An Innovative Approach for Isolating Faulty Parameters
Sachin Kumar, Eli Dolev, and Michael Pecht
Prognostics Health Management Group
Center for Advanced Life Cycle Engineering (CALCE)
University of Maryland, College Park, MD 20742
301-405-5323
pecht@calce.umd.edu

Abstract—This paper presents an innovative diagnostic approach that includes detection and fault isolation using the Mahalanobis distance (MD). The fault isolation approach is based on the analysis of residual MD values corresponding to performance parameters. The residual value is calculated by taking the differences between MD values estimated in two different scenarios: first, when a performance parameter is present, and second, when that performance parameter is absent. The residual of MD values for each parameter is obtained by using training data under several experiments planned by design-of-experiment concept, to analyze impact of each parameter. The distribution of residual MD values for each parameter is analyzed and a 95% probabilistic range is established. This range represents the expected contribution by parameters towards healthy systems MDs, and it is used to identify parameters that are responsible for the anomalous behavior of a system. Parameters that fall beyond the threshold limit are considered responsible for the anomalous behavior, and the parameter that has lowest residual value is isolated as the faulty parameter. A case study on notebook computers is presented to demonstrate and test the suggested new approach’s ability to isolate faulty parameters.1 2

Key terminologies used in fault diagnostics are fault, failure, and fault isolation. Fault is an unpermitted deviation from acceptable behavior of at least one performance property of a system. Failure is a permanent interruption of the system’s ability to perform a required function under specified operating conditions. It is not always easy to determine system failure because some faults may not lead to systems functionality loss; therefore, faults should be characterized and defined. Fault detection is the determination of faults present in a system. Fault isolation is the determination of the type and location of a fault, and it follows fault detection [4] [5]. One or more performance parameters may reflect different anomaly and failure modes of a system. Identification of performance parameters is essential for locating faulty components in a system that could cause the system to become unstable and even unusable [6] [7].

Fault isolation methods are broadly grouped into two classes: model-based and data-based. Model-based methods are generally functionality-dependent: system functionality is modeled into mathematical form and residuals are connected to specific faults [8] [9]. These methods are usually based on a deterministic process model that must be accurate to function effectively. This approach is suitable for isolating specific known faults. On the other hand, data-based methods rely exclusively on performance parameter measurements. For fault detection, under fault-free behavior, thresholds are created based on historical measurements to represent normal operation. For fault isolation, historical data obtained from the system under faulty behavior are used to define different unhealthy states of a system in order to distinguish different faults. Fault isolation is accomplished by comparing the current system

1 978-1-4244-2622-5/09/$25.00 ©2009 IEEE.
2 IEEEAC paper #1715, Version 4, Updated Nov 02, 2008
health in the state space with the known regions of the state
space and/or fault directions during faulty operation. Some
data-based approaches consider the contribution of
particular states to the overall shift from normal operation
for identifying and isolating faults [10] [11].

Many data-based fault isolation techniques take advantage
of principle component analysis (PCA) to more effectively
handle large amounts of data or to find relationships within
the data [12]. In this technique, the contribution plots are
observed in order to identify faulty parameters. Some
techniques define state models for parameters to capture
linear or nonlinear characteristics of faults and system
performance parameters. The residual generated by
estimation from models and observed data captures the
characteristics of faults [13]. In this case, a model is defined
for each fault type to isolate the fault. Defining a fault-
specific model is only useful for identifying specific known
faults. In the PCA approach, reduction in dimensionality is
at the cost of lost information [14]. The contribution plot, in
the PCA approach, provides the contribution of parameters
in the variability of an observation. However, higher
parameter contribution does not indicate that a parameter is
exhibiting anomalous behavior. The regression model of
performance parameters often fails to capture variability in
the parameters.

Another common data-based fault detection method utilizes
Mahalanobis distance (MD). The MD method is good for
two reasons; first, MD reduces a multivariate system to a
univariate system, and second, MD is sensitive to inter-
variable changes in a multivariate system. It includes all
observed performance parameters in defining the health of a
system (i.e., no information loss in the health estimation
process). MD based approach has been used for fault
detection in notebook computers [15]. In addition, fault
isolation using MD has been performed in some cases but in
these cases data from both healthy and unhealthy system
were used to isolate faults [16].

In this paper, a data-based approach utilizing the residual
Mahalanobis distance is proposed. It does not require
definition of a faulty state during training and fault
isolation, so it does not depend on any specific fault type.
The approach defines the threshold bound on parameter
residuals, which are estimated from healthy data. The
residual for each parameter is obtained by performing
experimental analysis planned by he design-of-experiment
(DoE). DoE is formed to evaluate the impact of each
parameter on MD value. In MD method, DoE concept has
been used to determine parameters that are contributing
most to Mahalanobis distance [17] [18].

