theory, statistics, etc., were therefore developed. In modern industry, reliabilities of devices and systems have been improved remarkably with the advancement of technology. Particularly, computer technology, large-scale integral circuit, etc., are widely used in devices and systems. The occurrence of faults becomes fewer and fewer. Engineers always complain that they can not acquire sufficient valid data for reliability assessment. The precision of reliability assessment can not be guaranteed for evaluating the features of devices and systems without sufficient data. Therefore, an effective approach is required by engineers to assess reliability accurately with only few data available. With respect to this issue, dimensional reduction method was proposed to undertake the reliability assessment with only few data by means of multiple factors.

In case of too-small data sets, several techniques can be possibly employed for reliability assessment, such as Bayesian analysis, bootstrap, neural network, etc. The Bayesian analysis is a methodology for formal incorporation of prior information with the current data⁽¹⁾. One of the fundamental problems in Bayesian analysis is the choice of a prior distribution, which should reflect the prior knowledge on the analysis of the parameter of interest. The bootstrap

method is a simple method of calculating the approximate dispersion of the estimator⁽²⁾. This method performs a re-sampling procedure of some data for very small samples. But to get a more accurate estimator, many more replications are needed. An expert system constructed by neural networks is applied for reliability evaluation of generating units and transmission system⁽³⁾. With actual data, the estimation of scheduled maintenance, mean time to failure (MTTF), mean time to repair (MTTR) and forced outage rate of generating units in the generation system, as well as the estimation of transient outage rate, permanent outage rate, permanent outage duration, normal rating, long term rating and short term rating of transmission equipments in the transmission system, are carried out. Although this method has many benefits for reliability estimation, it depends on the design of neural network architectures. Its extension for any other kinds of equipments and systems needs to be proved.

There are also several approaches that could be possibly used to conduct reliability assessment by means of multiple factors, such as principal component analysis (PCA), analysis of variance (ANOVA), etc.⁽⁴⁾⁽⁵⁾. However, factors related with the reliability assessment are quite different. If some of the factors are not statistical variables, these approaches are hard to use for reliability assessment by means of multiple factors. The dimensional reduction method analyzed in this paper, as a general reliability assessment methodology, conducts reliability assessment with few data by means of multiple factors. But it does not require all factors being statistical variables. Therefore, as a novel approach, it has more possibilities to use for reliability assessment.

The new concept of dimensional reduction for reliability assessment was firstly proposed by Nakamura⁽⁶⁾. This method was successfully used to build up a simple mathematical model for failure cases of submarine cables. With this model, reliability assessment and analysis on submarine cables were carried out. In addition, the dimensional reduction method was further employed for the maintenance

* Intelligent Systems Research Division, National Institute of Informatics
2-1-2, Hitotsubashi, Chiyoda-ku, Tokyo 101-8430

** Department of Advanced Systems Control Engineering, Graduate
School of Science and Engineering, Saga University
Honjomachi, Saga 840-8502

scheduling of pump systems in thermal power stations⁽⁷⁾. The mean time to failure of pumps was calculated using the dimensional reduction method and thereby reliabilities of pumps were assessed. From these applications, the dimensional reduction method has great potentials for reliability assessment of devices and systems under the situation of having only few data but with multiple factors. Since in the previous work dimensional reduction method is only explained and employed for specific objects, it is lack of more general, concrete definition and analysis on this method. Therefore, this paper presented a general definition of the concept, algorithm and application area of the dimensional reduction method for reliability assessment and analyzed the properties of this method, such as selection of factors, uniqueness of dimensional reduction results, determination of dimensional reduction sequence, etc.

2. Concepts and Data Processing

2.1 Reliability Assessment The term reliability is generally used to express a certain degree of assurance that a device or a system will operate successfully in a specified environment during a certain period of time. In general, for irreparable system, the reliability is quantitatively expressed by the probability $R(t)$, that devices or systems can be successfully operated over the period t without faults, as

$$R(t) = P\{T \geq t\} \dots\dots\dots (1)$$

For reparable system, the reliability is quantitatively expressed by the probability $A(t)$, called availability, that a reparable system can perform the defined functions correctly under specified conditions in a specified period even with faults⁽⁸⁾. Conceptually, it can be expressed as

$$\text{Availability} = \frac{\text{operating time}}{\text{operating time} + \text{down time}} \dots\dots\dots (2)$$

In actual systems, operating time is always described by mean time to failure. The down time is described by maintenance time.

In industries, most of devices and systems are reparable. The reliability assessment of reparable system is more complicated than that of irreparable system. In addition, the algorithm and the properties of reliability assessment of reparable system can be extended to irreparable system. Therefore, in this research, the object of analysis is the reliability assessment of reparable system. The term reliability assessment herein indicates mainly the availability assessment.

2.2 Dimensional Reduction The dimensional reduction method attempts to take data from a high dimensional space and map it onto a low dimensional space with the minimum of error. In most applications, dimensional reduction is carried out as a preprocessing step. At present, it is popularly used in clustering, classification, and many other data processing applications. The purpose of using dimensional reduction is to retain the most important dimensions (attributes), remove the noisy dimensions (irrelevant attributes) and reduce computational cost⁽⁸⁾. Three of the most common methods are used for dimensional reduction: Sammon mapping, principal component analysis (PCA) and neural networks.

