

이벤트 기반 지능형 선박엔진 결함분석[†]

(An event-driven intelligent failure analysis for marine diesel engines)

이 양 지*, 김 덕 영**, 황 민 순***, 정 영 수***

(Yang-Ji Lee, Duck-Young Kim, Min-Soon Hwang, and Young-Soo Cheong)

요약 본 논문은 운항중인 선박에서 기록되어지는 운항정보 및 엔진 가동정보 등을 실시간으로 모니터링하고, 문제 발생 시에 그 근본원인을 찾아내어 민첩하게 대응할 수 있는 일련의 결함원인 분석 및 예방시스템 개발을 목적으로 한다. 결함분석을 위해서는 선박엔진의 주요기관에 부착된 센서들로부터 장기간 수집된 정보를 사용하게 되는데, 이 양이 매우 방대하며, 잡음 및 중복정보(Redundancy)가 너무 많이 포함되어, 수집된 센서 데이터를 바로 고장분석에 사용하기에는 어려움이 있다. 따라서 본 논문에서는 방대한 양의 데이터 중, 정보의 손실을 최소화하고 중요한 정보만을 추출하기 위해 ‘Equal-frequency binning’과 ‘Entropy’ 기반의 데이터 필터링 방법에 관해 연구하였다. 실제로 시험운용 중인 선박엔진 데이터를 개발된 선박엔진 고장분석 소프트웨어를 이용하여 결함분석을 수행하여, 제안된 방법의 효용성을 검증한다.

핵심주제어 : 결함분석, 선박엔진, 데이터 이산화, 이벤트 로그

Abstract This paper aims to develop an event-driven failure analysis and prognosis system that is able to monitor ship status in real time, and efficiently react unforeseen system failures. In general, huge amount of recorded sensor data must be effectively interpreted for failure analysis, but unfortunately noise and redundant information in the gathered sensor data are obstacles to a successful analysis. This paper therefore applies ‘Equal-frequency binning’ and ‘Entropy’ techniques to extract only important information from the raw sensor data while minimizing information loss. The efficiency of the developed failure analysis system is demonstrated with the collected sensor data from a marine diesel engine.

Key Words : Failure analysis, Marine diesel engine, Data discretization Event log

1. 서 론

선박의 대형화, 전용화, 자동화가 급속화되고, 운용이 복잡해짐에 따라 안정성과 경제성이 더욱더 요구

된다. 이를 위해 선박의 운용 중 동작 상태를 감시하며, 동작의 이상 상태가 발생할 경우 이에 대한 빠른 조치를 할 수 있어야 한다. 따라서 선박에 대한 상태 정보를 실시간으로 감시하고 제어할 수 있는 선박엔진 모니터링시스템(Monitoring system)과 이러한 동작상태 정보를 이용하여 자동으로 의사결정을 할 수 있는 알고리즘 혹은 고장 발생 시에 고장의 근본적인 원인을 파악하여 빠르게 조치가 가능한 고장진단시스

[†] 본 연구는 “울산시 및 교육과학기술부의 울산과학단지 기초·원천 R&D 과제 지원 사업”의 지원을 받아 수행된 것임.
* UNIST 디자인 및 인간공학부
** UNIST 디자인 및 인간공학부, 교신저자
*** 현대중공업 통신운영부

템(Fault diagnosis system)의 구축이 필수적이다 [1-2].

