Parameter selection for health monitoring of electronic products

Sachin Kumar, Eli Doleva, Michael Pecht

Center for Advanced Life Cycle Engineering (CALCE), University of Maryland, College Park, MD 20742, USA

ABSTRACT

This paper presents an approach for selecting precursor parameters for health monitoring of electronic products. The approach includes failure modes, mechanisms, and effects analysis (FMMEA) and life cycle profile analysis of a product. The criticality of the failure mechanisms is established using a risk priority number (RPN), where the RPN for each failure mechanism is calculated as a product of the occurrence and the severity of each mechanism. Performance parameters that can be associated with the critical failure mechanisms should be selected for health monitoring of the product. These parameters could be used for diagnostic purposes. A case study is presented to demonstrate the parameter selection approach for a computer server system. FMMEA was performed on the server, and precursor parameters of the server were selected for monitoring based on the failure modes and mechanisms that posed the highest risk. The utilization of identified parameters for fault detection is presented through a diagnostic algorithm. This approach can be used to select parameters for health monitoring of any system.

1. Introduction

Electronic products are manufactured by putting together a logical combination of subassemblies, components, and parts. These products fail through various failure mechanisms in their life-cycle environment. The anomalous behavior of an electronic product often prompts a customer to return the product to a retailer or a manufacturer. The product often goes through extensive testing at the retailer's facility or the manufacturer's site. In many cases, these tests fail to recreate the exact anomalies and faults. Hence, the discovery of the root cause behind the anomalous behavior becomes difficult [1]. This fault examination process adds expense to manufacturers in terms of research investment, test equipment, and personnel.

The ability to detect anomalous behavior during a product's operation can free some resources invested in post failure analysis and identify the root cause behind failure events. Health monitoring is one such strategy that evaluates health during operation by measuring and recording the extent of deviation and degradation from a product's normal operating state. The benefits of health monitoring can be reaped in terms of advance warning of failure, prevention of catastrophic failure, reliability assessment, less unscheduled maintenance, fault identification, improved product qualification procedures, and improved design of future products [2]. Scanf et al. [3] have demonstrated that the implementations of prognostic technologies have a positive impact on maintenance costs due to the decrease in downtime and costs a unscheduled maintenance.

Health monitoring of electronic products has been conducted by using canaries to provide advance warning of failure, by monitoring precursors to impending failure (data-driven), and by modeling life-cycle environmental stress to compute accumulated damage (physics-of-failure) [4]. For estimating the remaining useful life of an electronic system, the use of sacrificial (non-functional) devices (i.e., canaries) that experience an operational environment similar to the functional electronic assembly has been suggested [5]. However, it is not always feasible to use canaries due to a product's architectural complexity and insufficient knowledge of its actual application environment. A product's remaining life can be estimated by analyzing its performance parameters, which indicate changes in the product's health status. Selection of necessary and sufficient performance parameters representing the product's operational condition is the key to the health monitoring procedure. In a case study on GPS degradation, a simulation-based approach was used to identify critical components [6]. However, this study did not provide an explanation of the consideration of components' failure modes. The efficiency and accuracy of a diagnostic and prognostic algorithm also depends on the type of data used. A formalized and informed approach for parameter selection based on failure modes and mechanisms would assist in accurate estimation of a product's residual life.

This paper presents a physics-of-failure (PoF) based approach for selecting parameters for product health monitoring. The
approach involves conducting failure modes, mechanisms, and effects analysis (FMMEA) on a product and identifying the parameters that can be used for in situ health monitoring of the product. The selected parameters would assist in detecting failures, identifying failure precursors, and calculating remaining useful life (RUL). These parameters can also assist in data-driven diagnostic approaches that have been built for the automatic testing of electronic systems [7,8]. These parameters can be used as an input to the prognostic models developed for the system based on different failure mechanisms. Failure modes, effects, and criticality analysis (FMECA) has been extended to select sensors for health monitoring and precursor identification to perform condition-based maintenance [9]. By analyzing trends within the performance data, precursors can be derived from one or more parameters that anticipate the changes that a product will experience in time. With the appropriate selection of precursors, one can extract features from the product that can be used to better describe the current and future states of product health.