Conventionally, dimensional reduction method is often

used when super-large data sets are available. However, the dimensional reduction method for reliability assessment proposed by Nakamura⁽⁶⁾ is quite different from the conventional dimensional reduction methods. In the process of reliability assessment, the use of dimensional reduction is to extract the characteristics of faults with few data by means of multiple factors through condensing the data.

2.3 Raw Data Processing The raw data obtained from actual systems generally can not be used directly for the dimensional reduction method due to their different formats. The description of all data with respect to the same factor must be therefore conformed to same patterns. For instance, one of simple patterns is to classify data according to whether there exist faults or not. If there have faults with respect to a specified factor, it can be defined as "Yes" or "1". On the contrary, it is defined as "No" or "0". By use of this pattern for all data with respect to the same factor, dimensional reduction process can be performed easily and understandably. Although there might exist other patterns to carry out raw data processing, as long as taking same patterns of raw data processing for each factor, it can be acceptable for the dimensional reduction method.

3. General Definition of Reliability Assessment by Dimensional Reduction Method

3.1 Essentials In a system, faults of devices and systems are always related with multiple factors. For example, concerning a power system, the factors related with faults include types of system, element, fault position, maintenance interval, mean time to failure, etc. In this paper, a multi-dimensional space is defined by use of these factors. The points in this space are supposed of the probabilities of faults. As illustrated in Fig. 1, a 3-dimensional space is described by a Cartesian coordinate, in which each axis represents a factor. Since there are actually few faults in devices and systems, in order to calculate the probability of fault at each point in this multi-dimensional space, dimensional reduction method is adopted to normalize probabilities of faults with the same base so that the multiple dimensions can be reduced to one dimension.

Fig. 1. Probabilities of faults in multi-dimensional space

The set of factors related with faults is defined as

$$\Sigma = \{X, Y, Z, \dots, U, T\} \dots\dots\dots (3)$$

which contains p factors. With respect to each factor, the number of faults is $N_f(\gamma)$, where γ denotes the factors $X, Y, Z, \dots, U, T$.

3.2 Algorithm

3.2.1 Dimensional Reduction Process From the point of view of probability theory and statistics, each fault of the device can be regarded as an event of the statistic process. Let $N_f(X_i, Y_j, \dots, T_m)$ and $N_t(X_i, Y_j, \dots, T_m)$ be the number of faults and the total number of elements exposed to faults at the level of $(X_i, Y_j, \dots, T_m)$ in the multi-dimensional space, where $X_i, Y_j, \dots, T_m$ denote the elements of $X, Y, \dots, T$, respectively. The probability of fault at the level of $(X_i, Y_j, \dots, T_m)^{(9)}$ is

$$P(X_i, Y_j, \dots, T_m) = \frac{N_f(X_i, Y_j, \dots, T_m)}{N_t(X_i, Y_j, \dots, T_m)} \dots\dots\dots (4)$$

Unfortunately, the number of faults at the level of $(X_i, Y_j, \dots, T_m)$ in the multi-dimensional space is often insufficient. The probability of fault at the level of $(X_i, Y_j, \dots, T_m)$ can be calculated as below by dimensional reduction method.

First, the dimensional reduction process is conducted to condense data from a multi-dimensional space into a one-dimensional space, by normalizing data into a base level for each factor. Thus, the equivalent data are distributed in a one-dimensional space.

If considering all the elements of factors $(Y, \dots, T)$, the total probability of fault of the element X_i in the multi-dimensional space is

$$P(X_i) = \sum_{j, \dots, m} P(X_i, Y_j, \dots, T_m) \dots\dots\dots (5)$$

In order to reduce the p -dimensional space to the $(p-1)$ -dimensional space, the probability of fault of X_i is transformed by the base X_1 for the factor X , and the relative weight $W(X_i|X_1)$ with respect to the base X_1 is calculated by

$$W(X_i|X_1) = \frac{P(X_i)}{P(X_1)} \dots\dots\dots (6)$$

Therefore, the total number of faults with respect to the base X_1 at the level of $(Y_j, \dots, T_m)$ is

$$\hat{N}_f(Y_j, \dots, T_m|X_1) = \sum_i N_f(X_i, Y_j, \dots, T_m) W(X_i|X_1) \dots\dots\dots (7)$$

Then, the probability of fault of the element Y_j of the second factor Y with respect to the base X_1 in the multi-dimensional space is

$$P(Y_j|X_1) = \frac{\sum_{i, k, \dots, m} N_f(X_i, Y_j, Z_k, \dots, T_m) W(X_i|X_1)}{\sum_{i, k, \dots, m} N_t(X_i, Y_j, Z_k, \dots, T_m)} \dots\dots\dots (8)$$

and its relative weight with respect to the base (X_1, Y_1) is

$$W(Y_j|X_1, Y_1) = \frac{P(Y_j|X_1)}{P(Y_1|X_1)} \dots\dots\dots (9)$$

where the element Y_1 is supposed as the base of factor Y .

