International Journal of Engineering and Technology Innovation, vol. 8, no. 3, 2018, pp. 191-199

Sensor Data Abstraction for Failure Prediction of Polymerase Chain Reaction Thermal Cyclers

Chan-Young Park, Mi-So Lee, Yu-Seop Kim, Hye-Jeong Song, and Jong-Dae Kim*

Department of Convergence Software, Hallym University, Korea

Bio-IT Research Center, Hallym University, Korea

Received 18 July 2017; received in revised form 20 November 2017; accepted 26 November 2017

Abstract

In this paper, the heating and cooling rates of polymerase chain reaction thermal cyclers was analysed over time, to predict their aging. Two kinds of methods were applied to calculate the rise times of the heating and cooling protocol sections, which were inversely related to the heating or cooling rates. The temporal changes of the rise times were investigated over several months. For three thermal cyclers with different structures, the increases of the rise times were fairly linear; therefore, the aging prediction is feasible.

Keywords: PCR thermal cycler, maintenance, aging, failure

1. Introduction

Equipment management is performed using corrective maintenance (CM) and periodic preventive maintenance (PM) before the failure [1-7]. A more sophisticated management approach to lower equipment maintenance costs is to predict when a failure occurs and to delay maintenance [6]. Equipment producers must balance user satisfaction and cost by designing equipment quality and reliability [7].

The development of the Internet of thing (IoT) technology allows, even low-cost equipment, to send sensor data to the cloud, and it changes the paradigm of equipment management [7]. It is possible to perform health monitoring and prognosis of the equipment by gathering and analyzing data in the cloud with the health information of equipment, such as equipment environment, equipment operation, and maintenance data [2]. The development of this technology will enable, even the costliest maintenance, to be carried out at a manufacturer's office, and automatic inspection will also be possible through the development of big data analysis technology.

With the development of cloud computing and data storage, a large amount of raw sensor data can be stored in the cloud. However, equipment can generate data from more sensors, as they have more and more functions, through the development of microelectromechanical systems and nanotechnology. In addition, the number of IoTs, real-time data, and data size of each sensor node are increasing rapidly [8]. For these reasons, it will become necessary to abstract information from sensor data in order to reduce the amount of data as much as possible. One more reason for abstraction is that it is more useful to increase the number and types of sensors that one can collect from than to increase the amount of data from each sensor to get valuable information. In any case, the complete abstraction of sensor data is worth investigating. This paper attempts to extract information on aging from the sensor data of a typical biochemical polymerase chain reaction (PCR) thermal cycler.

* Corresponding author. Email address: kimjd@hallym.ac.kr

PCR is a molecular biology technique that replicates and amplifies a desired portion of DNA. This technology is used in many areas, such as medical, criminal, investigation, and classification of biological organisms. PCR techniques can selectively amplify fragments of specific DNA in a very small amount of DNA solution in order to obtain genetic information. In addition, the time required for amplification is short, the experimental procedure is simple, and it can be amplified by a fully automatic machine; thus, it plays an important role in the handling of DNA [9].

In PCR, two strands of DNA are separated by the application of denaturation, via heat, and at a low temperature; the primer is annealed to the sequential terminal for amplification, and at slightly higher temperature, DNA is synthesized—this process is called polymerization or extension. The temperature of the PCR reagent is usually adjusted by adjusting the temperature of an aluminum block, where a tube containing the reagent is inserted. Several methods have been developed for controlling the temperature of the aluminum block, but the thermoelectric module (TEM) is usually employed. The performance of the PCR thermal cycler is most influenced by the temperature accuracy for denaturation, annealing, and extension [10-11]. However, when the TEM is aged, the heating and cooling rates of the aluminum block fall and eventually fail [12-14]. In particular, as in the target application of this paper, it is reported that TEM is aged more rapidly when it is utilized for thermal cycling [12]. In addition, deterioration of the heating and cooling rate is directly related to the performance of the PCR thermocycler because the overall protocol execution time is prolonged. From the above, it is essential to measure the heating and cooling rate of the aluminum block.

In this paper, data on the temperature sensor, installed to measure the temperature of the aluminum block, were stored and analyzed for several months. From the stored temperature data, the rise and fall intervals are extracted, and the rise time, which is inversely proportional to the temperature ramping speed, is calculated for each interval. Because rise time is directly related to aging or failure and the performance of the PCR thermal cycler, there is much room for data abstraction to anticipate the aging of the PCR thermal cycler.