Experiments were performed on notebook computers to
generate data and validate the data-based fault isolation
approach. The experimental details, the algorithmic
approach to fault isolation, and a case study are discussed
later in the paper. Both MD and DoE are briefly described
in the following sections.

Mahalanobis Distance

The Mahalanobis distance methodology is a process of
distinguishing multivariable data groups by a univariate
distance measure that is defined by several performance
parameters. The MD value is calculated using the
normalized value of performance parameters and their
correlation coefficients, which is the reason behind MD’s
sensitivity [17] [18]. A data set formed by measuring
performance parameters of a healthy product is used as
training set for MD. The collection of MD values for a
normal system is known as the Mahalanobis Space.

The parameters collected from a system are denoted as $X_i$, where
$i=1, 2, ..., m$. The observation of the $j$th parameter on the $i$th
instance is denoted by $x_{ij}$, where $i=1, 2, ..., m$, and
$j=1, 2, ..., n$. Here, $m$ is the number of parameters, and
$n$ is the number of observations. Thus the $(m \times 1)$ data
vectors for the normal group are denoted by $X_i$, where $j=1, 2, ..., n$. Each individual parameter in each data vector is
normalized by subtracting the mean of the parameter ($X_i$) and
dividing it by the standard deviation ($S_j$). These mean
and standard deviations are calculated from the healthy
data. Thus, the normalized values are as follows:

$$z_{ij} = \frac{x_{ij} - \bar{X}_i}{S_j}, \quad i=1, 2... m, j=1, 2... n,$$ 

(1)

where, $\bar{X}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}$ and $S_j = \sqrt{\frac{\sum_{i=1}^{n} (x_{ij} - \bar{X}_i)^2}{n-1}}$

Next, the values of the MDs are calculated for the normal
items using the following:

$$MD_j = \frac{1}{m} z_j^T C^{-1} z_j$$

(2)

where $z_j=[z_{1j}, z_{2j}, ..., z_{mj}]$ is a transpose vector of vector $z_j$
that comprises $z_{ij}$, and $C$ is the correlation matrix calculated as:

$$C = \frac{1}{(n-1)} \sum_{j=1}^{n} z_j z_j^T$$

(3)

For fault detection, a threshold MD is defined using training
(i.e., healthy) data. For a test system, the MD value is
calculated for each observation by using the performance
parameter’s mean, standard deviation, and a correlation
coefficient matrix obtained from the training data.

Design of Experiments

Design of Experiments (DoE) is a structured, organized
method for determining the relationship between different
factors (parameters) affecting a process and the output of that process. DoE is used to quantify indeterminate measurements of factors through observance of forced changes made methodically as directed by mathematically systematic tables. For evaluating factors, DoE involves designing a set of experiments in which all relevant factors are varied systematically. The factors that influence the process output most are identified by analyzing the output of these experiments [18].

Different performance parameters are considered as factors of the process while applying DoE approach to MD method and the MD value is used as an output of the process. In order to do fault isolation, DoE experimental methods implement “one change at a time” because it allows for a judgment on the significance to the output of input variables acting alone. Some kinds of dependency or interaction among parameters are considered in MD value since the MD method utilizes correlation among parameters.

2. THE SUGGESTED FAULT ISOLATION METHODOLOGY

A diagnostic algorithm should provide the capability of fault isolation in addition to fault detection. The detection capability identifies a fault (i.e., malfunction) in the system and allows for making a binary classification of the system health (faulty or non-faulty). The fault isolation capability identifies faulty parameters by the suggested methodology. The fault mode and location in the system can be determined by applying reasoning to the faulty parameters.

In the fault detection approach, a system’s health is classified by comparing the MD value with a threshold MD value (\(\tau\)) defined from training (healthy) data. A fault detection approach (Figure 1) starts with the monitoring of performance parameters followed by MD calculation and comparison with threshold value for classification of a system’s health. In the case that the system is unhealthy, further investigation on determining the type of fault will be performed by the fault isolation approach.

The fault isolation approach discussed here is based on the comparative analysis of residual MD. A residual MD bound (\(\Delta MD_{ia}\)) is defined by analyzing healthy systems. The threshold defining procedure (Figure 2) starts with constructing an orthogonal DoE considering “one change at a time.” Training data for each experiment is extracted from the training data collected from a healthy system. In each experimental run, the MD value corresponding to each observation is calculated where \(MD_{p}\) represents a MD calculated considering all parameters and \(MD_{ia}\) represents a MD calculated considering parameter \(i\) is absent.