선박의 경우 안정성을 평가하기 위해 실제 해상상태에서 손상 선박의 거동을 예측하거나 파랑 중 선박의 구조적인 결함여부를 나타내는 구조적안정성 평가 등을 수행하는 시뮬레이션 기반 시스템이 많이 구축되어 있다[3]. 또한 선박의 사고사례나 관련 데이터들을 이용하여 정의된 시나리오 및 변화하는 환경조건에 반응하여 선박의 고유 특성에 근거하여 안정성을 평가하는 등 구조적인 해석과 관련된 많은 연구가 진행되고 있다[4]. 이러한 선박의 구조적 해석 및 시뮬레이션 기반의 시스템은 많이 개발 되었지만 선박의 이상 유무를 판단하고 분석할 수 있는 시스템의 개발에 대한 연구는 부족하다. 현재 선박에 탑재된 시스템은 대부분 각 장치에 대한 신호가 안정 범위 이상을 넘었을 경우 고장 판별을 알리는 시스템으로 고장의 여부만을 알려주는 시스템이지만, 고장의 전조 증상을 미리 판단한다든지 인공 지능적으로 고장 확률, 부품의 교체 시기 등을 알려줄 수 있는 시스템이 개발된다면 결함을 미리 방지하여, 안정성과 경제성을 높일 수 있을 것이다.

일반적으로 고장진단시스템은 고장검출(Fault detection), 고장분리(Fault isolation), 고장평가(Fault identification)의 3단계 과정으로 구성된다. 고장검출의 단계에서는 발생한 고장의 유무를 판별하고, 판별된 고장은 고장분리의 단계에서 고장의 종류를 알아낸다. 그리고 고장의 정도를 고장검출 단계에서 결정한다.[5] 이러한 3단계의 프로세스 과정을 거치는 고장진단시스템에 관한 연구가 활발히 수행되고 있으며, 크게 수학적 모델을 기반으로 하는 방법과 모델을 기반으로 하지 않고 데이터를 기반으로 하는 방법으로 분류할 수 있다. 수학적 모델을 기반으로 하는 방법의 경우 일반적으로 입력(Input)과 출력(Output)에 관한 명확한 수학적 모델이 존재할 때 사용하며, 상태추정 접근법과 파라미터추정 접근법이 있다[6]. 반면에 수학적 모델이 없을 경우 실시간으로 들어오는 데이터나 계속적으로 수집된 과거(Historical) 데이터를 이용하는 데이터 기반의 방법에는 한계치 검사 기법, 전문가 시스템 및 통계적 분석 기법 등이 있다. 수학적 모델을 기반으로 하는 분석의 경우 시스템의 정확한 정량적 모델이 필요하기 때문에 선박의 운전환경에 따라 상태가 변하는 경우 정확도가 떨어지며, 수학적 모델을 얻

지 못하는 경우 적용하지 못하는 단점이 있다. 그렇기 때문에 본 논문에서는 정확한 수학적 모델이 없어도 적용 가능한 데이터 기반의 선박엔진 고장진단 시스템을 제안한다.

선박은 크랭크축, 실린더 블록, 실린더 라이너, 피스톤, 배기터빈 과급기 등 다양한 부품들로 구성되어있으며, 이러한 부품과 함께 다양한 종류의 센서들로 구성되어있다. 데이터 기반 고장 분석의 경우 일반적으로 엔진을 구성하고 있는 센서로부터 수집된 데이터를 이용한다. 오늘날 선박의 대형화, 디지털화와 함께 조선 IT의 급부상으로 점점 디지털 선박으로 변화하고 있다. 이러한 디지털 선박은 여러 센서들로부터 계속된 정보에 데이터 융합(Data convergence) 기법을 적용하여 선박 내 데이터베이스를 통해 자동으로 데이터를 처리하며 관리할 수 있는 유비쿼터스 기술을 기반으로 한 선박이다[7]. 이와 같이 선박을 통해서 센서 데이터 수집이 가능하며, 최근에는 고성능 센서 및 데이터 로깅 기술을 적용하여 시스템 운용 상태를 실시간으로 저장 및 모니터링 할 수 있다. 그러나 데이터 수집 장치를 통해 수집된 데이터를 직접 결함 분석에 사용할 경우 주변 환경 및 데이터 저장의 특성상 용량 및 정확성 등의 여러 문제들이 발생하기 되므로, 기존의 Raw data(원본 데이터)의 정보 손실을 최소화하며, 중요 정보만을 추출할 수 있는 데이터 필터링 기술이 필요하다.