By use of the same procedure, the probability of fault of element T_m of the last factor T with respect to the base $(X_1, Y_1, \dots, U_1)$ is

$$P(T_m|X_1, Y_1, \dots, U_1) = \frac{\sum_{i, \dots, s} [N_f(X_i, \dots, U_s, T_m) W(X_i|X_1) \dots W(U_s|X_1, \dots, U_1)]}{\sum_{i, \dots, s} N_t(X_i, \dots, U_s, T_m)} \dots\dots\dots (10)$$

3.2.2 Dimensional Expansion Process Since the weights of $W(X_i|X_1), \dots, W(T_m|X_1, \dots, U_1)$ are worked out in the dimensional reduction process, with using the dimensional expansion process, the probability of fault $P(X_i, Y_j, \dots, U_s, T_m)$ can be calculated by

$$\bar{P}(X_i, Y_j, \dots, U_s, T_m) = \frac{P(T_m|X_1, Y_1, \dots, U_1)}{W(X_i|X_1) \dots W(U_s|X_1, \dots, U_1)} \dots\dots\dots (11)$$

which is the equivalent probability of fault at the level of $(X_i, Y_j, \dots, U_s, T_m)$ in the multi-dimensional space.

3.2.3 Reliability Estimation Since the probability of fault $P(X_i, Y_j, \dots, U_s, T_m)$ is calculated by Eq. (11), with the following general algorithm, the availability at the level of $(X_i, Y_j, \dots, U_s, T_m)$ in the multi-dimensional space can be estimated.

First, the value of $\lambda(t_e|X_i, Y_j, \dots, U_s, T_m)$ with respect to the level of $(X_i, Y_j, \dots, U_s, T_m)$ is

$$\lambda(t_e|X_i, Y_j, \dots, U_s, T_m) = \int_{t_0}^{t_e} P(X_i, Y_j, \dots, U_s, T_m) \dots\dots\dots (12)$$

where t_0 and t_e denote the start time and the calculation time, respectively. It is used to form a histogram of failure rate.

Next, the failure rate is extrapolated by a probability distribution to estimate availability. Since the extrapolation of failure rate needs many different discrete values of $\lambda(t_e|X_i, Y_j, \dots, U_s, T_m)$, the fault intervals described by the variety of time t_0 and t_e in Eq. (12) are therefore divided into several levels. In general, the level of fault interval is also regarded as a factor in the multi-dimensional space and participates in the dimensional reduction process. Since it should be located at the last dimension, the factor T is defined as the level of fault interval.

In this research, Weibull distribution whose probability density function $p(t)$ characterizes with two parameters, the shape parameter σ and the scale parameter β , is selected as the extrapolating function⁽¹⁰⁾. It is expressed by

$$p(t) = \left(\beta t^{\beta-1} / \sigma^\beta \right) \exp(-t/\sigma) \left\{ \begin{array}{l} \sigma = \exp(-c/b), \quad \beta = b, \end{array} \right\} \dots\dots\dots (13)$$

where t denotes the time. With the least squares method, σ and β can be obtained by minimizing the errors between the Weibull distribution curve and the histogram of failure rate,

using the equations

$$\left. \begin{aligned} b &= \frac{N \sum_{i=1}^N x_i y_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i}{N \sum_{i=1}^N (x_i)^2 - \left(\sum_{i=1}^N x_i \right)^2} \\ c &= \frac{\sum_{i=1}^N (x_i)^2 \sum_{i=1}^N y_i - \sum_{i=1}^N x_i \sum_{i=1}^N x_i y_i}{N \sum_{i=1}^N (x_i)^2 - \left(\sum_{i=1}^N x_i \right)^2} \end{aligned} \right\} \dots\dots\dots (14)$$

$(N = N^T)$

where

$$\left. \begin{aligned} x_i &= \ln(t_i) \\ y_i &= \ln \ln(1/(1 - P(t_i))) \quad (i = 1, \dots, N^T) \end{aligned} \right\} \dots\dots\dots (15)$$

N^T denotes the number of the levels of fault intervals. Hence, mean time to failure ($M_{i,\dots,s}^f$) of a given device or system at the level of $(X_i, \dots, U_s)$ can be calculated using the formula with the calculated Weibull distribution parameters σ and β , as

$$\left. \begin{aligned} M_{i,\dots,s}^f &= \sigma \Gamma((1 + \beta)/\beta) \\ \Gamma(x) &= \int_0^\infty \theta^{x-1} \exp(-\theta) d\theta \end{aligned} \right\} \dots\dots\dots (16)$$

Finally, with the mean time to failure ($M_{i,\dots,s}^f$), the availability $a_{i,\dots,s}$ of a device or a system at the level of $(X_i, \dots, U_s)$ is

$$a_{i,\dots,s} = M_{i,\dots,s}^f / (M_{i,\dots,s}^f + M^r) \dots\dots\dots (17)$$

where M^r denotes the mean time to repair.