In general, the rise time is calculated from the difference between 10% and 90% of the temperature change [15-16]. However, measuring time can cause jitter for a variety of reasons, and the measured temperature is exposed to inherent noise. These two variables can cause the estimated rise time to be inaccurate. As one of the more accurate measurement methods, the width of the differential function of the temperature change function has been proposed [15]. In this paper, the rise time measured by these two measurement methods was analyzed statistically and the appropriate method was found out. The selected method was used to determine the rise time of three different types of PCR thermal cycler with different versions or adapted components. As a result of follow-up observation of the rise time change according to the number of times of use, it is confirmed that it increases linearly. From the experimental results, we concluded that the rise time extracted by the method presented in this paper could be used to measure the aging of the device.

2. Materials and Methods

2.1. Rise time calculation

In general, the rise time of a step response is calculated as the difference between the start and end times of each interval; specifically, the time to reach 10% and 90% of the response change. However, since there is fluctuation due to noise at the measurement temperature and measurement point, it can be estimated using the width of the derivative function of the step response, as shown in Fig. 1. The width of the function is obtained by its second order moment, as follows [15]:

$$t_r = \sqrt{\frac{\int (t-m)^2 s'(t)dt}{\int s'(t)dt}}, \quad m = \frac{\int t s'(t)dt}{\int s'(t)dt} \tag{1}$$

where t_r and $s'(t)$ are the rise time and the derivative of the step response $s(t)$, respectively.

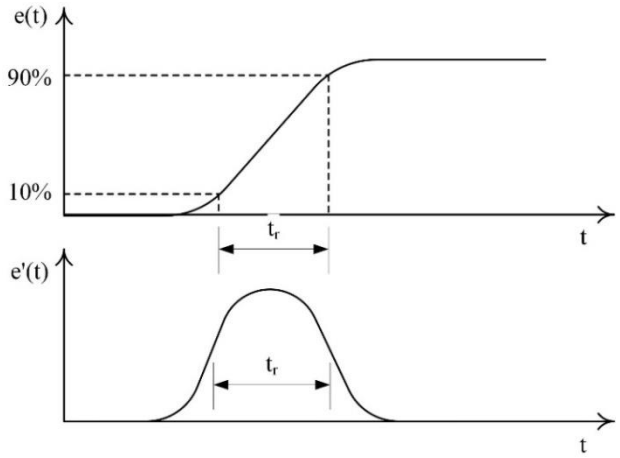

Fig. 1 Rise time estimation by the width of the derivative function of the step response

The two methods of finding the rise time are named as 'direct difference' and 'derivative width', respectively, hereafter. If the step response is analytic functions such as the exponential function or the error function described in the reference [15], the ratio of the rise times from the two methods will be constant regardless of the response time and it can be obtained analytically. For example, let the step response and its derivative are exponential functions with a time constant τ as follows:

$$s(t) = 1 - e^{-t/\tau}, \; s'(t) = \frac{1}{\tau} e^{-t/\tau} \tag{2}$$

Then, the width of the derivative function can be easily obtained by using the following definite integral equation;

$$\int_0^\infty t^n e^{-t/\tau} dt = n! \tau^{n+1} \tag{3}$$

And the width of the derivative function, t_r, is equal to the time constant τ.

$$t_r = \tau \tag{4}$$

On the other hand, the time from 10% to 90% of the step response is calculated from the following equations:

$$0.1 = 1 - e^{-t/\tau}, \; 0.9 = 1 - e^{-t/\tau} \tag{5}$$

The rise time t_r' is represented as follows:

$$t_r' = (\ln 0.1 - \ln 0.9)\tau \cong 9.1\tau = 9.1 t_r \tag{6}$$

That is, the rise time is about 9 times the width of the derivative function. This is true regardless of the magnitude or the time constant of the step response. However, in the thermal cycler, it is not clear whether the ratio is constant since the step response is controlled by a proportional-integral-derivative (PID) controller. Furthermore, it is difficult to obtain the ratio analytically because of the non-linearity of the PID control. In the result section of this paper, both methods are obtained by using MATLAB and statistically analyzed to show that the ratio is constant.