따라서 본 논문에서는 선박엔진 결함분석의 효율성을 높이기 위해 수집된 센서 데이터를 필터링 할 수 있는 방법을 제시하며, 필터링이 적용된 데이터를 기반으로 ‘선박엔진 고장분석시스템’을 개발한다. 개발된 시스템은 센서정보(각 센서 신호 표시), PCA(Principal Component Analysis), Correlation 분석, Entropy 정보를 제공하며, 또한 실제 선박 엔진 터미널 박스 데이터를 이용하여 고장의 원인을 분석하고, 그 효용성을 증명한다.

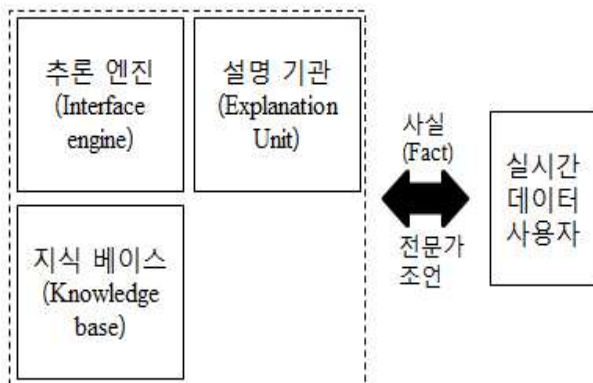
2. 관련 연구

2.1 기존의 선박엔진 고장진단시스템

기존의 선박엔진 고장진단시스템은 전문가 시스템, 신경망 방식, 상관관계 분석 등을 이용하는 방식이 주

를 이루고 있다[6].

전문가 시스템은 인공지능 연구의 한 분야로서 먼저 전문가가 가지고 있는 지식과 경험들을 컴퓨터 시스템에 입력시켜 지식베이스로 축적하고, 축적된 지식베이스를 이용하여 사용자의 요구를 추론하여, 그 결과를 사용자에게 전달해주는 방식이다[7]. <그림1>과 같이 지식베이스, 추론기관, 설명기관, 사용자 인터페이스(실시간 데이터 사용자)로 구성된다. 이는 전문가의 경험과 지식을 기반으로 컴퓨터에 지식 기반 정보를 저장해둬으로써 비전문가는 전문가의 도움 없이 컴퓨터에 저장되어 있는 정보를 이용할 수 있다. 전문가 시스템을 이용한 방식은 먼저 숙련된 전문가의 지식을 활용하여 데이터들 간의 상호연관성을 검토하고, 전문가에 의해 검토된 항목사이의 관계를 통계적 분석기법을 이용하여 정량화 한다. 그리고 이상상태를 감지하는 이상 감지 지식베이스와 이상 감지된 데이터와 서로 높은 연관관계를 가지고 있는 다른 데이터를 조사하여 종합적으로 진단할 수 있는 이상 진단 지식베이스를 구축하여 이상데이터를 자동으로 감지하고 진단할 수 있도록 한다[8]. 또한 일반적으로는 간단하게 전문가가 미리 각 센서데이터에 대한 정상 범위를 정해놓고, 실시간으로 데이터나 신호를 모니터링 하면서 정상 범위를 벗어났을 경우 알람이 발생할 수 있도록 하는 방법을 사용한다. 전문가 시스템을 통해서 전문적인 지식을 일관성 있게 이용할 수 있고, 지식을 쉽게 복사, 이동시킬 수 있다. 그러나 전문가로부터 얻어진 전문지식의 수준은 한계가 있기 때문에 범위를 벗어난 정보의 경우 해결하지 못하는 단점이 있다[6].



<그림 1> 전문가 시스템 구조[6]

위의 전문가 시스템에서 주로 쓰이는 통계적 방식은 신경망회로와 상관관계 분석으로 여러 가지 데이터를 축적하여 두고, 현재 엔진으로부터 실측 되는 정보와 기존의 데이터베이스의 정보를 비교하여 확률적으로 정상치에 대해 벗어난 정도를 판별하는 형태이다[9].

