Journal of KIIT. Vol. 15, No. 11, pp. 175-185, Nov. 30, 2017, pISSN 1598-8619, eISSN 2093-7571 175
http://dx.doi.org/10.14801/jkiit.2017.15.11.175

개선된 선박 엔진의 고장 예측 알고리즘

박준현*[1], 오은경*[2], 장민국**, 서영우***, 허성우****

Improved Forecasting Algorithm for Vessel Engine Failure

Jun-Hyun Park*[1], Eun-Kyung Oh*[2], Min-Kook Jang**, Young-Woo Seo***, and Sung-Woo Hur****

본 연구는 "과제번호 S0807-17-1021 지역 SW융합제품 상용화 지원 사업"의 지원을 받아 수행된 것임.

요 약

대형 선박의 자동화 및 고속화로 인하여 선박 운행 일정이 증가하고, 이에 따라 정박 시간이 짧아짐으로 이로 인해 정비시간의 부족현상이 나타나게 되었다. 이에 선박 고장의 확률이 증가하게 된다. 해상에서 선박 고장은 여러 가지 측면에서 심각한 결과를 초래하므로 이에 고장을 사전에 예측할 수 있는 기법이 필요하다. 이전 연구에서 제안한 VFF(Vessel Failure Forecasting) 알고리즘에서는 이러한 요구를 위해 선형회귀법을 이용하여 실시간으로 선박의 결함을 예측하도록 설계하였다. 본 논문에서는 오류 검출의 정확성을 향상시키기 위하여 VFF 알고리즘을 보완하여 데이터 정화, 자동 예측 구간 설정, 선형 회귀법을 개선한 스플라인 선형회귀법, 슬라이딩 윈도우 및 사용자 중심의 시계열 그래프를 제안한다. 실제 운항 중인 H해운사의 선박에서 수집된 데이터와 고장 보고서에 근거한 실험에서 개선된 알고리즘이 VFF 알고리즘보다 효과적인 결과를 실험적으로 입증하였다.

Abstract

Due to the automation and speeding up of large vessels, the vessel's operation schedule has been increased, and as a result, the docking time has become shorter, resulting in a shortage of maintenance time. This increases the probability of vessel failure. Vessel failure at a ocean causes serious consequences in many respects, so techniques that can predict the failure is necessary. A system to predict vessel failure using linear regression in real time has been suggested in previous paper. In this paper, we propose techniques for data-cleansing, automatic prediction interval setting, spline linear regression, sliding window, and for showing user-oriented time series graph to improve the accuracy of error detection. Results obtained by simulation using the real data and the failure report generated by a vessel, which belongs to H marine company, prove that the improved algorithm is more effective in predicting engine failure in advance compared to the methods used in previous paper.

Keywords
algorithm, correlation, regression, vessel engine failure prediction, big data

* 동아대학교
- ORCID[1]: http://orcid.org/0000-0003-1444-0982
- ORCID[2]: http://orcid.org/0000-0002-8539-3514
** 지마린서비스
- ORCID: http://orcid.org/0000-0002-6705-8501
*** 투그램시스템즈
- ORCID: http://orcid.org/0000-0002-2184-1058
*** 동아대학교 컴퓨터공학과 교수(교신저자)
- ORCID: http://orcid.org/0000-0002-9701-4008

· Received: Oct. 20, 2017, Revised: Nov. 15, 2017, Accepted: Nov. 18, 2017
· Corresponding Author: Sung-Woo Hur
Dept. of Computer Engineering, Donga University
University, Donga-Univ, Hadan 2-doong, Saha-gu, Busan, Republic of Korea
Tel.: +82-51-200-7776, Email: swhur@dau.ac.kr

Ⅰ. 서 론

최근 선박 발전 추이를 보면 대형화, 지능화, 자동화가 급속도로 가속화되어 승조원 수가 감소하게 되었고 승조원의 실수에 의한 사고가 증가했다. 선박 운항 일정이 많아짐에 따라 정박할 수 있는 시간이 짧아지고 이로 인한 정비시간의 부족현상도 나타나게 되었다. 또한, 글로벌 해상물동량과 선박 보유량이 매년 지속해서 증가하고 있으며 이에 따른 선박 관리 산업의 역할이 중요해지고 있다. 따라서 실시간으로 선박의 결함을 예측하여 예방할 수 있는 시스템의 개발이 필수적이다[1]-[5].

[6]에서는 선형회귀법을 이용하여 실시간으로 선박의 결함을 예측하도록 설계하였다. 본 논문에서는 오류 검출의 정밀성과 정확성을 향상시키기 위하여 [6]의 기법을 보완하여 데이터 정화, 자동 예측 구간 설정, 선형 회귀법을 개선한 스플라인 선형회귀법, 슬라이딩 윈도우 및 사용자 중심의 시계열 그래프를 제안한다. 개선된 내용을 정리하면 다음과 같다.

첫째, 개선된 데이터 정화방법으로 [6]에서는 데이터 정화는 특정 시간에 하나이상의 센서의 데이터가 누락이 되면 해당 시간의 데이터는 사용하지 않았다. 개선된 알고리즘에서는 통신 또는 센서의 문제로 인해 변형된 데이터를 효과적으로 정화 하도록 추가 설계하였다.

둘째, 상관도에 따른 자동예측 구간 설정 기법으로, [6]에서는 예측구간을 구하기 위해 t분포를 사용하였으며 회귀선에 맞는 예측구간을 설정하기위해 특정 값을 입력으로 주었다. 개선된 알고리즘에는 자동으로 예측구간을 효과적으로 설정할 수 있도록 하였다. 결함 예측 과정에 있어 모든 선박은 같은 환경을 가질 수 없기 때문에 각 선박에 맞는 환경을 설정해 주어야 한다. 결함 예측을 위해 필요한 상관계수 계산 주기, 결함 예측 주기, 결함 판단 기준 등을 매개변수로 두어 현장에서 선박의 환경에 맞게 설계하도록 하였다. 마지막으로 기존의 선형회기법을 센서간의 값에 대하여 구간별로 나누어 기존의 알고리즘보다 정밀하고 정확하게 결함을 예측할 수 있도록 하였다.