2.2. Experimental setup

Fig. 2 shows the thermal cyclers tested in this study. Both devices are proven products from the same company [16-17]. Both devices have similar with structure and the MyPCRv2 is the newer version. Fig. 3 shows their block diagram. The upper surface of the TEM module becomes a heat absorbing surface when cooling, causing the temperature of the aluminum block to cool down and vice versa under heating. The tube with the reagent is put into the aluminum block, and the aluminum lead is

pressed on with the aluminum lead above 100 degrees to prevent the evaporation of the solution in the tube. Three interchangeable thermistors with standard accuracy were inserted to measure the temperature of the aluminum leads, heat block and heat sink. In this paper, to observe the temperature of an aluminum block, the middle thermistor was monitored.

Both devices have a similar structure, but the heat dissipation structure is different, and the MyPCRv2 is the newer version. In this paper, we analyzed the temperature regulation speed for several months, for three devices: two for the MyPCRv2, of which the Peltier's are different from each other, and one for the MyPCRv1. Implementation of high performance GUI functions in embedded systems requires considerable development time as well as maintenance cost when compared with the implementation of basic functions. Therefore, the PCR thermal cyclers adopted in this paper have host-local structure and PC's as the host. The local system controls the temperature of the aluminum heating block and the lid heater, according to the host's commands. The PCR protocol is managed and processed using a host PC to reduce the time, cost, and maintenance cost required for product development. Several of these products can be controlled simultaneously using a PC application with a powerful graphical user interface (GUI) environment. Therefore we could easily create an independent program to handle the rise time.

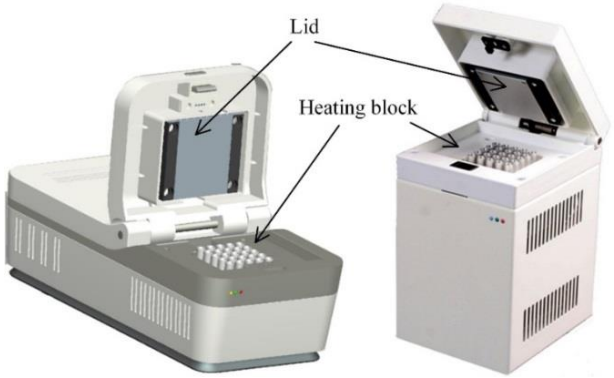

Fig. 2 Employed PCR thermal cyclers (MyPCRv2 (left) and MyPCRv1 (right), LabGenomics, Co., Ltd.)

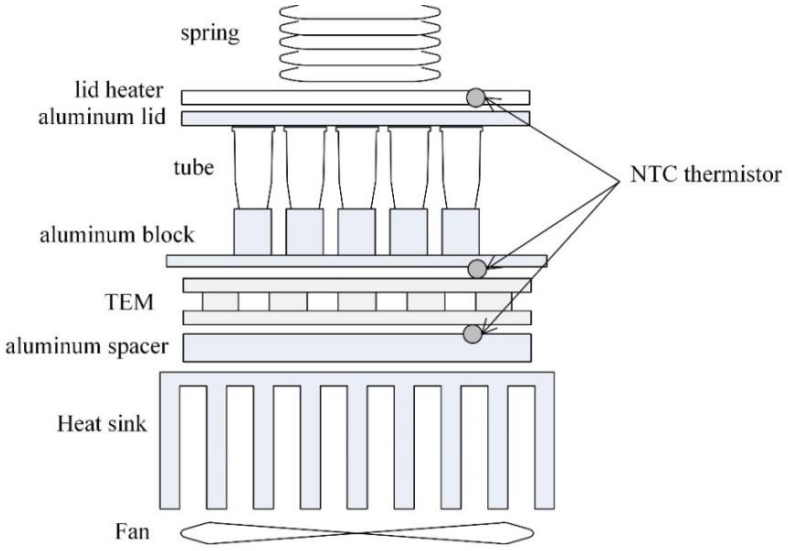

Fig. 3 The block diagram of the internal structure of the PCR thermal cyclers

Table 1 is a typical example of a PCR protocol. The first row labeled '1' in the table is the preheating step to keep the solution at 95 ° C for 30 seconds to activate the enzyme before normal cycling. The next three rows form one cycle of amplification, each representing denaturation, annealing and polymerization. The 'GOTO' label is for flow control and represents the cycle returning to the row labeled '2' for 34 times. This example, therefore, consists of a total of 35 cycles of amplification steps. The last row contains the extra extension step to finalize the PCR. This step allows unbonded strands to combine with comp entary strands.