4. Properties Analysis

The dimensional reduction method for reliability assessment introduced in section 3 is a general method for any kinds of devices or systems. With this method, reliability assessment is conducted with only few data by means of multiple factors. In this section, several properties of this method are analyzed for the utilization of the proposed method.

4.1 Selection of Factors For different kinds of devices or systems, there may have many kinds of factors related with faults. Actually, all available factors are expected to contribute themselves to the reliability assessment in order to obtain sufficient information on faults, so that the precision of reliability assessment can be increased. Unfortunately, the factors, which are only related with the part of faults, can not be employed to participate in the dimensional reduction process because they can not include the features of all faults. It means that each selected factor should relate with all faults and it has the same number of faults as those of other selected factors, i.e.,

$$N_f(X) \equiv N_f(Y) \equiv \dots \equiv N_f(T) \dots\dots\dots (18)$$

where $N_f(\cdot)$ denotes the number of faults of each factor. Since the more factors are employed, the more precise reliability assessment can be achieved, therefore, when using dimensional reduction method, we could not ignore any useful factors for reliability assessment.

4.2 Uniqueness of Dimensional Reduction Results for Different Bases

In the dimensional reduction process introduced in section 3, the base $(X_1, Y_1, \dots, U_1)$ is supposed to calculate the probabilities of faults of any devices or systems. In fact, with any base, the probability of fault of a device or a system is unique. This property can be proved as below.

If the base $(X_1, Y_1, \dots, U_1)$ was selected, the probability of fault $P(X_i, Y_j, \dots, U_s, T_m)$ can be described by

$$\begin{aligned} P^* \{X_i, Y_j, \dots, U_s, T_m\} \\ = \frac{P\{T_m | X_1, Y_1, \dots, U_1\}}{W(X_i | X_1) \dots\dots\dots W(U_s | X_1, \dots, U_1)} \dots\dots\dots (19) \end{aligned}$$

If we change the base as $(X_\alpha, \dots, U_\mu)$, the probability of fault can be also calculated by

$$\begin{aligned} P^{**} \{X_i, Y_j, \dots, U_s, T_m\} \\ = \frac{P\{T_m | X_\alpha, \dots, U_\mu\}}{W(X_i | X_\alpha) \dots\dots\dots W(U_s | X_\alpha, \dots, U_\mu)} \dots\dots\dots (20) \end{aligned}$$

Based on Eq. (11), we can obtain

$$P\{X_\alpha, \dots, U_\mu, T_m\} = \frac{P\{T_m | X_1, Y_1, \dots, U_1\}}{W(X_\alpha | X_1) \dots\dots\dots W(U_\mu | X_1, \dots, U_1)} \dots\dots\dots (21)$$

Besides, by Eq. (11), it is easy to deduce that

$$P\{X_\alpha, \dots, U_\mu, T_m\} = P\{T_m | X_\alpha, \dots, U_\mu\} \dots\dots\dots (22)$$

If substituting (21) in (20), and according to the definition of weight, it can be obtained that

$$P^* \{X_i, Y_j, \dots, U_k, T_m\} = P^{**} \{X_i, Y_j, \dots, U_k, T_m\} \dots\dots\dots (23)$$

Therefore, it concludes that the probability of fault is independent with the selection of base.

4.3 Determination of Dimensional Reduction Sequence

The main impact on the results of dimensional reduction method comes from the dimensional reduction sequence. And it is also the most difficult point to determine. In this section, we attempt to propose several valuable criteria for the selection of dimensional reduction sequence.

Criterion I: Restriction of physical meaning.

When we select a dimensional reduction sequence, restriction of physical meaning should be taken into account. Probability of fault of a factor should be reasonable during the course of dimensional reduction. If intermediate results violate the restriction of physical meaning obviously, the final results can not be adopted for the further work even they seems reasonable. Therefore, when we select the sequence of dimensional reduction process, we must consider the restriction of physical meaning to guarantee the reasonability of the final probability of fault of the desired factor. With respect to various objects of reliability assessment, the restrictions of physical meaning are different. It is hard to define a general description of restrictions in terms of physical meaning. In the simulation part of this paper, we will give an example to illustrate the importance of this criterion when selecting the sequence of dimensional reduction process.

Criterion II: Calculation accuracy of the probability of

fault of a factor.

Another criterion for the determination of dimensional reduction sequence is the calculation accuracy of probability of fault of a factor. Since the probability of fault of the current dimension will sequentially employ the weights representing the relation of probabilities of faults of the former dimension with respect to the base, the probability of fault of the former dimension must be more accurate than that of the current dimension in order to reduce the error propagation.

Since the numbers of faults of all factors are similar, the smaller the number of elements of the factor is, the higher the calculation accuracy of the probability of fault of each element is. Therefore, before we select the sequence of dimensional reduction process, we can order the numbers of elements of factors from small to large, as

$$N^1 \leq N^2 \leq \dots \leq N^p \dots\dots\dots(24)$$

and then calculate the probability of fault of the factor from the factor with the small number of elements to that with the larger number of elements.

Criterion III: Other regulations of dimensional reduction sequence on some factors.