Ⅱ. 관련 연구

지금까지 결함을 예측하는 방법에 관한 연구는 많이 이루어졌지만, 기계가 스스로 기존의 데이터를 학습하여 실시간으로 수집되는 데이터에 대한 분석을 통해 결함의 발생 가능성 여부를 판별하는 연구는 아직까지도 부족하다. 결함을 예측하는 다양한 방법 중 대표적인 방법으로는 수학적 모델, 임계치 검사 방법, 상관계수, 편차를 이용한 방법 등이 있다[7].

임계치 검사 방법은 구현이 간단하고 적용하기가 쉬워 가장 많이 사용되고 있는 방법으로 실시간으로 수집되는 데이터 중에서 전문가가 지정한 센서의 정상 허용범위를 벗어나면 결함으로 판단하는 방법이다[8]. 임계치 검사 방법은 결함 유무를 쉽게 판단할 수 있다는 장점이 있으며, 센서의 값이 임계치와의 차이가 많이 나야지만 결함 판단이 가능한 단점이 있다. 임계치 검사 방법은 통신 문제나 기상이변으로 인한 잡음이 발생하여 데이터 값에 간섭할 수 있다. 따라서 시스템의 작은 간섭에도 데이터의 변형을 초래할 수 있어 고려되어야 하며, 임계치를 설정할 때 유의해야만 한다[9][10].

상관분석법에 의한 방법은 수치로 센서 간의 관계를 나타내는 것으로 변수 간의 연관 정도를 수치로 나타내는 상관계수를 이용하여 센서 간의 연관성을 정량적으로 표현하는 것이다. 상관분석법에 따른 방법은 선박기관시스템이 많은 기관으로 구성되어 있어 설치된 모든 센서로부터 발생하는 데이터를 처리하기 위해서는 한계가 있다. 상관관계를 분석하여 데이터를 축소함으로써 정확한 계산과 시간을 단축할 수 있다[11].

Ⅲ. 회귀분석을 이용한 결함 예측 알고리즘

본 논문에서 제안하는 알고리즘은 그림 1과 같이 구성되어 있으며 표 1은 개선된 고장 예측 알고리즘이다. [6]에서 제안한 알고리즘 보다 본 논문에서 개선된 알고리즘은 단계1의 데이터 정화 부분, 단계3의 상관관계에 따른 자동 예측 구간 설정, 단계 4의 구간 결합 예측, 기본 결합 예측 분류, 단계 5-1 구간별 회귀분석 (스플라인 회귀 분석) 적용이다.

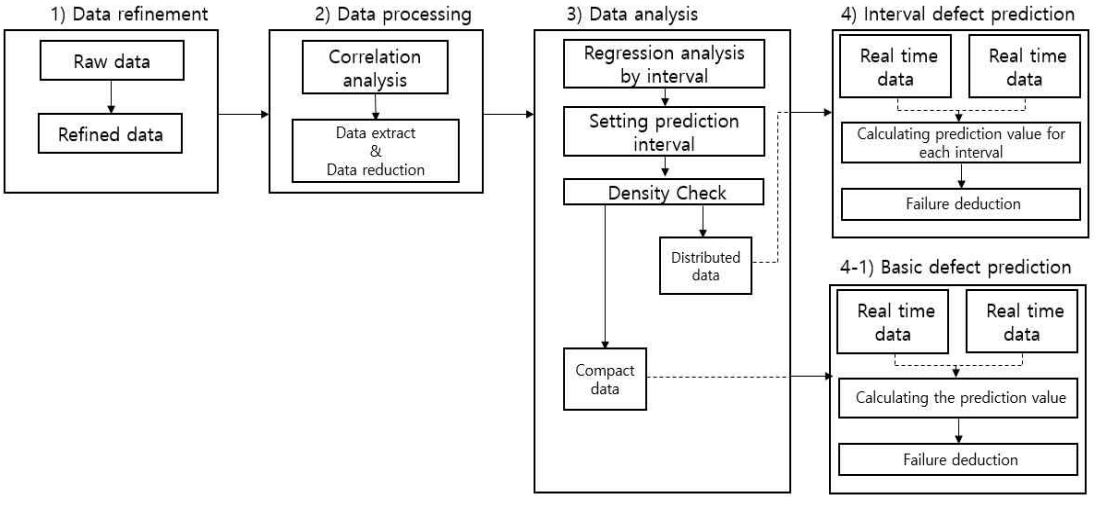

그림 1. 결함 예측 알고리즘 모델
Fig. 1. Failure forecasting algorithm model

표 1. 고장 예측 알고리즘
Table 1. Failure forecasting algorithm

Step 1. Refine noisy data
Step 2. Filter sensors of |Correlation coefficient| > alpha
Step 3. Set up automatic prediction interval according to correlation after regression analysis
Step 4. Classify data into interval defect prediction and basic defect prediction after checking data density
Step 5-1. Deduct failure by regression analysis for each interval
Step 5-2. Deduct failure by simple regression analysis
Step 6. Print a fault prediction message when an error occurs over a certain range.

3.1 데이터 정화

선박으로부터 수집된 데이터는 환경에 따라 쉽게 변할 수 있다. 변형된 데이터를 사용한 결함 예측은 효과적이지 못하므로 데이터를 효과적으로 정화해 주는 단계가 필요하다. [6]에서 제안한 방법에 추가적으로 통신 장애 또는 배의 출항 정지, 도선작업등으로 인해 발생 가능한 불특정한 주기로 일반적인 센서의 데이터 값을 무시하였다.

3.2 데이터 처리

3.1절의 과정이 끝나면 정화된 데이터를 사용하여 상관분석(Correlation Analysis)을 하고 상관계수를 계산한다. 상관계수를 통해 높은 관계가 있는 센서 쌍만 추출하여 3.3절에서 효과적인 분석을 할 수 있다. 데이터의 밀집도에 따라 단순 선형 회귀분석과 구간별 회귀분석으로 나뉘며 식 (1)과 식 (2)를 통해 상관계수를 구한다.