Table 1 Typical protocol example

Label	Temperature (℃)	Duration (sec)
1	95	30
2	95	30
3	55	30
4	72	30
GOTO	2	34
5	72	180

After one typical protocol was completed, the temperature of the aluminum block was recorded by performing the protocol in Table 2 which was simplified from the original protocol to calculate the rise time. The original system records the thermistor temperature every second, but the resolution is poor to analyze the rise time. Therefore, the temperature data was saved every 100ms by separately creating the program for this purpose. The protocol reduced to have only one cycle and the stay time was reduced to 10 seconds.

Table 2 The protocol for rise time calculation

Label	Temperature (℃)	Duration (s)
1	50	10
2	95	10
3	50	10
4	72	10
5	95	10
6	50	10

Fig. 4 shows the temperature profile obtained by performing the protocol shown in Table 2. As mentioned earlier, the temperature obtained from the thermistor located on the TEM top surface in Fig. 3 was stored at intervals of 100ms and the temperature was plotted using MATLAB. The temperature change intervals from the left were labeled as '1' to '4'. Since the '3' and '4' intervals were also the heating intervals as '1', only the '1' interval was analyzed to obtain the rise time of the heating, and the rise time of the cooling was obtained for the interval '2'. When analyzing in the cloud, it would be better to upload the rise time of all the sections together with the interval information to the cloud together. This is because different PCR protocols may be used depending on the place and time. The heating and cooling rise times were calculated from the 10% to the 90% arrival times of the protocol temperature change. In the figure, the 10% and 90% arrival points of each interval are marked with 'O' and 'Δ', respectively.

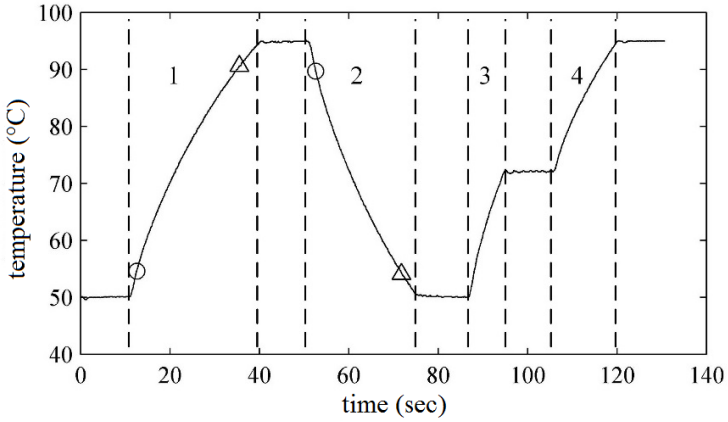

Fig. 4 Intervals of the protocol for rise time calculation

As the equipment ages, it is expected that the rate of heating and cooling will decrease; thus, the rise time is expected to increase. The method described above was employed to collect data for a total of eight months and used all the data up to the date when the device was judged abnormal. The increase rates of the rise times for each device were investigated using linear regression. If the coefficient of determination of the regression is close to 1, aging can be predicted using the method proposed in this paper.

3. Result

Three thermal cyclers were named as: PCR-A (MyPCRv2, the old version of Peltier), PCR-B (MyPCRv2, the new version of Peltier), and PCR-C (MyPCRv1). The number of experiments for PCR-A, PCR-B, and PCR-C was 10 for 74 days, 13 for 95 days, and 9 for 60 days.

To compare the rise time calculation method, rise time was calculated in two methods for all of the intervals, devices and experiments. Table 3 shows the comparison of two rise time calculation methods of the heating interval for the PCR-A device. The fourth row shows the ratio of the 'direct difference' result of the second row to the 'derivative width' of the third row. The fifth row labeled 'error' shows the difference between the 'direct difference' result and the 'derivative width' result multiplied by the average ratio shown in the 'ratio' row of the last column. The last row showed a relative error to the 'direct difference' value. Since the relative error was smaller than 1.83%, the value from the 'derivative width' multiplied by the average ratio was compatible to that by the 'direct difference' method. Figure 5 is a graphical representation of this situation. In the figure, the 'direct difference' result, marked with the 'Δ' mark, overlaps the result of the 'derivative width' multiplied the mean ratio, denoted by the '+' mark.