신경망회로 방식은 데이터의 정성적인 것과 정량적인 것을 동시에 처리하는 방법으로 부하와의 상대적인 중요도를 이용한다. 데이터의 학습을 통하여 자기 자신의 데이터와 부하간의 상대적 중요도를 출력할 수 있으며, 출력된 데이터가 정상 범위 이상일 경우 이상 데이터로 판별한다. 선박 엔진의 경우 어느 한 데이터가 표준 운전범위를 벗어났다고 하여 계통고장이라고 단정할 수 없으므로 표준 운전범위를 벗어난 데이터를 중심으로 상호 영향을 미치는 관련 데이터를 체계적으로 조사하여 계통 고장 가능성을 진단하는 것이 이상적이다[10].

상관관계를 이용한 방식 또한 신경망회로 방식과 비슷하다. 선박의 이동성으로 인하여 데이터 값들은 같은 부하조건이라도 외적환경변화에 따라 다르게 나타난다[6]. 따라서 선박을 구성하는 여러 조건들 간의 관계를 측정해야 할 필요가 있다. 선박엔진 시스템을 구성하는 다양한 부하의 변동에 따라 계측항목별 상관관계를 검토해 상관관계가 높은 항목과 낮은 항목으로 분류한다. 정상상태의 계측항목은 부하와 높은 상관관계를 가지고, 고장이 발생하였을 경우에는 기계의 특성이 정상이 아니기 때문에 낮은 상관관계를 가진다[11]. 이러한 점을 이용하여 정상상태에서 부하의 상태에 따른 상관계수의 크기를 측정하여 이상 데이터를 감지한다.

선박엔진 고장진단시스템의 경우 시스템을 구성하고 있는 여러 센서 데이터 중에서 주로 감지가 용이한 진동 센서 데이터를 이용하는 연구가 많다[12]. 이는 주로 FFT(Fast Fourier Transform)를 중심으로 여러 가지 확률적인 방법들을 적용하여 수치적 비교를 한다. 정상 데이터를 먼저 수집하고 정상상태의 엔진 고유 진동 주파수를 기준으로 각 채널을 통해 수집되는 엔진 진동 주파수를 비교하여 엔진의 고장유무를 판별한다[9]. 즉 기존에 실시간으로 수집되는 진동 센서 데이터를 FFT를 이용하여 주파수 영역으로 변환시킨 후 데이터의 값이 임계치(Threshold)를 넘기면 이상 데이터로 판별하는 방식을 이용한다. 그러나 샘플링 간격을 아주 작게 할 경우 주파수를 모두 비

교하는 것은 연산에 있어 비효율적이다. 이를 해결하기 위해 각 주파수 전체 파형을 분석에 이용하는 것이 아닌 파형 중 일부 이상이 있는 부분만 추출해내어 비교하는 방식을 사용하여 데이터베이스의 용량을 감소시킨다.

2.2 Raw data 축소를 위한 데이터 필터링

기존의 데이터 필터링 방법은 일반적으로 데이터들 사이의 거리를 이용한 분류와 각 센서 데이터들의 파형 이용하는 방식이 있다.

2.2.1 데이터 분류를 이용한 데이터 필터링

미리 훈련되어진 데이터를 통해 결함(Fault)과 무결함(Non-fault)으로 분류하고, 이 데이터들과 실제 들어오는 데이터들 간의 거리를 측정하여 결함과 거리가 가까운 데이터만 수집한다. 즉 결함과 무결함으로 이벤트의 특징을 미리 분류하여 놓고, 연속적으로 들어오는 데이터들을 고정된 윈도우로 묶은 후 결함 이벤트가 있는지 여부를 확인한다. 결함 이벤트가 있는 부분과 학습(Training)을 통해 기존에 저장되어있는 각 결함 패턴들과의 거리를 측정해 가장 근접한 거리의 데이터를 결함으로 판단하고, 그 종류를 파악한다[13]. 비슷하게 이전에 미리 시스템의 상태 값이나 제어 값을 변동시켜 민감하게 변하는 변수를 측정하여 무결함 기준모델을 만들고, 이 모델과 실시간으로 측정되는 값의 잔차(Residual)를 계산하여 잔차가 허용범위를 벗어난 상태의 데이터만 분석하는 방법이 있다[14].