In the life cycle of a product, several failure mechanisms may be activated by different environmental and operational parameters acting at various stress levels, but in general only a few operational and environmental parameters and failure mechanisms are responsible for the majority of failures. High-priority mechanisms are those with high combinations of occurrence and severity. Prioritization of failure mechanisms provides an opportunity for effective utilization of resources. In the following section, a discussion of the FMMEA approach is presented, followed by a case study on a computer server system.

### 2. Failure modes, mechanisms, and effects analysis (FMMEA)

The traditional failure modes and effects analysis (FMEA) methodology is a procedure used to identify and evaluate the potential failure modes of a product and their effects. FMEA is also used to identify actions that could eliminate or reduce the likelihood of the occurrence of a potential failure [10]. Many organizations within the electronics industry have employed or required the use of FMEA, but in general this methodology has not provided satisfaction, except for the purpose of safety analysis [11]. One limitation of the FMEA procedure is that it does not identify the product failure mechanisms and models in its analysis and reporting process. Failure mechanisms and their related physical models are important for planning tests and screens to audit nominal design and manufacturing specifications, as well as the level of defects introduced by excessive variability in manufacturing and material parameters. Cassanelli et al. [12] suggest that the study of failure mechanism adds value to ordinary FMEA. Failure modes, mechanisms, and effects analysis (FMMEA) methodology overcomes the weaknesses of the traditional FMEA process [11]. FMMEA is a physics-of-failure (PoF) based methodology for assessing the root causes of failures and failure mechanisms of a given product [13]. A schematic diagram showing the steps in FMMEA is shown in Fig. 1 [11].

A potential failure mode is the manner in which a failure can occur—that is, the ways in which the item fails to perform its intended design function, or performs the function but fails to meet its objectives [14,15]. Failure modes are closely related to the functional and performance requirements of the product. Failure causes are due to defects in design, process, quality, or part application, and are the underlying cause of the failures; or they initiate a process that leads to failure and can help to identify the failure mechanism driving the failure mode. Failure mechanisms are the processes by which a specific combination of physical, electrical, chemical, and mechanical stresses induces failures. A failure effect is an effect that a failure has on an entire product or system.

FMMEA prioritizes failure mechanisms based on information about the environmental and operational stress levels that must be considered in system design. In FMMEA, a risk priority number (RPN) is calculated to represent the criticality of each failure mechanism. The RPN is the product of the occurrence and severity of each failure mechanism with the assumption that the failure mechanism is predetermined for a failure mode (i.e., likelihood of detection is 1). Occurrence describes how frequently a failure mechanism is expected to result in failure. Severity describes the seriousness of the effect of the failure caused by a mechanism. FMMEA is based on an understanding of the relationships between the product requirements and the physical characteristics of the product (and their variation in the production process); the interactions of product materials with loads (stresses at application conditions); and their influence on product susceptibility to failure with respect to use conditions. This involves finding the failure mechanisms and reliability models to predict probability of failure.

FMMEA is a major improvement over traditional design for reliability methods since it internalizes the concept of the failure mechanism at every step of the decision-making. Utilization of failure mechanisms as the basis of reliability assessment has been accepted by major technical organizations such as EIA/JEDEC, SEMATECH, and IEEE. Examples of standards that require estimation of the failure mechanism for reliability analysis from JEDEC are given in Refs. [16–21]. SEMATECH publications [22–25] utilize the concepts of failure mechanism–based reliability assessment for semiconductors. IEEE Standard 1413.1 promotes the reliability prediction method of using the load (stress) and damage models that determine when a specific failure mechanism will occur for a product in a given environment [15].

The various failure mechanisms identified for a product in a given environment can be used to calculate the cumulative damage in order to estimate remaining useful life for prognostic purposes. The damage estimation due to the presence of a failure mechanism involves using the geometry and material properties of the product, together with the measured life-cycle loads such as temperature, voltage, etc., acting on the product. Failure mechanisms are categorized as either overstress or wear out. Overstress failure occurs because of a single load (stress) condition that exceeds a fundamental material strength. Wear out failure occurs because of cumulative damage due to loads (stresses) acting over time. The wear out failure mechanisms for electronics that can be modeled and used are summarized in Table 1 [26]. In a situation where parameters associated with a failure model cannot be monitored, one of the two approaches: first, one can establish a correlation between an available or measurable parameter to the parameter.
used in the model, or second, one can establish a model for measurable parameters that represents similar degradation or wear out rate. This analytical model and correlation can be established using an independent experiment on a separate test bed.