The residual MD corresponding to a parameter \(i\) is represented by \(\Delta MD_{i}\). The \(\Delta MD_{i}\) is obtained by subtracting the \(MD_{p}\) from \(MD_{ia}\). For each observation, the \(\Delta MD_{i}\) is obtained and the distribution of \(\Delta MD_{i}\) represents a distribution of a parameter’s contribution in MD. A threshold value (\(\Delta MD_{ia}\)), corresponding to a 95% bound on the \(\Delta MD_{i}\) for each parameter is defined as an expected range of \(\Delta MD\) for parameter \(i\).

\[
\Delta MD_{i} = MD_{ia} - MD_{p}
\]

Each parameter is evaluated to identify important parameters that should be considered first during parameter isolation. To identify important parameter, for each experimental run the signal-to-noise ratio (S/N ratio) is calculated using the MD values corresponding to each observation of the training data. When the true levels of abnormality are not known, larger-the-better-type S/N ratios are used because, for the abnormality, the MD value should be large. Taguchi’s larger-the-better S/N ratio is defined as

\[
-10\log \left( \frac{1}{n} \sum_{i=1}^{n} \left( \frac{1}{MD_i} \right) \right)
\]

where, \(n\) is the number of observations for an entity.

\[\text{(5)}\]
For a given parameter, \( i \), the average value of the S/N ratio is determined over all the runs with the parameter present, \( S/N_p \), and parameter absent, \( S/N_a \). The loss in the S/N ratio is defined as the difference between two S/N ratios: \( S/N_a - S/N_p \). If this loss is negative, then the parameter \( i \) is considered important, otherwise, it is not.

For a test system, the procedure for isolating a faulty parameter is shown in Figure 3. Each observation from a test system is analyzed for each experimental run. The MD value for each run of test observations is calculated where the mean, standard deviation, and correlation matrix used are from a healthy system. The \( \Delta MD_i \) for each parameter of an observation is calculated and compared with the parameter threshold \( \Delta MD_{th} \). The parameter with an \( \Delta MD_i \) beyond the threshold (\( \Delta MD_{th} \)), is picked as the potential faulty parameter, and the highest \( -\Delta MD_i \) is determined to be the faulty parameter. Other parameters are evaluated as well in order to validate that the identified faulty parameter has been correctly identified.

### 3. A Case Study

The fault detection and faulty parameter isolation methodology was applied to a notebook computer. Experiments were conducted on ten brand new identical notebook computers with an assumption that these systems were representative of normal/healthy systems. The data from the healthy computers was used to define a healthy baseline and to identify specific parameter behaviors. The performance parameters monitored during experiments were fan speed, CPU temperature, motherboard temperature, video card temperature, two power saving states (%C2, and %C3), %CPU usage, and %CPU throttle. They are represented as P1 to P8 in the following discussion.

The experiment was designed to replicate the real-time usage of computers. The computers were exposed to six different environmental conditions, as shown in Table 1. For each temperature-humidity combination, four usage conditions and three power supply conditions were considered. A set of user activities was defined to execute different usage condition on the laptop computers. In total, 72 experiments were conducted. The same usage conditions were applied simultaneously to all computers to achieve time synchronization between the laptop and software application responses. The computer’s power mode was always set to ON. The screen saver and hibernation option were disabled to prevent these functions from occurring during the experiment.

<table>
<thead>
<tr>
<th>Temperature and Humidity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. 5°C with uncontrolled RH</td>
</tr>
<tr>
<td>2. 25°C with 55% RH</td>
</tr>
<tr>
<td>3. 25°C with 93% RH</td>
</tr>
<tr>
<td>4. 50°C with 20% RH</td>
</tr>
<tr>
<td>5. 50°C with 55% RH</td>
</tr>
<tr>
<td>6. 50°C with 93% RH</td>
</tr>
</tbody>
</table>

The Mahalanobis distance values were obtained for the experimental (i.e., training) datasets. These values were used to create a baseline for healthy system. The baseline was used to compare MD values of test systems and detect anomalies. The next objective was to isolate the parameter that was responsible for the fault in the computer.

The orthogonal DoE was designed with “one change at a time.” For this study, 8 parameters were monitored and with “one change at a time,” 9 experimental runs were defined, as listed in Table 2. Each cell entry represents parameters status by MD calculation: “1” represents that a parameter is included, and “2” represents that a parameter is excluded from the MD calculation.