이와 같이 클러스터링 되어진 Raw data 중 결함 기준 모델과 거리가 가까운 데이터들을 추출하여 결함 분석에 이용함으로써 결함진단 시스템의 시간적 공간적 효율성을 높인다.

2.2.2 신호의 파형을 이용한 필터링

신호의 파형을 이용하는 방식에는 연속적으로 들어오는 센서 데이터들의 신호를 이용하여 임계치 이상의 신호만 추출해내는 방법이나 입력되는 파형에 따라 결함을 분류하는 방식이 있다. <그림2>와 같이 연속적으로 입력되는 데이터 신호 중에서 결함인 부분의 파형을 'A', 'B', 'C'와 같이 특징별 Sub-sequence로 미리 분류한 후 실시간으로 입력되는 센서 데이터 신호를 윈도우를 통해 구간을 나누고, 구간으로 분류된 데이터와 미리 정해져 있는 Sub-sequence와의 충돌수를 비교하여 Weight를 측정한다. 측정된 Weight 중 가장 많은 주파수(Frequency)를 가지는 윈도우 구간을 결함 관련 데이터로 추출한다. 즉 모든 센서 데이터 신호를 수집하는 것이 아닌 기존의 결함 패턴 파형과 비슷한 신호만을 추출해내어 데이터의 저장 공간을 축소한다. 또한 연속적으로 들어오는 신호를 벡터 Segment로 나눌 수 있다. 예를 들어 <그림 3>과 같