데이터의 밀집도가 골고루 분포되어 있는 경우는 개선된 구간별 선형회귀 분석(스플라인 선형 회귀 분석)을 사용하여 예측값을 설정하고, 밀집도가 한 영역으로 포진되어 있는 경우는 [6]에서 제시한 기존의 방법대로 예측값을 설정한다.

$$r = \frac{S_{XY}}{S_X S_Y} \tag{1}$$

$$S_{XY} = \frac{\Sigma(X_i - \overline{X})(Y_i - \overline{Y})}{n-1}$$

S_X : X의 표준편차
S_Y : Y의 표준편차
r : 표본의 상관계수
S_{XY} : 표본의 공분산

$$r_k = \frac{(S_{XY})_k}{(S_k)_X (S_k)_Y} \quad (2)$$

$$(S_{XY})_k = \frac{\sum((X_k)_i - \overline{X_k})((Y_k)_i - \overline{Y_k})}{n_k - 1}$$

$(S_k)_X$: k구간에 대한 X의 표준편차

$(S_k)_Y$: k구간에 대한 Y의 표준편차

3.3 데이터 분석

3.3.1 표본 회귀식을 이용한 예측값 설정

3.2절의 과정이 끝나면 추출된 센서 상들에 데이터 분석단계를 적용한다. 데이터 분석 단계에서는 회귀분석을 사용하여 실시간으로 수집되는 값에 대하여 결함을 예측한다.

$$\hat{Y} = a + bX \quad (3)$$

$\hat{Y}$:종속변수, a :절편, b :기울기, X :독립변수

식 (3)은 표본회귀식을 나타낸 것이다. 표본 회귀식에서 절편 a와 기울기 b가 회귀식을 결정한다. a와 b 값을 알면 독립변수의 값을 주었을 때 $\hat{Y}$을 예측할 수 있다. 여기서 a와 b를 회귀 계수라고 한다. 본 논문에서는 최소자승법을 이용하여 회귀식을 구하였다.

$$a = \overline{Y} - b\overline{X} = \frac{\sum Y_i}{n} - b\frac{\sum X_i}{n} \quad (4)$$

$$b = \frac{\sum X_i Y_i - n\overline{X}\overline{Y}}{\sum X_i^2 - n\overline{X^2}}$$

a : 절편, b : 기울기

$\overline{X}$: X의 평균, $\overline{Y}$: Y의 평균

n : 데이터의 개수

식 (4)는 최소자승법을 통하여 구해진 절편 a와 b이다. 식 (4)에서 계산되어진 회귀계수를 이용하여 관찰치의 예측값을 구할 수 있다. 하지만 실시간으로 관측된 데이터가 예측값에 정확하게 일치하기는 힘들어서 예측구간을 사용하여 결함 검출에 적용한다. 예측구간의 범위는 식 (5)와 같다[12].

$$b_1 \pm t_{a/2,\, n-2} \sqrt{\frac{\sum(y - \hat{y})^2 / n - 2}{\sum(x - \overline{x})^2}} \quad (5)$$

b_1 : 예측값, α : 오차율, n : 자유도(데이터 개수)

x : 독립변수, y : 종속변수,

3.3.2 상관도에 따른 자동 예측 구간 설정

[6]에 제시된 방법은 예측구간을 구할 때는 상관계수에 상관없이 t분포값을 센서 쌍에 적합하도록 수동적으로 설정하여 사용한다. 수동으로 사용자가 설정하지 않을 시에는 디폴트 값을 사용하여 각 센서의 상관도와 상관없이 일정한 예측구간이 적용되어 정밀한 예측 구간 설정에 한계성을 가졌다. 또한 수동적으로 설정한다 하더라도 정밀성을 위해 매번 각 센서 쌍 마다 데이터를 분석하여 적절한 t분포값을 구해야 하는 문제가 있다. 본 논문에서는 이러한 문제점을 해결하기 위해 상관계수에 따라 t분포값을 자동화 할 수 있는 식을 설계하였다. 여기서 고려되는 데이터는 3.2절에서 설명한 과정을 통해 선별된 상관계수가 $|cor| \geq 0.7$인 데이터만을 사용하였으며 상관계수에 비례하여 t분포값을 가지도록 하였다.

$$Tvalue = a - (|cor| - 0.7) \times b \quad (6)$$

a :초기값, b :상관도조정 계수, $|cor|$:상관계수

3.3.3 데이터 분포에 따른 선형 회귀식 선택

본 절에서는 개선된 새로운 기법을 사용하기 위해 데이터들의 분포를 분석하는 과정이다. 이 과정을 통해 데이터들의 분포가 골고루 펼쳐져 있다면 새롭게 제시된 구간별로 선형 회귀식을 적용하여 예측값을 설정하는 구간별 선형회귀식(스플라인 선형회귀식) 방법을 선택하고, 한쪽 값에 그 분포가 치우쳐 있으면 [6]에서 제시된 전체구간 선형 회귀식 방법으로 예측값을 설정한다. 한쪽으로 밀집되어 있는 데이터일 경우는 구간별로 선형회귀식을 따로 구하는 것이 무의미하기 때문이다.

예측 구간을 설정한 뒤 전체구간에 대한 선형회귀와 구간별 선형회귀 중 어떤 방법을 사용할지 정한다. 데이터의 분포가 특정 구간에 치중되어 있다면 전체구간에 대한 선형회귀를 사용하고 균등하게

분포가 되어 있다면 구간별 회귀분석을 사용하여 예측값과 예측구간을 계산한다. 구간별 회귀분석이란 데이터의 범위를 n등분하여 각 구간에 대한 회귀식을 구하고 분석 하는 것을 말한다. 구간별 회귀분석은 전체구간에 대한 선형회귀분석에 비해 정확하고 정밀하다. 구간별로 회귀식을 적용 할 경우 구간의 특징을 반영할 수 있어 배의 상태에 따른 데이터들의 값을 민감하게 반응할 수 있어 예측값 설정에 정밀성을 기대할 수 있다.