The health monitoring process begins with gathering information on a product's failure mechanisms, modes, environmental conditions, and performance parameters that can be monitored. Each monitored parameter is evaluated in terms of its ability to detect the initiation of a failure mode. Parameters that are related to a product's failure modes are chosen for continuous monitoring. These parameters enable application of diagnostic and prognostic methods for product health monitoring and reliability estimation. A failure precursor is an event or series of events that is indicative of an impending failure. A failure is predicted by correlating the change in the monitored precursor parameter and the effect of this change on the operation of a product.

For products with multiple usage conditions, a precursor parameter must be selected such that a change in combined loading is accounted for in the reasoning model. For example, the solder joints of components on an electronic circuit board used under temperature cycling and vibration conditions will fail earlier than when the circuit board is subjected to only temperature cycling or vibration. In such a situation the reasoning model should account for the combined effects of temperature cycling and vibration on the monitored parameter.

### 3. Case study on a computer server

A blade server was analyzed for this study. A blade server is a server chassis housing multiple thin, modular electronic circuit boards, known as server blades. Each blade is a server in its own right, often dedicated to a single application. The blades are literally servers on a card, containing processors, memory, integrated network controllers, an optional fiber channel host bus adapters (HBAs), and other input/output (I/O) ports. These blades share common power supplies and air-cooling resources [27]. This allows more processing power in less rack space, simplifies cabling, and reduces power consumption. The server can hold up to 10 individual blades (see Fig. 2). The backside of the server (Fig. 3) shows the power supplies, fans, I/O devices, and network management devices that are shared by all 10 blades. The individual blades are actually very similar to personal computers.

In general, the phases of a life cycle profile include manufacturing, transportation, operation, and storage. A life cycle profile involves both environmental and operational loads that a system is exposed to throughout its life. These loads may be, for example, temperature and humidity, vibration, shock, power, or corrosion. It is important to understand a system’s life cycle profile to determine the actual loads that will affect the system’s performance and when those loads will occur. Additionally, information about the life cycle profile can be used to eliminate failure modes that may not occur under the given application conditions [28].

For this project, a server manufacturer provided a typical life cycle profile of a server (see Table 2). A life cycle profile for other loads, such as vibration and shock, could also be obtained based on field data, handbooks, and standards. The percentage duration mentioned in Table 2 is based on the total hours spent during that phase of the product’s life cycle.

FMMEA is a tool used during the reliability design review stage by people who are responsible for the design, manufacture, management, and maintenance of the product, since it internalizes the concept of the failure mechanism at every step of the decision-making process. The product of an FMMEA is a table of information that summarizes the analysis of all possible failure modes and mechanisms. A semi-quantitative approach that is used for risk and safety assessment [29,30] is considered to prioritize the issues identified by FMMEA, and all failure modes and mechanisms are quantified based on their occurrences, severity, and detectability.

Severity is a classification category assigned to each failure depending on its effects on equipment and/or system operation. A severity level can be assigned to each failure mechanism based on certain criteria. It can be expressed in terms of economic or...
image loss. Qualitative criteria used to assign severity levels for a server failure on a scale of 1–5 are as follows: server failures resulted in safety-related catastrophic failures = 5; server failed to provide intended service = 4; server performance gradual degrading = 3; server operates at reduced specification = 2; and server suffers minor damage = 1.

Occurrence describes how frequently a failure mode is projected to occur as a result of a specific cause. Like severity levels, occurrence levels can be assigned to each failure mode based on certain criteria. Qualitative criteria used to assign ratings for occurrence of a server failure event on a scale of 1–5 are as follows: server failure is almost inevitable = 5; server suffers repeated failures = 4; server encounters occasional failures = 3; server has relatively few failures = 2; and server failure is unlikely = 1. The numbers used to represent severity and occurrences are application-dependent. For every other mission, these numbers would vary.