For each experimental run (Table 2), an S/N ratio was calculated from the training data. The change in S/N ratio with the inclusion and exclusion of each parameter in MD calculation is presented in Figure 4. The parameters that have downward slopes (i.e., a loss of information due to a parameter’s absence) were considered to be significant ones, and the degree of slope defined the degree of the parameter’s importance in MD calculation. Figure 4 suggests that at least 4 out of 8 parameters must be considered for MD calculation, including fan speed, CPU temperature, motherboard temperature, and video card temperature.
Table 2. Orthogonal Design of Experiment

<table>
<thead>
<tr>
<th>No.</th>
<th>Performance Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
</tr>
</tbody>
</table>

The MD value for the first experiment contains all parameters. The second, third, and subsequent experiments each exclude one parameter and the MD is calculated for these experiments. Subtracting the first experiment’s MD from these experiments gave the quantitative measure $\Delta MD_i (= MD_{a} - MD_{p})$ of a parameter’s contribution. $\Delta MD_i$ due to the absence of different parameters in an observation’s MD value is shown in Figure 5.

Figure 5 – $\Delta MD$ values for an observation of training data.

From the training data, $\Delta MD$s for each observation were calculated. Sample distributions of the $\Delta MD$ for four parameters from the training data are shown in Figure 6.

A field-returned notebook computer was used to verify the anomaly detection methodology. Five thousand data points from the field returned computer were analyzed. Figure 7 shows a baseline healthy system (lower plot) and test system (above plot). This demonstrates that the test system had an anomaly that was detected by the MD method.

At the second stage, in-order to isolate faulty parameters, $MD_i$ analysis were performed. The observation of test data’s $MD_i$ analysis is shown in Table 4. The parameter that exhibited the highest - MD was most likely the faulty parameter, and components related to that parameter are most likely the cause of the faults. The result from Table 4 indicates that fan speed is the most likely faulty component (i.e., the fan was suspected to be faulty). The fan speed $\Delta MD$, followed normal distribution and a 95% confidence range was used to define the threshold boundary for each parameter to isolate faulty parameters, as shown in Table 3.

ΔMD, followed normal distribution and a 95% confidence range was used to define the threshold boundary for each parameter to isolate faulty parameters, as shown in Table 3.

Table 3. Threshold Bound (95% range) on Parameters MD from Training Data

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\Delta MD$ at 2.5%</th>
<th>$\Delta MD$ at 97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1 - Fan Speed</td>
<td>-0.85</td>
<td>0.34</td>
</tr>
<tr>
<td>P2 - CPU temperature</td>
<td>-0.86</td>
<td>0.33</td>
</tr>
<tr>
<td>P3 - Mother board temperature</td>
<td>-0.61</td>
<td>0.28</td>
</tr>
<tr>
<td>P4 - Video card temperature</td>
<td>-0.86</td>
<td>0.34</td>
</tr>
<tr>
<td>P5 - %C2 state</td>
<td>-0.66</td>
<td>0.28</td>
</tr>
<tr>
<td>P6 - %C3 state</td>
<td>-0.67</td>
<td>0.34</td>
</tr>
<tr>
<td>P7 - %CPU usage</td>
<td>-0.58</td>
<td>0.37</td>
</tr>
<tr>
<td>P8 - %CPU throttle</td>
<td>-0.64</td>
<td>0.34</td>
</tr>
</tbody>
</table>
parameter was identified as being faulty approximately 75% of the time, followed by video card temperature for 24% of the time.

Figure 7 – Comparison of MDs of test system with baseline.

CPU-related parameters in the MD analysis fell beyond the threshold bound almost 100% of the time, which suggested that there was a potential problem with the CPU. The analysis of the correlation between CPU-related parameters and other parameters indicated that the CPU parameters were not as correlated to fan or temperature parameters as the fan speed and temperature parameters. The S/N ratio estimation suggested that the parameters related to the CPU were of least importance in MD calculation. These two analyses suggested that CPU parameters were not causing faults and were not picked as the faulty parameters.