3.4 결함 예측

3.4.1 결합 예측을 위한 로그 기록

선박의 상태가 정상일 때 수집된 데이터 값을 이용하여 3.3절 단계에서 데이터의 표본 회귀식과 예측구간을 설정한다. 실시간으로 전송되는 데이터에서 독립변수에 해당하는 데이터를 3.3절에서 설명한 함수에 적용하여 종속변수의 예측구간을 구한다. 실제로 수집된 종속변수가 예측구간을 벗어나면 해당 시점에 결함이 있다고 판단하여 결함 로그를 남긴다.

예를 들어, 센서 1의 값 10에 대한 센서 2의 예측 구간이 0 ~ 20으로 계산 되었을 경우, 관측된 센서 1의 값이 10일 때 센서 2의 정상 데이터는 0 ~20의 값으로 해석한다. 이 때 센서 2의 값이 15이면 정상으로 판단하고 22일 경우는 결함으로 판단한다.

3.4.2 윈도우 슬라이딩과 0점 조정 실시간 그래프

고장예측 알고리즘이 적용될 E-CBM(E-Condition Based Monitoring System)에서 오류 검출 기간을 사용자 중심으로 쉽게 파악하도록 하기 위해 슬라이딩 윈도우를 설정하는 기능을 추가하였다.

또한 사용자가 실시간으로 입력되는 센서 값이 예측값, 예측 구간 사이에 있는지 여부를 쉽게 분석할 수 있도록 그림 3과 같은 기존의 시계열그래프를 그림 4와 같이 0점 조정된(예측값이 0점 조정된) 그래프로 개선하였다. [6]에서 제시한 방법에서 슬라이딩 윈도우와 사용자 중심 그래프 표현을 추가하였다.

슬라이딩 윈도우의 기본크기는 2일 이다. 사용된 기간만큼 데이터를 유지하기 위해 새로운 데이터가 수집되면 가장 오래된 데이터를 내보내고 데이터베이스를 통해 입력 받은 새로운 데이터를 윈도우에 포함시킨다. 이때 로그 기록의 범위도 윈도우 슬라이딩의 기간 동안만의 기록을 가지고 있다. 해당 윈도우 동안 디폴트 값이나 사용자의 입력에 따라 정해진 오류 비율 % 이상 로그 기록을 가지는 센서는 고장이 발생할 수 있다고 예측 메시지 보낸다. 윈도우의 기간이 넓을 경우 오류 검출의 민감성은 떨어지나, 일시적인 장애나 상태(폭풍우)로 인해 발생하는 기계들의 센서 값들에 대해서는 오류 메세지를 검출하지 않는다.

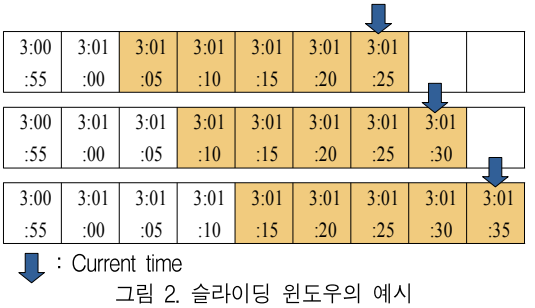

그림 2. 슬라이딩 윈도우의 예시
Fig. 2. Example of sliding window

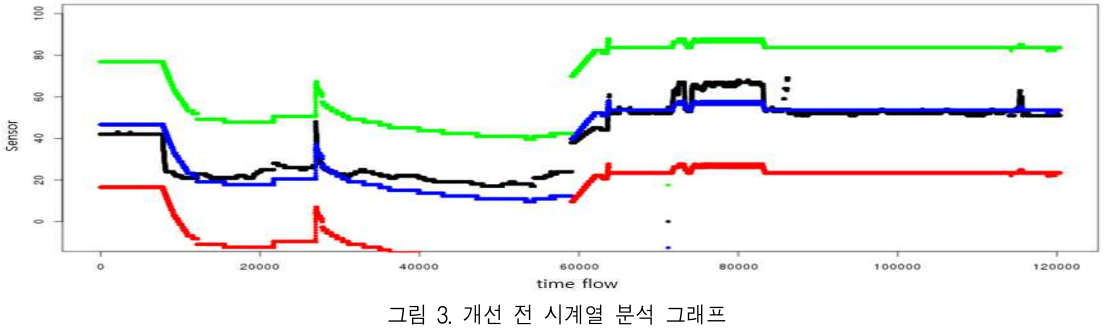

그림 3. 개선 전 시계열 분석 그래프
Fig. 3. Time series analysis graph before improvement

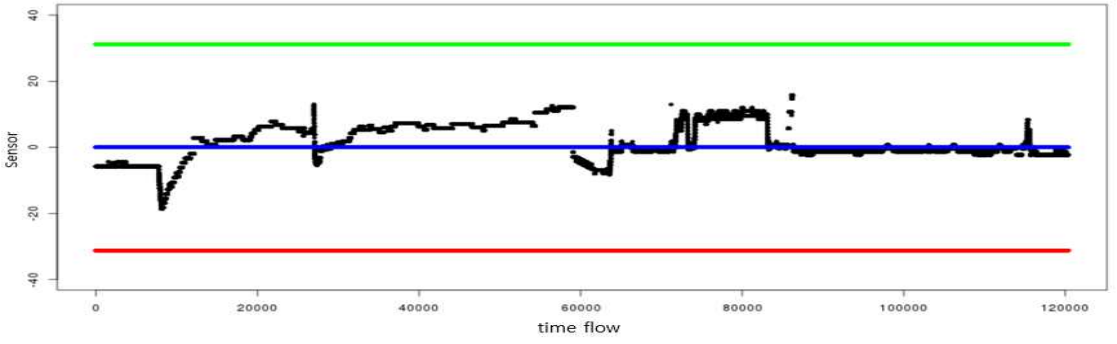

그림 4. 개선 후 시계열 분석 그래프
Fig. 4. Time series analysis graph after improvement

윈도우의 기간이 좁으면 오류검출의 민감성이 높아지는 장점이 있지만 일시적인 통신오류를 모두 결함으로 처리될 단점이 있다.

사용자 중심 그래프 표현은 예측값을 항상 0으로 두어 일직선으로 표현이 되어 사용자 입장에서 현재 상황에 대한 빠른 판단을 할 수 있도록 하였다.