Detectability is the ability to identify a failure mode either by inspection or by measurement. In health monitoring an assumption of 100% detectability of a failure mode would not be invalid. Selection of the health monitoring parameters is based on the correlation of measurable quantity to failure mode because of analysis of historical evidence.

To identify a high-risk failure mechanism, a risk matrix can be formulated based on severity and occurrence rating levels. A risk priority number (RPN) is calculated from the product of severity, occurrence, and detection (i.e., = 1). There are three categories of RPN: 0–4 = low risk; 5–14 = moderate risk; and 15–25 = high risk. Likewise, a fan operates on a regulated power supply so probability of the overstress event is also low. A similar analytical approach is taken to assign severity and occurrence level for other components as well.

3.1. Fan

Operating the electrical and mechanical components of a server produces heat, which must be removed to ensure proper function of all these components. Fans are the most common components used to remove heat from servers. One of the failure modes of a fan is when the fan fails to rotate. This failure can occur due to any of the three failure mechanisms mentioned in Table 3. But these are not the only failure mechanisms. The failure of a fan would lead to high temperatures in the server chassis, which, in turn, could cause the server to shut down. This effect is considered to be of high severity, and so the number five is assigned to it in Table 2. The probability of this failure mode due to different mechanisms varies, and this can be obtained from the historical data. In our case study, different numbers are assigned to the occurrences. These numbers are for illustration purposes only.

3.1.2. Hard disk drive

Data storage on a server is provided by hard disk drives. Many storage connection methods such as FireWire, SATA, SCSI, DAS, Fibre Channel, and iSCSI are also used to store data outside the server [37]. One of the failure modes of a hard disk drive is when the hard disk’s head crashes. This failure can occur due to physical shock or wear out. This type of failure can lead to spontaneous (intermittent) failures of the server. Disk drive failure can be identified by observing changes in the current and rotation measurement. Electronic failures usually relate to problems on the controller board of the actual hard disk. A rise in the hard disk drive’s temperature can indicate failure or degradation of electronic or electrical components [38].
3.1.3. Memory card

Every server comes with a certain amount of physical memory, usually referred to as main memory or random access memory (RAM). Most servers have slots for adding and replacing memory sticks. Memory does not spin or move, but due to its fragile nature a failure can have devastating effects. A possible failure mode of a memory card is an electrical short. Failure of a memory card could be due to corrosion on the metal contacts of the memory. This can lead to damaged memory and failure to boot the server. The measurements that can be performed on memory card include its temperature, wattage used, and a memory capacity. The power consumption (wattage) by memory varies widely depending on how frequently it is accessed. In case of any damage, degradation in memory would manifest itself as a change in memory capacity and a temperature rise.

3.1.4. Network card

The network interface card provides networking capability to the server. It can fail in different modes, including no connectivity, packet loss in I/O, and incompatibility with new hardware/software. Electrical failure in a network card can result in a loss of data processing ability and connectivity to the network. The network card becomes unstable when the components on the card are exposed to high temperatures, resulting in loss of memory instructions. Packet loss in I/O can occur due to inappropriate encoding or addressing when information packages are transferred over the network. The network card can also fail due to different I/O formats of different components.