Table 4. Observation of Parameters MDs from Test Data.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Observations out of ΔMD₀</th>
<th>Observations highest - ΔMD</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1 - Fan Speed</td>
<td>4544</td>
<td>3654</td>
</tr>
<tr>
<td>P2 - CPU temperature</td>
<td>1300</td>
<td>155</td>
</tr>
<tr>
<td>P3 - Mother board temperature</td>
<td>2112</td>
<td>0</td>
</tr>
<tr>
<td>P4 - Video card temperature</td>
<td>2255</td>
<td>1190</td>
</tr>
<tr>
<td>P5 - %C2 state</td>
<td>4989</td>
<td>0</td>
</tr>
<tr>
<td>P6 - %C3 state</td>
<td>4939</td>
<td>0</td>
</tr>
<tr>
<td>P7 - %CPU usage</td>
<td>4957</td>
<td>1</td>
</tr>
<tr>
<td>P8 - %CPU throttle</td>
<td>4631</td>
<td>0</td>
</tr>
</tbody>
</table>

At the 2,700th data point, the drop in MD value was observed because the computer shutdown by itself and was then restarted manually, which resulted into temporary fan start-up and a drop in three temperatures. However, these temperatures were above the nominal/expected value of the CPU and video card temperatures during that period. Therefore, these temperatures were identified as anomalous as well. Intuitively, if the fan is not functioning well, the temperature of the system would rise and become abnormally high. The manual investigation of the data set also indicated that the fan was not operating and that the three monitored temperatures were higher by ~10°C from the nominal temperature values. Therefore, it can said that the suggested approach identified the faulty parameter of the computer.

4. SUMMARY AND CONCLUSIONS

A new approach for fault isolation utilizing the Mahalanobis distance has been demonstrated. The work expands the applicability of the Mahalanobis distance from detection to fault isolation. In literature, unhealthy system data were used to scale the MD values for fault isolation. In this suggested approach, data from an unhealthy system is not required; rather, a threshold bound based on healthy training data for each parameter is defined. The strength of the MD method is that it preserves all the information available because it does not reduce the dimensionality of the data.

A set of experiments was conducted to establish the “healthy” or “normal” operation on a set of notebook computers subjected to a range of usages and environmental conditions. A field-returned computer was evaluated in-situ using Mahalanobis distance techniques. The Mahalanobis-distance-based faulty parameter identification method was used to identify the fault and key parameters for root cause analysis of anomalies. The suggested fault isolation approach utilized MD and DoE Three critical parameters were identified: fan speed and two temperature components (CPU temperature and video card temperature). The fan speed parameter was identified as being faulty. The results showed that the suggested approach could be used for quick fault detection at a system level. Faulty parameter isolation can also be achieved using the same distance measure.

ACKNOWLEDGEMENTS

This work is sponsored by the members of the CALCE Prognostics and Health Management Consortium at the University of Maryland, College Park.

REFERENCES


**BIOGRAPHY**

**Sachin Kumar** received the B.S. in Metallurgical Engineering from the Bihar Institute of Technology and the M.Tech. in Reliability Engineering from the Indian Institute of Technology, Kharagpur, India. He is currently pursuing the Ph.D. degree in Mechanical Engineering at the University of Maryland, College Park. His research interests include reliability, electronic system prognostics, and health and usage monitoring of electronic systems.

**Eli Dolev** received a B.S. in Mechanical Engineering, an M.S. in Mechanical Engineering, and an MBA from the University of the Negev, Beer Sheva, Israel. He is chief system engineer at the Engineering and Production Department at the Nuclear Research Center–Negev, Israel. He is involved with the research, design, and development of mechanical systems. Previously he was head of the Electronic Packaging Department at the Nuclear Research Center–Negev, and engaged in the development, design and manufacturing of electronic handheld instruments and
systems for radiation detecting and monitoring. He has authored or co-authored over 30 conferences and refereed journal publications. Currently he is a guest researcher at the CALCE Electronic Products and Systems Center at the University of Maryland, College Park where he is involved in evaluating models and algorithms for prognostics and health management.

Michael Pecht has a B.S. in Acoustics, an M.S. in Electrical Engineering and an M.S. and Ph.D. in Engineering Mechanics from the University of Wisconsin at Madison. He is a Professional Engineer, an IEEE Fellow, and an ASME Fellow. He has received the 3M Research Award for electronics packaging, the IEEE Award for chairing key Reliability Standards, and the IMAPS William D. Ashman Memorial Achievement Award for his contributions in electronics reliability analysis. He has written over twenty books on electronic products development, use, and supply chain management. He served as chief editor of the IEEE Transactions on Reliability for eight years and on the advisory board of IEEE Spectrum. He has been the chief editor for Microelectronics Reliability for over eleven years and an associate editor for the IEEE Transactions on Components and Packaging Technology. He is a Chair Professor and the founder of the Center for Advanced Life Cycle Engineering (CALCE) and the Electronic Products and Systems Consortium at the University of Maryland. He has also been leading a research team in the area of prognostics, and formed the Prognostics and Health Management Consortium at the University of Maryland. He has consulted for over 50 major international electronics companies, providing expertise in strategic planning, design, test, prognostics, IP, and risk assessment of electronic products and systems.