In the dimensional reduction method, there may have other regulations of dimensional reduction sequence on some factors, especially requiring some additional computation during the course of dimensional reduction process. Therefore, the sequences of some factors in the dimensional reduction process are fixed. We must obey the regulations of these factors.

With our experience, if we select the sequence of dimensional reduction process according to the above three criteria, the reasonability and the accuracy of probabilities of faults of factors calculated by the dimensional reduction method can be guaranteed.

4.4 Effectiveness for the Case of Few Data and Accuracy Improvement of Reliability Assessment From the general definition of dimensional reduction method, dimensional reduction method not only can adapt to implement reliability assessment with few data, but also can improve accuracy of reliability assessment. In order to realize these purposes, several means are employed in this method. One is the data collection. For each element of a factor, all available data are summed up to calculate the probability of fault of this element, as Eq. (5). Another is the data condensation. With respect to a same base, probabilities of faults of all elements of a factor are transformed and summed up to calculate the probability of fault of the element of another factor, as Eq. (8). As calculating the probability of fault of the element of the last factor, it should have enough data to work out accurate value of probability of fault. With the same level of accuracy as the last factor, probabilities of faults of all elements of all factors can be calculated by using dimensional expansion process. Moreover, the more data and factors there have, the higher accuracy probability of fault is. Therefore, by use of dimensional reduction method, the reliability assessment with few data can be implemented and the accuracy of reliability assessment can be improved.

4.5 Application Areas Since the reliability of a device or a system assessed by use of the dimensional reduction method is interrelative with multiple factors, therefore,

the application area of this reliability value is wider than the conventional reliability value, such as fault analysis, reliability design, optimization of maintenance scheduling, etc. In addition, this dimensional reduction method is not limited to some special objects, though it was originally developed for the device of power delivery⁽⁶⁾.

5. Simulation

5.1 Preliminary In order to verify the effectiveness of dimensional reduction method and discuss some important properties of it, we made simulations based on the practical request and actual data from Kyushu Electric Power Co. of Japan.

The objects of simulation are 60 different types of electrical elements in 18 thermal power stations. The raw data contain the information on examination time, fault time, maintenance and exchange time, and so on. Fig. 2 illustrates an actual table including part of raw data of electrical elements of the 18th thermal power station, one of thermal power stations in Kyushu Electric Power Co. of Japan, in which “ ” indicates the examination time, “ ” indicates the fault time and black rectangle indicates the maintenance and exchange time. The raw data are recorded from the year of 1989 and all of them have same formats. In order to conduct reliability assessment, based on the raw data, the input data were tabulated as given in Table 1, including the element number Y_j , maintenance interval and mean time to failure (MTTF).

Next, the significant factors for the dimensional reduction method were defined as below.

- Thermal power station $X_i (i = 1, \dots, 18)$.
- Electrical element $Y_j (j = 1, \dots, 60)$: it is classified by Kyushu Electric Power Co.
- Type of average maintenance interval $Z_k (k = 1, \dots, 4)$: it is classified into four types based on the investigation of fault data and balancing the distribution of fault data in a predetermined period: Z_1 denotes that the maintenance interval smaller than 70% of average maintenance interval, Z_2 70%–90%, Z_3 90%–150%, Z_4 150%–250%.
- Level of fault interval $T_m (m = 1, \dots, 6)$: it is classified into defined as 6 levels (unit: hrs): 0–

MATSUMURA No.13 (Electric Elements)

Items	60	61	62	63	1	2	3	4	5	6	7	8	9	10	11	12
	1.151	1.514	1.514	1.514	1.514	1.514	1.514	1.514	1.514	1.514	1.514	1.514	1.514	1.514	1.514	1.514
Star																
Star																
Bearing																
Sealing																
Thyristor																
Magnetic Separator																
Rectifier																
Regulation Valve																
Main Pump																
Recycle Pump																
Vacuum Pump A, B																

Fig. 2. Part of raw data of electrical elements of the 18th thermal power station

Table 1. Input data of elements of the 1st thermal power station for reliability assessment

Y_j	Maintenance Interval[hrs]	MTTF[hrs]
1	1685.8	5052.3
2	1740.7	4540.8
3	1740.7	3744.5
4	1529.6	2921.5
8	1898.3	0.0
9	1898.3	0.0
10	1727.3	0.0
12	1727.3	0.0
13	3454.5	3333.0
15	3454.5	4638.5
17	3454.5	5944.0
21	3454.5	4638.5
22	2303.0	0.0
23	2303.0	0.0
24	3032.8	2088.3
27	2878.8	4636.0
28	2303.0	0.0
31	2303.0	0.0
34	2303.0	0.0
35	2303.0	0.0
36	2303.0	0.0
38	2303.0	0.0
40	4236.3	7360.8
41	3166.7	6060.1
44	1868.0	0.0
45	4337.0	0.0
47	3576.0	6426.5
48	5222.0	9520.0
51	1930.0	3333.0
53	1930.0	3333.0
54	1930.0	0.0
55	3173.3	3373.5
59	2148.1	2327.4
60	2790.8	2201.4

20,000, 20,000–40,000, 40,000–60,000, 60,000–80,000, 80,000–100,000 and longer than 100,000.