그림 3의 그래프보다는 그림 4의 그래프가 실시간으로 입력되는 데이터를 구별하기 쉬우며, 또한 0점인 예측값으로부터 떨어진 정도를 쉽게 파악할 수 있다.

Ⅳ. 개선된 결함 예측 알고리즘 적용

적용단계에서 사용한 데이터는 H해운사의 H-Mar (HNMW)호에서 2015년 4월 16일부터 2017년 10월 현재까지 5~30초 간격으로 기록된 데이터이다. 해당 기간 동안 피스톤 링과 실린더 라이너 결함 그리고 엔진오일 누유가 보고되었다. 보고된 결함과 관련된 센서를 제안하고 있는 알고리즘에 적용하여 하여 실제로 보고된 날짜 이전에 결함을 탐지하여 본 논문에서 제안하는 알고리즘의 효율성을 실험적으로 검증하였다.

4.1 피스톤 링, 실린더 라이너 결함

2016년 1월 1일 피스톤 링 및 실린더 라이너의 결함이 발생하여 새 제품으로 교체된 기록이 있다. 해당 부품과 연관성이 높은 센서를 분석해본 결과 센서 1~4에서 결함의 징조를 찾을 수 있었다.

표 2. 센서 정보
Table 2. Sensor information

Number of sensor	Sensor description
sensor 1	Temperature of C/A entering main engine
sensor 2	Temperature of receiver collecting C/A
sensor 3	Outlet temperature of cooling oil for piston
sensor 4	Temperature of main bearing lubricant oil
sensor 5	Fuel viscosity
sensor 6	Temperature of the fuel before entering the supply unit

4.1.1 센서 1 vs 센서 2

그림 5~8은 센서 1~2의 상관관계를 나타낸 그래프이다. 두 센서 사이의 상관계수는 0.794로 실시간 분석 시 잘 확인해볼 필요가 있는 센서이다. 그림 5는 2015년 5월~8월의 결함이 보고되지 않은 정상 기간의 데이터로 표본 회귀식과 예측구간을 구한 것이다. 그림 6은 결함이 보고되기 30일 이전의 데이터로 상관관계를 나타낸 것이다. 그림 5는 측정된 데이터가 정상 판단 범주에 모두 포함된다. 그림 5는 [6]의 기법에 따라 정상적인 데이터이고 이를 토대로 상관계수를 구하고 선형회귀식으로 예측값을 설정하였다. 그러나 x축 센서 30 ~35 사이 데이터와 50이상의 데이터의 분포는 예측 회귀선과는 다소 다른 실측값을 보이고 있다. 45 ~50 사이의 데이터에 대해서만은 본 회귀식의 예측값을 따르고 있는 것으로 보인다. 그림 6은 구간별로 회귀 분석을 통해 그림 5보다는 각 구간의 정상 데이터의 분포의 값을 반영하는 형태의 예측값이 나오는 것을

확인할 수 있다. 예측값의 정밀성이 개선되었다 할 수 있다.

그림 7은 [6]의 기법대로 오류 검출을 한 데이터의 분포이고 그림 8은 본 논문에서 제시된 개선된 구간별 선형 회귀 분석으로 오류 검출 데이터 나타내고 있다. 실측된 데이터 중 50 이상의 데이터의 경우는 그림 5과 그림 6에서 모두 정상적인 범위의 데이터로 처리되었으나, 그림 5에서는 예측 구간내의 데이터 값으로 표시되나 그림 6에서는 거의 예측값에 가까운 값임을 할 수 있다. 따라서 예측구간의 경계에서는 그림 5로 표시되는 [6]의 기법보다는 그림 6으로 표시되고 있는 개선된 방법이 정밀성이 높다. 구간별 예측 구간의 경계선이 비연속적인 이유는 구간별 상관도가 달라져, 예측구간이 자동 설정되기 때문이다.

그러나 그림 7~8을 보게 되면 신뢰구간을 벗어나는 데이터를 확인 할 수 있다.

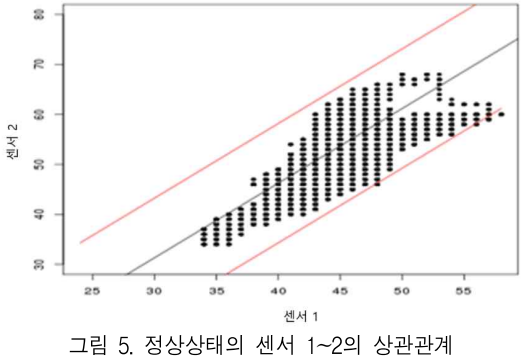

그림 5. 정상상태의 센서 1~2의 상관관계
Fig. 5. Correlation between sensor 1~2 at normal condition

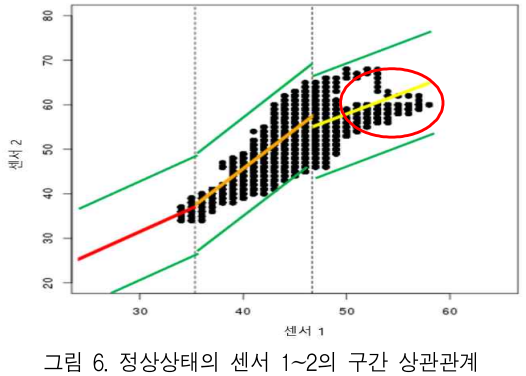

그림 6. 정상상태의 센서 1~2의 구간 상관관계
Fig. 6. Correlation by division between sensor 1~2 at normal condition

그림 8은 본 논문에서 제안하는 구간별 회귀분석을 센서 1~2에 적용한 그래프이다. 각 구간별로 회귀식 및 예측구간을 산출하여 기존 방법인 그림 7 보다 정확하고 효율적으로 결함을 예측하는 것으로 확인했다. 따라서 해당 센서의 결함이 발생한 시점을 알기 위해 그림 9와 같이 시계열 분석이 필요하다.