Table 3 Comparison of two rise time calculation methods

Time (day)	0	7	14	21	29	35	42	56	66	74	Mean
Direct difference	23.56	25.44	26.59	28.61	29.12	30.14	28.14	36.82	40.87	45.41	
Derivative width	6.80	7.34	7.64	8.25	8.38	8.69	8.10	10.79	11.61	12.90	
Ratio	3.47	3.47	3.48	3.47	3.47	3.47	3.47	3.41	3.52	3.52	3.48
Error	0.07	0.07	0.05	0.06	0.00	0.07	0.00	0.68	0.53	0.58	0.21
Relative error (%)	0.28	0.27	0.19	0.20	0.00	0.23	0.01	1.83	1.29	1.27	0.56

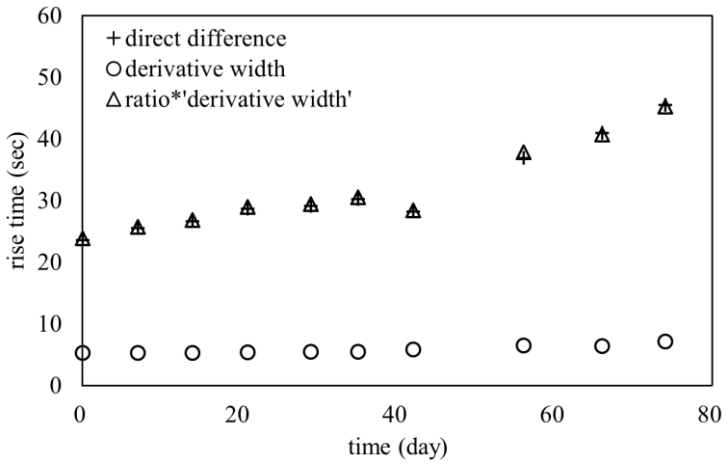

Fig. 5 Relation of two rise time calculation methods

Table 4 summarizes the statistical characteristics of the rise time ratios obtained by the two methods described above. The mean of the ratios is 3.14~ 3.47, and the relative standard deviation, ie, coefficient of variation (CV), was less than 0.88%. This means that the ratio of rise time obtained by both methods is very constant. The last column in Table 4 is the mean, standard deviation, and CV for all data, regardless of device or heating cooling. The mean ratio was 3.45 and the CV was 0.83%, so it was concluded that either method could be used. Therefore, when the rise time was calculated using the 'derivative width' method, the calculated value was multiplied by 3.45, that is the average ratio all over the intervals, devices, and the experiments. As shown in the table, the rise times calculated using both methods were almost the same; therefore, all remaining results were presented using the 'direct difference' method. The relatively simple 'direct difference' method provides almost the same rise time as the statistical method to reduce the influence of noise. This result is considered to be possible because the temperature measurement error of the devices used in the experiment is sufficiently small.

Table 4 Statistics of ratio of rise times calculated by two methods

	PCR-A		PCR-B		PCR-C		Total
Ratio statistics	heating	cooling	heating	cooling	heating	cooling	
Mean	3.48	3.45	3.47	3.45	3.45	3.41	3.45
std	0.03	0.03	0.02	0.02	0.02	0.01	0.03
cv (%)	0.88	0.73	0.58	0.48	0.57	0.38	0.83

Fig. 5 shows the increasing rate trend of the rise times of the heating and cooling intervals. As shown in Fig. 5, heating performance decreased more rapidly and this trend was common to all three devices.

Figs. 6-8 show the heating and cooling rise time changes of the devices. The rise time increases were fairy linear for all of the devices regardless of heating or cooling. Table 5 summarizes the slope and coefficient of determination (R^2) of the trend lines. R^2 was more than 0.82, and, notably, was more than 0.90 for the heating interval. The increasing rate of the rise time was 1.02~1.70 for heating and 0.39~0.46 for cooling. The increasing trend in rise time in the table is very linear and shows that the change in rise time for heating is greater. It has also been found that the growth rates are similar, even if the structure or components of the equipment are different.

The performance of the device is directly related to the DNA amplification performance, but the longer the rise time, the longer the PCR execution time. Therefore, the increase in the rise time itself is also closely related to the deterioration or failure of the device. Therefore, it is necessary to observe the rise time of the heating section more closely.