### Table 3

<table>
<thead>
<tr>
<th>Item</th>
<th>Potential failure modes</th>
<th>Potential failure mechanisms</th>
<th>Possible effects</th>
<th>Severity</th>
<th>Occurrence</th>
<th>RPN</th>
<th>Monitoring parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fan [34–36]</td>
<td>Poor cooling</td>
<td>Wear out of fan bearings</td>
<td>Damaged fan and improper cooling of server components</td>
<td>3</td>
<td>5</td>
<td>15</td>
<td>Temperature of server chassis, current, voltage, fan blade RPM</td>
</tr>
<tr>
<td>Failure to rotate</td>
<td>Thermal aging of insulating material in winding</td>
<td>Server heats up and shuts down</td>
<td></td>
<td>5</td>
<td>2</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Hard disk drive [37,38]</td>
<td>Hard drive crash</td>
<td>Mechanical overstress</td>
<td>Loss of data</td>
<td>5</td>
<td>4</td>
<td>20</td>
<td>Temperature of hard disk drive, disk platter RPM, motor current, shock/vibration, power-on hours</td>
</tr>
<tr>
<td>Overheating of hard drive</td>
<td>Electro-migration</td>
<td>Spontaneous (intermittent) failures of server</td>
<td>3</td>
<td>4</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrical surge</td>
<td>Electrical overstress</td>
<td>Knocks out the controller board</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logical error</td>
<td>Not identified</td>
<td>Severely fragmented hard drive</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network card [41,42]</td>
<td>No connectivity</td>
<td>Corrosion of metal contacts</td>
<td>Loss of data to network card and inability to process data</td>
<td>4</td>
<td>3</td>
<td>12</td>
<td>Data throughput, network utilization, current, voltage</td>
</tr>
<tr>
<td>High temperature</td>
<td>Thermal fatigue</td>
<td>Loss in memory instructions</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Packet loss in I/O</td>
<td>Error in encryption</td>
<td>Loss of data</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incompatibility with new hardware</td>
<td>Inaccurate formatting</td>
<td>Cannot connect server to clients</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memory card [39,40]</td>
<td>Electrical short due to impurities</td>
<td>Contamination in memory</td>
<td>Diminishes ability to release heat and stay cool</td>
<td>3</td>
<td>5</td>
<td>15</td>
<td>Memory usage capacity, temperature on memory card</td>
</tr>
<tr>
<td>Failure to boot</td>
<td>Corrosion of metal contacts</td>
<td>Electrostatic discharge</td>
<td>Damaged memory</td>
<td>3</td>
<td>3</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Static electricity</td>
<td>Damage to power supply</td>
<td>Damage to power supply unit</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>Temperature of power supply unit, voltage of power supply unit</td>
<td></td>
</tr>
<tr>
<td>Short circuit</td>
<td>Electrical overstress</td>
<td>No power output</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss of power</td>
<td>Electrical overstress</td>
<td>Excessive power output</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unregulated output</td>
<td>Thermal fatigue</td>
<td>Damage to the CPU</td>
<td>5</td>
<td>2</td>
<td>10</td>
<td>Temperature of CPU die, CPU usage, CPU throttle</td>
<td></td>
</tr>
<tr>
<td>CPU [45]</td>
<td>Power degradation</td>
<td>Voltage degradation</td>
<td>Reduced performance</td>
<td>3</td>
<td>3</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

3.1.3. Memory card

Every server comes with a certain amount of physical memory, usually referred to as main memory or random access memory (RAM). Most servers have slots for adding and replacing memory sticks. Memory does not spin or move, but due to its fragile nature a failure can have devastating effects. A possible failure mode of a memory card is an electrical short. Failure of a memory card could be due to corrosion on the metal contacts of the memory. This can lead to damaged memory and failure to boot the server. The measurements that can be performed on memory card include its temperature, wattage used, and a memory capacity. The power consumption (wattage) by memory varies widely depending on how frequently it is accessed. In case of any damage, degradation in memory would manifest itself as a change in memory capacity and a temperature rise.

3.1.4. Network card

The network interface card provides networking capability to the server. It can fail in different modes, including no connectivity, packet loss in I/O, and incompatibility with new hardware/software. Electrical failure in a network card can result in a loss of data processing ability and connectivity to the network. The network card becomes unstable when the components on the card are exposed to high temperatures, resulting in loss of memory instructions. Packet loss in I/O can occur due to inappropriate encoding or addressing when information packages are transferred over the network. The network card can also fail due to different I/O formats of different components.

3.1.5. Power supply

A power supply function is critical because it is a single power source for all blades within the server enclosure. There are three main failure modes of power supplies: temperature-induced failure, loss of power, and unregulated output power. Temperature-induced failure occurs when the power supply unit experiences an electrical short, causing the temperature in the server chassis to increase. This can lead to a damaged power supply. Electrical overstress can result in power loss to the server. Unregulated output power occurs when there is an internal rectifier or condenser
failure in the power supply due to electrical overstress. This failure can lead to excessive power output.