그림 9는 시간의 흐름에 따른 센서 1의 값과 정상데이터로 구해진 예측구간을 나타내고 있다. 12월 20일까지 정상적인 범위를 유지하지만 12월 20일 이후에는 한계치를 벗어나는 것을 확인할 수 있다. 전문가의 의견에 따르면 20일에 피스톤 링의 절손 발생 시점으로 판단된다. 20일부터 센서 1의 값이 상한값을 벗어나는 것을 확인할 수 있다.

[6]의 기법은 약 10일 이전에 결함이 예측되었고 개선된 기법 그림 9는 약 13일 전에 예측된다.

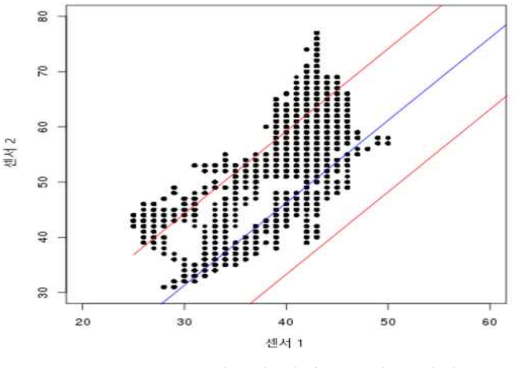

그림 7. 결함 보고 직전의 센서 1~2의 상관관계
Fig. 7. Correlation between sensor 1~2 just before failure report

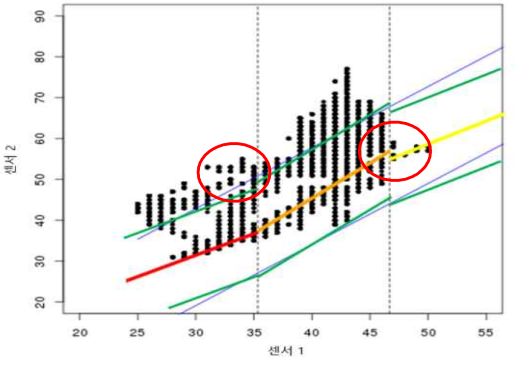

그림 8. 결함 보고 직전의 센서 1~2의 구간 상관관계
Fig. 8. Correlation by division between 1~2 just before failure report

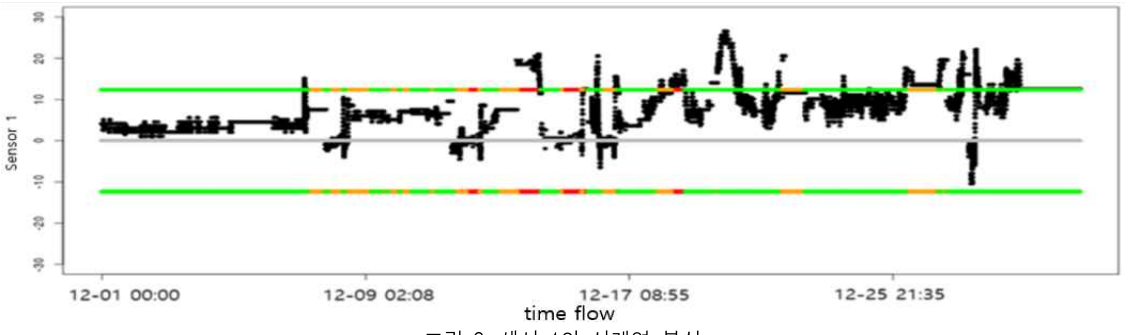

그림 9. 센서 1의 시계열 분석
Fig. 9. Time series analysis of sensor 1

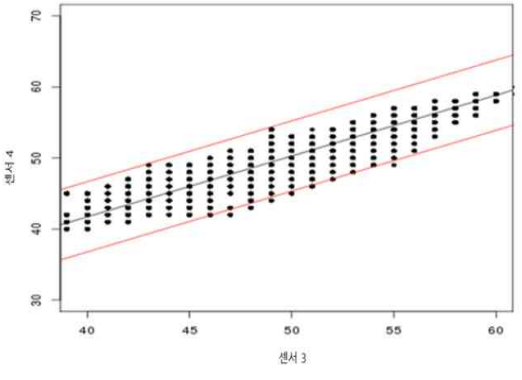

그림 10. 정상상태의 센서 3~4의 상관관계
Fig. 10. Correlation between sensor 3~4 at normal condition

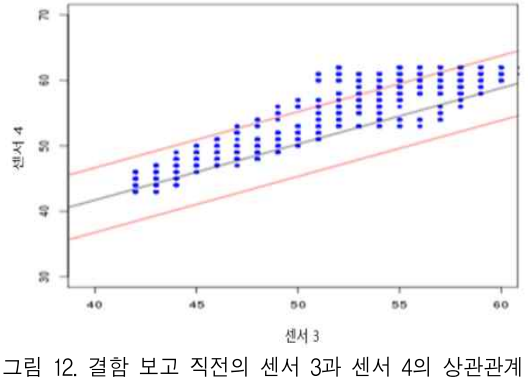

그림 12. 결함 보고 직전의 센서 3과 센서 4의 상관관계
Fig. 12. Correlation between sensor 3 and sonsor 4 just before failure report

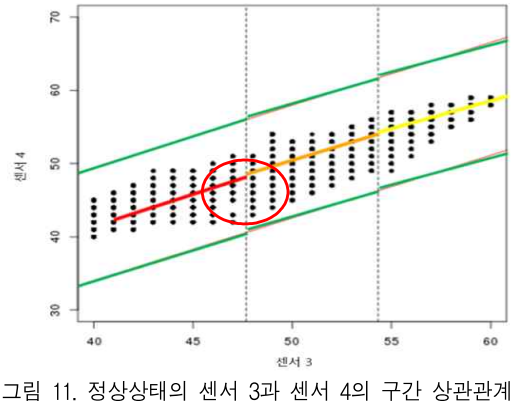

그림 11. 정상상태의 센서 3과 센서 4의 구간 상관관계
Fig. 11. Correlation by division between sensor 3 and sensor 4 at normal condition

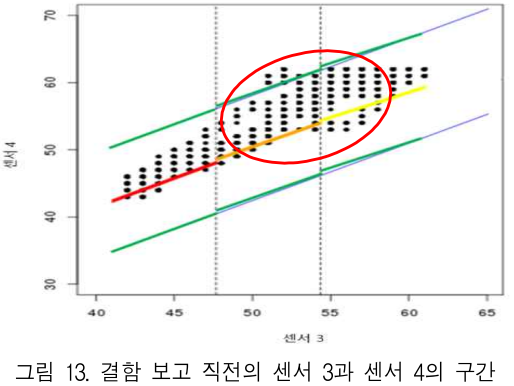

그림 13. 결함 보고 직전의 센서 3과 센서 4의 구간 상관관계
Fig. 13. Correlation by division between sensor 3 and sensor 4 just before failure report

4.1.2 센서 3 vs 센서 4

4.1.1에서 피스톤 링 결함 분석이 의심되므로 그에 관한 다른 두 센서를 분석하였다.