5.2 Simulation Results of Reliability Assessment

First of all, the dimensional reduction sequence was determined by the criteria introduced in 4.3. The level of fault interval T was definitely in the fourth dimension. According to the restriction of physical meaning for each element, i.e., the shorter maintenance interval is, the smaller its probability of fault is, the type of average maintenance interval Z was selected as the third dimension. A discussion on this point will be given in section 6. According to the criterion II, since the number of elements of system X was smaller than that of electrical element Y , X was selected as the first dimension and Y as the second dimension considering about the accuracy of statistic process. Therefore, the sequence of dimensional reduction was $X \rightarrow Y \rightarrow Z \rightarrow T$.

Owing to the uniqueness of dimensional reduction results, (X_1, Y_1, Z_1) was selected as the base of the dimensional reduction method for simplicity. Then the dimensional reduction process was carried out with the all input data of 18 thermal power stations. Table 2 illustrates the dimensional reduction results of X . Table 3 illustrates the dimensional reduction results of Y . Table 4 illustrates the dimensional reduction results of Z . And table 5 illustrates the dimensional reduction results of T . In these tables, $N_f, N_t, \hat{N}_f$ denote the number of faults, total number of elements exposed to faults and the number of faults with respect to the base, respectively. W denotes the relative weight, and $P(\cdots)$ denotes the relative probability of fault.

Therefore, the probability of fault at each level in the 4-dimensional space with respect to the base level (X_1, Y_1, Z_1) can be obtained by the dimensional reduction expansion given in (11), with the actual data of electrical elements acquired from 18 thermal power stations in Kyushu Electric Power Co. Based on the probabilities of faults, availabilities of each electrical element in each system with the various types of average maintenance intervals can be obtained. In Fig. 3, the estimated availabilities of each electrical element

Table 2. Dimensional reduction results of X

X_i	N_t	N_f	$P(X_i)$	W
1	112.0	31.0	0.277	1.000
2	159.0	60.0	0.377	0.733
3	137.0	50.5	0.369	0.751
4	101.0	22.0	0.218	1.271
5	156.0	53.5	0.343	0.807
6	179.0	70.5	0.394	0.703
7	144.0	43.0	0.299	0.927
8	115.0	36.0	0.313	0.884
9	121.0	39.0	0.322	0.859
10	130.0	56.5	0.435	0.637
11	159.0	60.0	0.377	0.733
12	142.0	48.5	0.342	0.810
13	200.0	84.5	0.422	0.655
14	86.0	35.5	0.413	0.671
15	79.0	23.0	0.291	0.951
16	91.0	33.5	0.368	0.752
17	92.0	32.0	0.348	0.796
18	218.0	92.5	0.424	0.652

Table 3. Dimensional reduction results of Y

Y_j	N_t	N_f	$\hat{N}_f$	$P(Y_j X_1)$	W
1	40.0	19.0	15.2	0.380	1.00
2	21.0	7.5	5.9	0.283	1.34
3	36.0	17.5	13.6	0.378	1.00
4	39.0	15.5	12.1	0.309	1.23
5	11.0	1.5	1.3	0.117	3.24
6	10.0	3.5	2.3	0.228	1.67
7	1.0	0.5	0.3	0.328	1.16
8	19.0	6.0	4.3	0.229	1.66
9	20.0	7.5	6.2	0.311	1.22
10	24.0	8.0	6.7	0.279	1.36
11	1.0	0.5	0.3	0.318	1.19
12	14.0	4.5	3.2	0.229	1.66
13	17.0	7.5	6.4	0.375	1.01
14	3.0	1.5	1.0	0.326	1.16
15	27.0	11.0	8.8	0.327	1.16
16	5.0	1.0	0.8	0.154	2.47
17	18.0	5.5	4.1	0.227	1.67
18	19.0	7.0	5.2	0.271	1.40
19	4.0	1.5	1.1	0.270	1.41
20	3.0	1.0	0.8	0.252	1.51
21	14.0	5.0	4.3	0.305	1.24
22	50.0	16.0	12.2	0.243	1.56
23	65.0	25.0	18.3	0.281	1.35
24	22.0	9.0	6.9	0.316	1.20
25	5.0	0.0	0.0	0.000	0.00
26	6.0	2.5	1.9	0.311	1.22
27	19.0	8.5	6.7	0.354	1.07
28	21.0	5.0	3.7	0.176	2.15
29	9.0	1.5	1.2	0.129	2.95
30	4.0	1.0	0.8	0.202	1.88
31	18.0	7.5	6.1	0.336	1.13
32	11.0	1.5	1.2	0.106	3.58
33	9.0	3.5	2.5	0.276	1.38
34	24.0	4.5	3.1	0.127	2.99
35	22.0	5.0	3.6	0.164	2.32
36	6.0	0.5	0.4	0.059	6.48
37	3.0	0.0	0.0	0.000	0.00
38	11.0	3.0	2.2	0.195	1.94
39	8.0	1.0	0.7	0.093	4.06
40	72.0	38.5	29.7	0.413	0.92
41	32.0	14.5	11.3	0.354	1.07
42	24.0	10.0	7.1	0.296	1.28
43	10.0	4.0	2.6	0.262	1.45
44	20.0	9.0	6.3	0.317	1.20
45	36.0	15.0	11.1	0.308	1.23
46	2.0	1.0	0.7	0.367	1.04
47	72.0	37.0	29.0	0.403	0.94
48	65.0	33.0	25.1	0.387	0.98
49	56.0	26.0	18.6	0.332	1.14
50	4.0	2.0	1.4	0.361	1.05
51	11.0	5.0	3.6	0.331	1.15
52	16.0	4.0	3.0	0.185	2.05
53	30.0	13.0	10.2	0.340	1.12
54	5.0	0.0	0.0	0.000	0.00
55	61.0	30.5	22.6	0.371	1.02
56	13.0	7.5	5.6	0.427	0.89
57	8.0	4.0	2.7	0.341	1.11
58	5.0	2.0	1.3	0.261	1.45
59	599.0	240.5	184.8	0.308	1.23
60	621.0	147.5	118.2	0.190	2.00