3.1.6. Central processing unit

The central processing unit (CPU) is considered to be the brains of the computer. The CPU fetches instructions from the memory, decodes their binary contents, and executes them. It references memory and I/O ports as necessary in the execution of instructions. In addition, the CPU also recognizes and responds to certain external control signals [45]. Failures of the CPU are generally very severe, meaning their effects lead to loss of function or system failure. Two basic failure modes of the CPU are burning and power degradation. Burning of the CPU occurs due to thermal fatigue. This can result in significant damage to the CPU. Power degradation of the CPU occurs when the voltage substantially decreases below a rated threshold value. This can lead to reduced performance of the server.

3.2. System parameters and failure precursors to be monitored

The results of the FMMEA of the blade server indicated that three different high-risk failures existed within the server system: rotation failures of the fans; head crashes in the hard disk drives; and electrical shorts of the memory cards. Consequently, these high-risk failures were considered for diagnostics and prognostics. Monitoring parameters and/or failure precursors that are related to these high-risk failures could indicate or provide warnings in advance. Each of these failure modes is associated with one or more measurable variables that, if monitored with a sensor and analyzed using predictive algorithms, could indicate impending failure. Table 3 presents the high-risk failure modes and the corresponding failure precursors that can be monitored by the existing system sensors. The parameters that are monitored are in bold letters in the monitoring parameters column of Table 3 (see also Table 4).

4. Monitoring and data analysis plan

This section discusses how to monitor the precursors identified in the last section. To monitor the failure precursors, one or more sensor devices are required. Fortunately, most current-generation computer systems, including servers, have several built-in sensors. Thermal diodes are typically installed near critical electronic components including the hard disk drive, CPU, and memory card to monitor their temperatures. These and other sensors that exist within the computer system, along with performance parameters such as CPU usage, CPU throttle, and memory usage capacity, can be interrogated via software. This information can be recorded in a log file. Now that the monitoring plan is established, in order to study the evolution of the system responses to the life cycle loading of the data collected is analyzed for diagnosis of the blade servers. Diagnostics pertain to the detection and isolation of faults or failures in a system.

Several measurements are made of a system to understand the system's behavior. It becomes difficult to figure out what is happening with a system when the collected data appears clouded, unclear, and even redundant. This data set can be re-expressed to extract the most meaningful information and filter out the noise. The principal component analysis (PCA) is used for re-expressing the data in lower dimensions such that similarities and differences in a system's performance can be highlighted. The advantage of PCA approaches is that once patterns in the data are identified, the data can be compressed further by reducing the number of dimensions without much loss of information. PCA transforms the data by a linear transformation to a new coordinate system [46]. PCA is useful for analyzing a system where most observed variables are monotonically related to underlying factors and to each other. In addition, PCA allows the use of variables that are not measured in the same units (e.g., temperature, voltage, rpm). PCA, except for special cases, always gives a unique solution, and it can be used for both continuous and binary data [47].

In this study, PCA is preferred for purposes of data reduction, since the measurements have more than three dimensions with different units. Although individual parameters may remain below their respective specification limits, a system can behave anomalously due to changes in correlation among parameters. These changes should be identified before any failure occurs. The PCA can identify the changes in parameters, since it takes a cloud of data points and rotates it such that the maximum variability is visible. It identifies the most important gradients, which is necessary for prioritizing the corrective action. The singular values derived from the PCA can be used to perform system diagnostics. The component of the vector with the highest variance value is the component that describes most of the changes in a multivariate system.

Based on the FMMEA of the server, five critical monitoring parameters were chosen. With this data, diagnostics can be performed by considering PCA in an iterative fashion while analyzing in situ system data. The multivariate data is reduced to a vector, and principal parameters are identified. These principal components represent more than 95% of variations in the original parameters. With each new record, a new PCA will be performed to find a principal parameter. If the principal parameter changes to a new parameter, then the algorithm identifies an anomaly and prompts a further investigation into the health of the system.

In order to observe the effect of new observations and to have sufficient historical data for PCA, an overlapping windowing technique is used. The window number can be used to partition the entire data set into overlapping sections, where the overlap ratio is 0.8. This approach provides real-time analysis of collected data. The data set from a server was imported and five monitoring parameters were selected as T1, T2, T3, MU, and VIB, which represent the temperatures at the hard disk drive, server chassis, and memory card; memory usage capacity; and vibration at the hard drive, respectively. The first and second principal components were computed using the resulting eigenvector from the singular value decomposition, and the original data set was plotted against the two principal axes.