그림 10~13은 센서 3~4의 상관관계를 나타낸 그래프이다. 두 센서간의 상관계수는 0.9883582로 결함 예측에 필요한 데이터 쌍이다. 그림 10과 11은 정상상태에서의 표본회귀식과 예측구간을 나타낸 것이다.

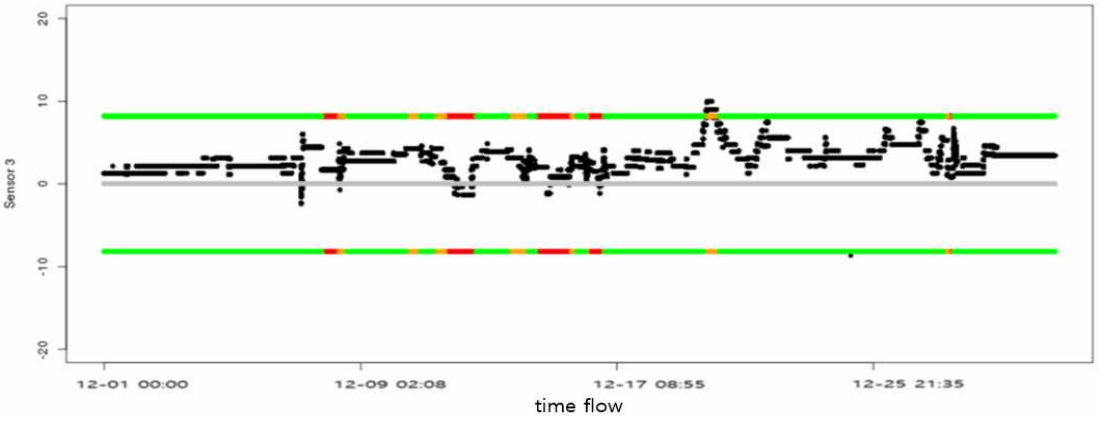

그림 14. 결함 보고 직전 센서 3의 시계열 분석
Fig. 14. Time series analysis of sensor 3 just before failure report

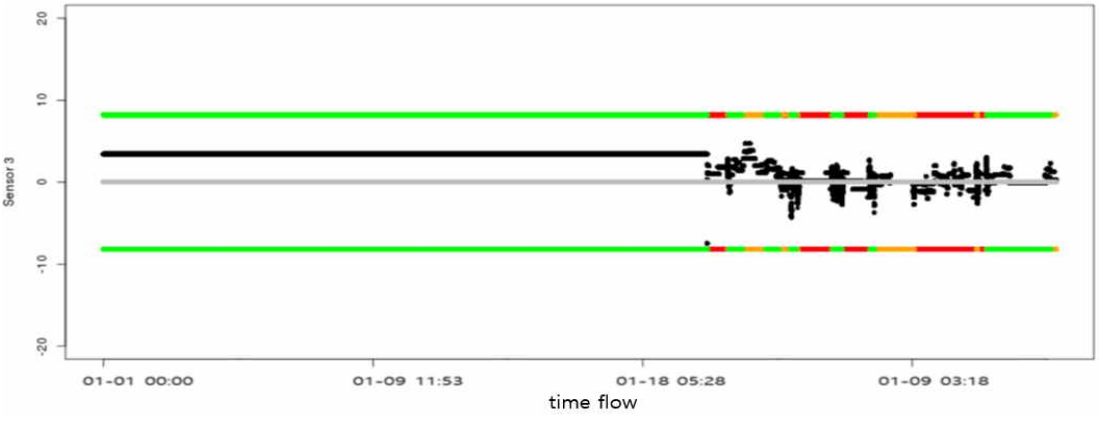

그림 15. 수리 직후 센서 3의 시계열 분석
Fig. 15. Time series analysis of sensor 3 just after repairing

그림 10 데이터의 분포의 경우 45 이하의 경우의 분포 55이상의 분포는 예측값을 나타내는 그래프와 약간의 상이함이 나타난다.

그림 11의 경우는 이러한 점이 반영된 정밀한 예측값 그래프가 그려진다. 그림 12에서는 오류 데이터라고 하는 부분을 그림 13에서는 정상적인 예측 구간에 포함이 된다. 실험을 통해 전체적으로 정밀성이 높아져 오류 검출의 정확성이 개선됨을 확인하였다.

그림 14와 그림 15는 센서 3~4에 대한 시계열 분석을 나타내는 그래프이다. 시계열 분석 또한 구간을 나누어 어느 구간 데이터가 어느 시점에 결함이 발생한 것인지 알 수 있다.

각 구간에 대하여 더 정확한 선형회귀식을 적용하게 되면서 결함을 더 효율적으로 탐지할 수 있었다. [6]의 기법에서는 약 10일 이전에 결함을 탐지할 수 있었으나, 본 논문에서 제시한 방법으로 약 12일 이전에 예측할 수 있다는 것을 확인했다.

V. 결론 및 향후 과제

본 논문에서는 [6]에서 제시한 선형회귀 분석의 정밀성과 정확성을 개선하기 위해 자동 예측 구간 설정과 구간별 회귀분석 기법을 제시하였다. 그리고 정상적인 데이터들의 판별을 위하여 데이터 정화 방법을 개선하였다. 정화된 데이터와 자동 예측 구

간 설정기법을 이용하여 구간별 회귀 분석을 통한 예측값 예측 구간을 이용하여 오류검출을 위한 정상적인 데이터 범위를 설정하여 검출의 정밀성을 향상시켰다.