Table 5 The slope of the increasing rate

		PCR-A	PCR-B	PCR-C
heating	slope	1.1216	1.7032	1.0150
	R^2	0.9000	0.9489	0.9497
cooling	slope	0.4591	0.3915	0.4092
	R^2	0.8591	0.9402	0.8164

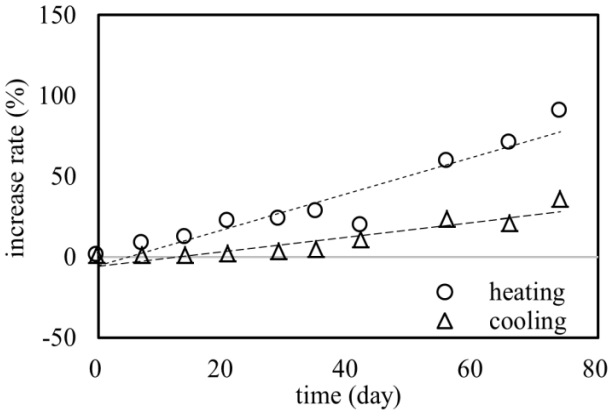

Fig. 6 Rise time increase rate of PCR-A

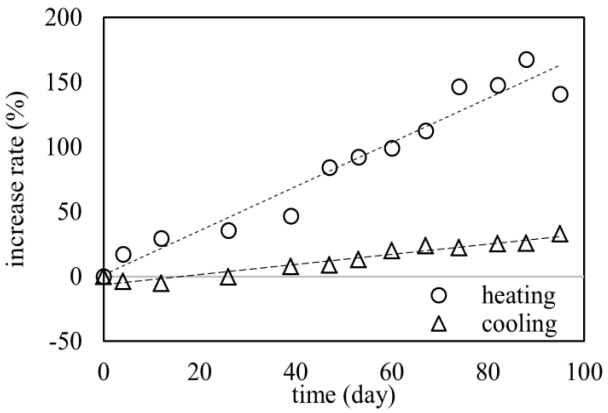

Fig. 7 Rise time increase rate of PCR-B

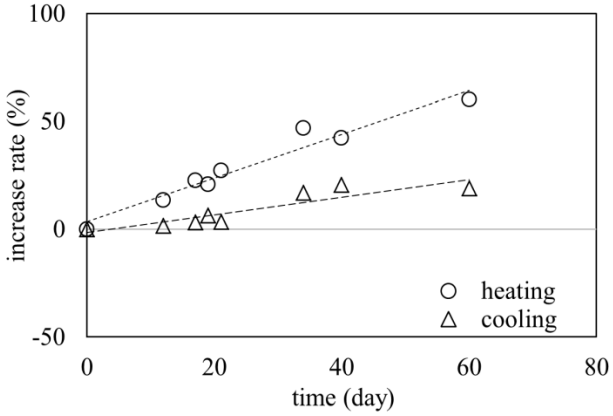

Fig. 8 Rise time increase rate of PCR-C

4. Conclusions

In this paper, the heating and cooling times of PCR thermal cyclers were analyzed over time, to predict the aging of the PCR thermal cyclers. In the temperature profile, the heating and cooling sections were separated and the rise times of each section were measured and tracked for several months. A simple rise time calculation method and a noise-resistant statistical method were compared considering temperature measurement noise. Experimental results, however, showed that simple method delivered the compatible results. Experimental results for rise time showed that all rise times increased linearly regardless of heating and cooling, and the rise time of the heating interval increased more rapidly. These results suggested that the rise time could be used to predict the aging and failure of the PCR thermal cycler and the heating rise time should be monitored more carefully.

When all the devices were connected to the cloud, the sensor data abstraction method proposed in this paper could reduce the amount of data and the cloud is able to accommodate more equipment and sensors. However, caution is required because the details of the raw data will be lost.

In this paper, the aging and failure of equipment was able to be predicted by extracting information from the temperature measurement data of embedded sensors. The same data abstraction methodology can be applied to sensor data added for the purpose of equipment.

Acknowledgements

This research was supported by the Leading Human Resource Training Program of Regional Neo Industry, through the National Research Foundation of Korea (NRF), and was funded by the Ministry of Science, ICT, and Future Planning (NRF-2016H1D5A1909654).