From Fig. 4 it can be seen that the first principal component explains more variation in the original data set. Specifically, it accounts for a data range from −3 to 7, whereas the second component accounts for a variation with a range from −1 to 2.5. This was the first indication that the first component could be chosen as the data direction. The data direction is the direction in which the data changes the most. A histogram of the three principal components was also assessed (Fig. 5), and it can be seen that 90% of the variation can be explained by the first two principal components, and 70% of the variation can be explained by the first principal. The curve (Fig. 5) represents cumulative percentage variance explained.

An explained variance plot (Fig. 6, '+', T1, 'o', T2, and '□', T3) displays the health status to flag changes in system dynamics in

| Table 4 |
|-----------------|-----------------|
| **High-risk failures** | **Precursors to be monitored** |
| Rotation failures of the fan | Temperature of server chassis |
| Head crashes in the hard disk drive | Temperature of hard disk drive, shock/vibration |
| Electrical shorts on the memory card | Memory usage capacity, temperature of memory card |

real time and simultaneously identify the critical parameters. A flag is raised when a change is observed in system dynamics. Notice the change that occurs before the 100th window period. The principal parameters of the system change from 1 and 2 to 3 and 1. The observations after the 100th observation are similar to the initial stage of analysis because few observations from the beginning are used for completing the analysis cycle.

The approach discussed in this paper takes advantage of sensors embedded in the chips for health monitoring, but it is not limited to the use of embedded sensors. In the absence of embedded sensors, an additional sensor system can be deployed to collect information on parameters of interest. This paper lists parameters that are readily available in computer systems and that can be correlated to the available physical models for different failure mechanisms. The paper uses the data-driven approach for identifying potential problems in a system. Associating these parameters with the physics-of-failure model would be the next step so that cumulative damage can be estimated. For estimating cumulative damage, a state-based model can also be used in the absence of physical models.

The health monitoring of a system provides different types of parameter measurements. Some parameters vary with each observation and some parameters do not. The physical interpretation of each parameter, in the context of system failure, is needed to apply damage estimation techniques. The data analysis approach should take into account the data type to be analyzed; a simple if–else logic sequence can be suitable for analyzing parameters that do not vary over time, whereas PCA type analysis is required for parameters that vary overtime and exhibit correlation among themselves. For health monitoring, the parameters that exhibit variations at every observation are of more interest, because a trend in system performance can be established, and a pro-active decision can be made to improve the system’s availability. PCA is best suited for parameters that exhibit variation, and it will assist in identifying the right parameters that can be used for quantifying a system’s health overtime. Nevertheless, in our approach, any parameter that shoots up at anytime would be identified as the parameter contributing the most and would be flagged. This allows the detection of sudden changes in any parameter that might represent an overstress condition.

5. Conclusions

This paper presents an approach to select precursor parameters for health monitoring of electronic products. The approach begins with failure modes, mechanisms, and effect analysis (FMMEA) of a product under consideration. A discussion on selecting performance parameters that are associated with a product’s critical failure modes and mechanisms is presented. A principal component analysis (PCA) approach is presented to identify a few of the most varying parameters from several available parameters to reduce the dimensionality of a problem. The approach is illustrated using a computer server system. Based on the historical data, the severity and occurrences of each failure mode and mechanism are defined. These failure modes and mechanisms are ranked in order to select a critical few. The precursor parameters that are associated with critical failure modes and mechanisms are identified for health monitoring of the computer server. The results from the FMMEA analysis on the computer server are bridged with a multivariate technique for analyzing and diagnosing the server’s health. A multivariate statistical technique, principal component analysis, is employed to identify the changes in the server’s performance parameters. It also is a good way to identify the most influential parameters of the server. This parameter selection and isolation methodology can be extended to any product’s health monitoring scheme.
Acknowledgments

The members of the Prognostics and Health Management Consortium, CALCE, University of Maryland, College Park and The Prognostics and Health Management Center at City University of Hong Kong supported this work. The authors would like to thank Mark Zimmerman of CALCE for copyingedit this paper.

References