향후 과제로는 정상적인 데이터를 위한 데이터 정화 작업을 위해 더 많은 연구가 진행되어야 한다. 실제로 제시된 데이터 정화 기준 이외에서 선상에서 발생하는 많은 예외들이 존재한다. 이것에 대한 체계적인 연구를 통해 선박을 운항하는데 있어 발생 가능한 케이스별 센서들의 값 변동사항들을 분석하여 효과적인 데이터 정화 기준으로 설정하는 것이 더 정밀하고 정확한 예측값 계산 가능하도록 하는 핵심 기술이 될 것이다.

References

[1] MAN B&W Diesel A/S, "Project Guide and S80MC/S80MCE", Copenhagen Denmark, pp. 1.1-1.20, 1990.

[2] New Sulzer Diesel Ltd, "Engine Selection and Project Manual", RTA84(C,T,M), Winterthur Switzerland, pp. 1-9, 1994.

[3] Y. I. Kim, H. K. Oh, and Y. H. Yu, "The Development of Diesel Engine Room Fault Diagnosis System Using a Correlation Analysis Method", Korean Society of Marine Engineers, Vol. 30, No. 2, pp. 253-259, Mar. 2006.

[4] J. C. Bae, G. H. Kang, Y. C. Ha, K. S. Hwang, H. S. Lee, and S. M. Kim, "A Study of a Marine Engine Monitoring System base on IEEE 802.15.4", The Korean Society of Marine Engineering (KOSME), pp. 383-385, 2010.

[5] D. H. Kim, "A study on Implementation of a Real-Time Fault Diagnossis System for Marine Diesel Engines", Master Thesis, University of KOREA Marine and Ocean University, 2001.

[6] Jun-Hyun Park, Min-Kook Jang, Gang-Hui Lee, Eun-Kyung Oh, and Sung-Woo Hur, "Forecasting Algorithm for Vessel Engine Failure", Journal of Korean Institute of Information Technology, Vol. 14, No. 11, pp. 109-117, Nov. 2016.

[7] Junichi Hirata and Koji Takasaki, "Support Scheme in Collaboration with Manufacturers Making Use of the Diagnostic Platform Based on Big Data Analysis", CIMAC World Congress, No. 74, 2016.

[8] J Gertler, "Fault Detection and Diagnosis in Engineering System", CRC, Virginia, USA. 1998.

[9] Y. J. Lee, D. Y. Kim, M. S. Hwang, and Y. S. Cheong, "An event-driven intelligent failure analysis for marine diesel engines", Korea Industrial Information System Research, Vol. 17, No. 4, pp. 71-85, Aug. 2012.

[10] Y. J. Lee, D. Y. Kim, M. S. Hwang, and Y. S. Cheong, "An event-driven Failure Analysis System for Real-Time Prognosis", Journal of the Korea Industrial Information System Society, Vol. 18, No. 4, pp. 250-257, Aug. 2013.

[11] K. Y. Kim, Y. I. Kim, and Y. H. Yu, "The Fault Diagnosis of Marine Diesel Engines Using Correlation Coefficient", The Institute of Electronics and Information Engineers, pp. 41-45, 2010.

[12] S. I. Pak and T. H. Oh, "Correlation and Simple Linear Regression", The Korean Society of Veterinary Clinics, Vol. 27, No. 4, pp. 427-434, 2010.

저자소개

박 준 현 (Jun-Hyun Park)

2015년 2월 : 동아대학교 컴퓨터공학과(공학사) 스마트그리드(공학사)

2017년 2월 : 동아대학교 컴퓨터공학과(공학석사)

2017년 3월 ~ 현재 : 동아대학교 컴퓨터공학과(박사과정)

관심분야 : 알고리즘, 딥러닝, 이미지/영상 처리

오 은 경 (Eun-Kyung Oh)

2000년 ~ 현재 : 컴퓨터공학과 강사
1997년 2월 : 동아대학교 컴퓨터공학과(공학사)
2003년 2월 : 동아대학교 컴퓨터공학과(공학석사)
2010년 2월 : 동아대학교 컴퓨터공학과(공학박사)
관심분야 : Cad Algorithm, 전산 수학, VLSI 알고리즘, 최적화 알고리즘

장 민 국 (Min-Kook Jang)

2007년 3월 : 한국해양대학교 기관시스템공학부
2007년 3월 ~ 2012년 5월 : 한진해운 해상직원 근무
2012년 6월 : 한진 SM 육상직원 근무
2015년 1월 : 유수 SM으로 사명변경
2017년 2월 : 한국해양대학교 기관공학과 석사학위 취득
2017년 9월 : 지마린서비스로 사명 변경 (現 기술연구소 근무 중)
관심분야 : 스마트 선박. 빅데이터, 엔진, 고장 진단, 친환경 선박

서 영 우 (Young-Woo Seo)

2004년 2월 : 방송통신대학 컴퓨터공학(학사)
2006년 7월 : Univ Of Salford, UK(석사)
2006년3월 ~ 2007년 2월 : 라파앤컴퍼니, 서울 한국
2007년 4월 ~ 2010년 12월 : 라파앤컴퍼니, 산호세, 미국
2011년 1월 ~ 2011년 12월 : 액센추어, 뉴욕, 미국
2014년 4월 ~ 현재 : (주)투그램시스템즈 대표이사

허 성 우 (Sung-Woo Hur)

1981년 2월 : 경북대학교 전자공학과(공학사)
1983년 2월 : 한국과학기술원 전산학과(공학석사)
2000년 6월 : Univ. Illinois (Chicago) EECS (Ph.D.)
2002년 2월 ~ 2006년 8월 : 미국 인텔 연구소 (산호세) 기술자문위원
1986년 3월 ~ 현재 : 동아대학교 컴퓨터공학과 교수
관심분야 : Cad Algorithm, Combinatorial Optimization