References

[1] D. D. Reitze, "Using cloud computing to enhance automatic test equipment testing and maintenance capabilities," 2013 IEEE AUTOTESTCON, IEEE Press, October 2013, pp. 1-6.
[2] D. Galar, A. Thaduri, M. Catelani, and L. Ciani, "Context awareness for maintenance decision making: A diagnosis and prognosis approach," Measurement, vol. 67, pp. 137-150, 2015.
[3] W. Wang, "An overview of the recent advances in delay-time-based maintenance modelling," Reliability Engineering & System Safety, vol. 106, pp. 165-178, May 2012.
[4] X. Jin, Y. Sun, Z. Que, Y. Wang, and T. W. Chow, "Anomaly detection and fault prognosis for bearings," IEEE Transactions on Instrumentation and Measurement, vol. 65, no. 9, pp. 2046-2054, September 2016.
[5] S. O. Guclu, E. U. Warriach, T. Ozcelebi, and J. J. Lukkien, "Improving failure prediction accuracy in smart environments," 2016 IEEE International Conf. Consumer Electronics, IEEE Press, March 2016, pp. 317-318.
[6] V. H. Coria, S. Maximov, F. Rivas-Dávalos, C. L. Melchor, and J. L. Guardado, "Analytical method for optimization of maintenance policy based on available system failure data," Reliability Engineering & System Safety, vol. 135, pp. 55-63, March 2015.
[7] D. Kwon, M. R. Hodkiewicz, J. Fan, T. Shibutani, and M. G. Pecht, "IoT-based prognostics and systems health management for industrial applications," IEEE Access, vol. 4, pp. 3659-3670, 2016.
[8] A. Kankanhalli, J. Hahn, S. Tan, and G. Gao, "Big data and analytics in healthcare: introduction to the special section," Information Systems Frontiers, vol. 18, no. 2, pp. 233-235, April 2016.
[9] P. Singleton, DNA methods in clinical microbiology, New York: Springer Science & Business Media, 2013.
[10] Y. H. Kim, I. Yang, Y. S. Bae, and S. R. Park, "Performance evaluation of thermal cyclers for PCR in a rapid cycling condition," Biotechniques, vol. 44, no. 4, pp. 495-496, 498, 500, and passim, April 2008.
[11] G. C. Saunders, J. Dukes, H. C. Parkes, and J. H. Cornett, "Interlaboratory study on thermal cycler performance in controlled PCR and random amplified polymorphic DNA analyses," Clinical Chemistry, vol. 47, no. 1, pp. 47-55, January 2001.

[12] M. T. Barako, W. Park, A. M. Marconnet, M. Asheghi, and K. E. Goodson, "Thermal cycling, mechanical degradation, and the effective figure of merit of a thermoelectric module," Journal of Electronic Materials, vol. 42, no. 3, pp. 372-381, March 2013.

[13] E. Hatzikraniotis, K. T. Zorbas, I. Samaras, T. H. Kyratsi, and K. M. Paraskevopoulos, "Efficiency study of a commercial thermoelectric power generator (TEG) under thermal cycling," Journal of Electronic Materials, vol. 39, no. 9, pp. 2112-2116, September 2010.

[14] H. C. R. L. Tenorio, D. A. Vieira, C. P. de Souza, E. C. T. de Macêdo, and R. C. S. Freire, "A thermoelectric module thermal-cycling testing platform with automated measurement capabilities," IEEE International 2016 of the Instrumentation and Measurement Technology Conference Proceedings (I2MTC), IEEE Press, 2016.

[15] W. C. Elmore, "The transient response of damped linear networks with particular regard to wideband amplifiers," Journal of Applied Physics, vol. 19, no. 55, pp. 55-63, 1948.

[16] C. Y. Park, J. D. Kim, Y. S. Kim, H. J. Song, J. M. Kim, and J. Kim, "Cost reduction of PCR thermal cycler," International Journal of Multimedia and Ubiquitous Engineering, vol. 7, no. 2, pp. 389-39, 2012.

[17] C. Y. Park, Y. H. Park, Y. S. Kim, H. J. Song, and J. D. Kim, "Performance evaluation of cost-optimized thermal cycler," Technology and Health Care, vol. 24, no. s1, pp. S179-S185, 2015.