Table 4. Dimensional reduction results of Z

Z_k	N_t	N_f	$\hat{N}_f$	$P(Z_k X_1, Y_1)$	W
1	435.0	165.0	156.3	0.359	1.000
2	756.0	304.5	297.8	0.394	0.912
3	649.0	242.5	268.4	0.414	0.869
4	328.0	113.0	137.5	0.419	0.857

Table 5. Dimensional reduction results of T

T_m	N_t	N_f	$\hat{N}_f$	$P(T_m X_1, Y_1, Z_1)$
1	580.0	285.5	282.4	0.487
2	448.0	208.5	203.1	0.453
3	254.0	121.0	111.3	0.438
4	184.0	93.5	80.5	0.437
5	129.0	66.5	58.0	0.450
6	98.0	50.0	44.0	0.448

in the 12th thermal power station are illustrated. There are 60 electrical elements. For each element, there are four availabilities with respect to four types of average maintenance intervals.

Fig. 3. Estimated electrical element availabilities of the 12th thermal power stationTable 6. Probability of fault $P\{T_m|X_i, Y_j, Z_k\}$ of the 20th element in the 8th thermal power station with different bases

T_m	Base (X_1, Y_1, Z_1)			
	(X_8, Y_{20}, Z_1)	(X_8, Y_{20}, Z_2)	(X_8, Y_{20}, Z_3)	(X_8, Y_{20}, Z_4)
1	0.125	0.137	0.144	0.146
2	0.215	0.236	0.247	0.251
3	0.264	0.290	0.304	0.308
4	0.300	0.329	0.345	0.350
5	0.326	0.358	0.375	0.380
6	0.345	0.378	0.397	0.402

T_m	Base (X_9, Y_{10}, Z_3)			
	(X_8, Y_{20}, Z_1)	(X_8, Y_{20}, Z_2)	(X_8, Y_{20}, Z_3)	(X_8, Y_{20}, Z_4)
1	0.125	0.137	0.144	0.146
2	0.215	0.236	0.247	0.251
3	0.264	0.290	0.304	0.308
4	0.300	0.329	0.345	0.350
5	0.326	0.358	0.375	0.380
6	0.345	0.378	0.397	0.402

6. Discussion

In this section, with simulation results some properties introduced in section 4 are demonstrated. Firstly, we have changed the base to perform dimensional reduction process again in order to prove the uniqueness of dimensional reduction results. One selected base is (X_9, Y_{10}, Z_3) and another is (X_1, Y_1, Z_1) . The selection of the base is arbitrary. In Table 6, the probabilities of fault of $P\{T_m|X_i, Y_j, Z_k\}$ of the 20th element in the 8th thermal power station with different bases are given. Since the probabilities of faults with different bases are same, the availabilities at each level in the 4-dimensional space with different bases should be the same.

Next, with simulation results the determination of dimensional reduction sequence by considering the restriction of physical meaning is illustrated. In the availability assessment of electrical elements of thermal power stations, the restriction of physical meaning is defined that, the shorter maintenance interval is, the smaller its probability of fault is. In table 7, we can see the calculation results of the probabilities of faults of $P(Z)$ with different dimensional reduction sequences. According to the definition of the type of average maintenance interval Z , the probabilities of fault of the various types of average maintenance intervals should be as

$$P\{Z_1\} \leq P\{Z_2\} \leq P\{Z_3\} \leq P\{Z_4\} \dots \dots \dots (25)$$

From the simulation results given in table 7, only in the cases of $X \rightarrow Y \rightarrow Z$ and $Y \rightarrow X \rightarrow Z$, the probabilities of faults of the various types of average maintenance intervals have the relation as (25). Therefore, the type of average maintenance interval Z should be only selected as the third dimension.

Besides, in Fig. 2 and Table 1 there are quite few raw data

Table 7. Calculation results of probabilities of faults $P(Z)$ with different dimensional reduction sequences

	$X \rightarrow Y \rightarrow Z$	$X \rightarrow Z \rightarrow Y$	$Y \rightarrow X \rightarrow Z$
1	0.359	0.295	0.331
2	0.394	0.295	0.353
3	0.414	0.297	0.370
4	0.419	0.274	0.382

	$Y \rightarrow Z \rightarrow X$	$Z \rightarrow X \rightarrow Y$	$Z \rightarrow Y \rightarrow X$
1	0.441	0.379	0.379
2	0.515	0.403	0.403
3	0.505	0.374	0.374
4	0.514	0.345	0.345

on faults of devices and systems. Some devices even have no faults and their MTTFs are zero. From the simulation results reliabilities of all devices and systems can be worked out corresponding to various types of average maintenance intervals, as illustrated in Fig. 3. Concerning the accuracy of reliability assessment of dimensional reduction method, it has been evaluated by engineers who use these reliability values for maintenance interval scheduling. Comparing with the conventional maintenance interval scheduling, the scheduling with using reliability assessed by dimensional reduction method is more reasonable and fit for the actual situation.

7. Conclusions

The dimensional reduction method is a powerful algorithm to assess reliability with only few data by means of multiple factors. In this paper, a general definition of this algorithm was given. In particular, the properties of dimensional reduction method, such as selection of factors, uniqueness of dimensional reduction results for different bases, determination of dimensional reduction sequence, etc., were discussed. With the actual data on electrical elements in Kyushu Electric Power Co., the effectiveness of the proposed method has been verified and some properties of this method were also demonstrated. The dimensional reduction method can be widely used in reliability assessment of devices and systems. The results of reliabilities assessed by this method are greatly useful for industries.

Acknowledgment

The authors would like to thank Dr. H. Hatazaki, general manager of Thermal Power Operation Group, Kyushu Electric Power Co., Japan, for his valuable suggestions, and Mr. Hanato and Mr. Koki, master students of Graduate School of Science and Engineering, Saga University, for their help in data processing.

(Manuscript received Sep. 22, 2003,

revised Feb. 25, 2004)

References

- (1) S. Zacks: Introduction to reliability analysis: probability models and statistical methods, Springer-Verlag, New York (1992)
- (2) T. Seki and S. Yokoyama: "Robust parameter-estimation using the bootstrap method for the 2-parameter Weibull distribution", *IEEE Trans. Reliability*, Vol.45, No.1, pp.34–41 (1996)
- (3) N. Amjady and M. Ehsan: "Evaluation of power systems reliability by an artificial neural network", *IEEE Trans. Power Syst.*, Vol.14, No.1, pp.287–292 (1999)
- (4) R.C. Gonzalez and R.E. Woods: Digital image processing, Addison Wesley Publishing Company (1992)
- (5) T.J. Rick and J. Thayer: Introduction to analysis of variance, Sage Publications, Thousand Oska, CA (2001)
- (6) M. Nakamura, N. Nanayakkara, H. Hatazaki, and K. Tsuji: "Reliability analysis of submarine power cables and determination of external mechanical protections", *IEEE Trans. Power Delivery*, Vol.7, pp.895–902 (1992)
- (7) M. Nakamura, T. Katafuchi, and H. Hatazaki: "Decisions for maintenance-intervals of equipment in thermal power stations, based on few data", *IEEE Trans. Reliability*, Vol.50, No.4, pp.360–364 (2001)
- (8) C. Ding, X.F. He, H.Y. Zha, and H.D. Simon: "Adaptive dimension reduction for clustering high dimensional data", Proc. of 2nd IEEE International Conference Data Mining, pp.147–154 (2002)
- (9) H. Cramer: Mathematical methods of statistics, pp.137–165, Kaigai Publications, Tokyo (1973)
- (10) M. Mackisack and R. Stillman: "A cautionary tale about Weibull analysis", *IEEE Trans. Reliability*, Vol.45, No.2, pp.244–248 (1996)

Tao Zhang (Member) was born in Beijing, China, in 1969. He received B.S., M.S. and Ph.D. degrees in Electrical Engineering from Tsinghua University of China, in 1993, 1995 and 1999, respectively. In 2002, he received his second Ph.D. degree in electrical engineering in the Department of Advanced Systems Control Engineering, Graduate School of Science and Engineering, Saga University of Japan.

At present, he is a researcher in the National Institute of Informatics of Japan. His current research interests include information engineering, nonlinear system control theory, power system control and maintenance scheduling, robotics and neural networks.

Masatoshi Nakamura (Member) was born in Fukuoka, Japan, in 1943. He received B.S., M.S. and Ph.D. degrees in electrical engineering from Kyushu University, in 1967, 1969 and 1974, respectively.

From 1973 to 1974, he was a Research Associate in Kyushu University. Since 1974, he has been with the Faculty of Science and Engineering, Saga University, where he is currently a Professor in the Department of Advanced Systems Control Engineering in the Graduate School. His research interests include systems control theory and its applications, especially power system control, thermal flow-control, robotics and biomedical engineering.

Prof. Nakamura is a fellow of the Society of Instrument and Control Engineers of Japan, and a member of the IEEE Control Systems Society, the Institute of Electrical Engineers of Japan, the Robotics Society of Japan and the Institute of Systems, Control and Information of Japan.