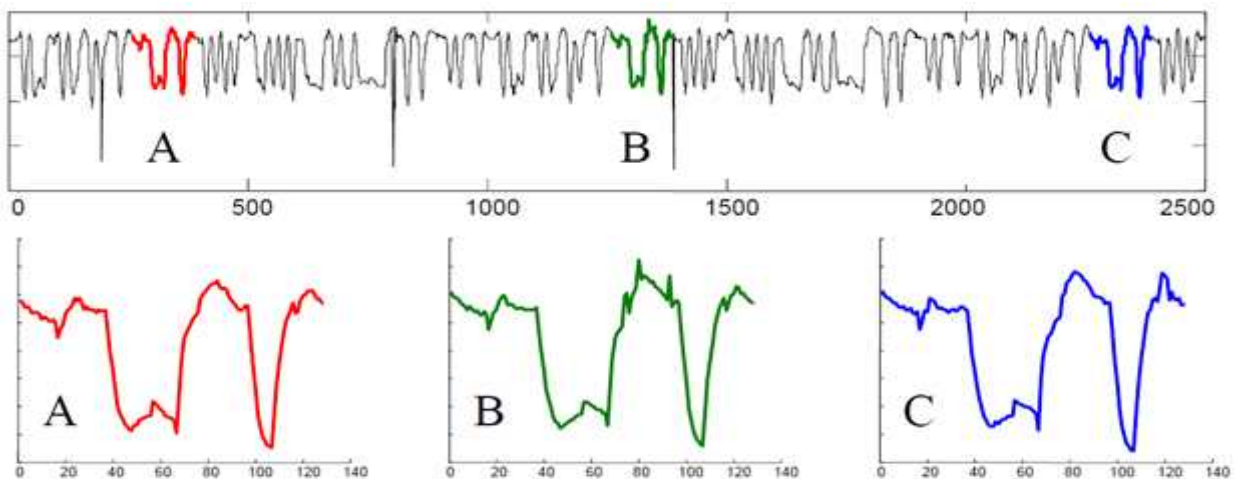
$$X = [-1, -2, -1, 0, 2, 1, 1, 0]$$

$$n = |X| = 8$$

$$\bar{X} = (\text{mean}(-1, -2, -1, 0), \text{mean}(2, 1, 1, 0))$$

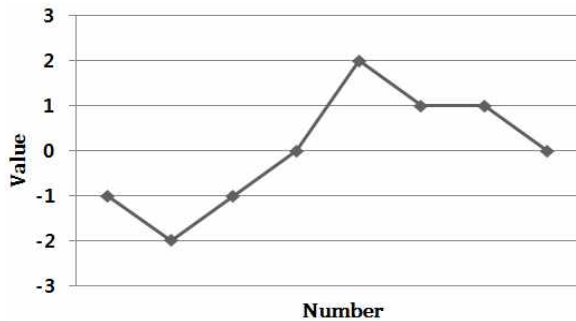
$$X = (-1, 1)$$

$$n = |X| = 2$$



<그림 2> 파형에 따라 Sub-sequence로 나누어진 센서 신호[15]

은 기존의 신호는 아래와 같이 표현된다.



<그림 3> 8개의 Points로 구성된 센서 신호[16]

즉, 기존의 신호를 8개의 Points(X)로 나누고, 고정된 윈도우 크기 '4'로 Segment를 나눈 후 나누어진 데이터들의 Mean value(평균)를 이용하여 8개의 Points를 2개(\bar{X})로 축소하였다. 축소된 Dataset을 이용하여 결합분석시스템에 사용하여, 분석의 효율성을 증가시킨다.

이와 같은 선박엔진 결합진단시스템의 효율적인 처리를 위한 데이터 필터링 방법을 기반으로 본 연구에서는 '데이터 이산화(Discretization)' 기법[14]을 적용하여 기존의 Raw data 중 중요 정보만을 추출하여 이벤트 로그로 변환하고, 축소된 이벤트 로그를 통해 결합을 분석하는 시스템을 개발하였다.

3. 데이터 필터링의 필요성 및 데이터 이산화

3.1 데이터 필터링의 필요성

일반적으로 결합의 경우 순간적인 기계결합의 경우보다는 서서히 기계적인 소모에 의한 결합이 발생하는 경우가 더 빈번하다. 즉 결합을 위해 수집된 데이터는 짧게는 1개월 길게는 수년 동안 수집된 데이터로 그 양이 매우 방대하다. 센서의 종류 또한 시스템을 구성하는 온도센서, 습도센서, 가속도 센서 등과 같이 매우 다양하기 때문에 각 센서의 정상(Normal state) 범위 및 단위가 다르므로 전체 센서들에 대한 통일된 기준이 없어 결합 분석의 정확도 측면에서 떨어진다. 데이터의 양, 기준이 없는 다양한 센서뿐만 아니라 센서로부터 축적된 데이터는 주변 환경의 영

향으로 잡음(Noise), 결측치(Missing value), 중복 정보(Redundancy)를 포함한다. 이러한 이상 데이터는 센서에서 측정된 값이 데이터 저장 장치로 옮겨질 때에 발생하는 오류이다. 즉 Raw data를 결합분석에 직접 사용한다면 저장 공간적인 측면과 연산을 위한 시간적 측면에서 비효율적이고, 정확도와 신뢰도가 떨어지므로 이를 위한 데이터 필터링 방법이 적용되어야 한다.

이양지 외[17]에서는 이와 같은 Raw data의 필터링을 위해 통계적 분석 기법인 PCA와 Correlation 분석을 적용하였다. PCA는 기존의 데이터에 많은 양의 변수가 있을 때 그 변수의 양보다 작은 주성분으로 전체 변동 중 상당 부분을 설명할 수 있으며, 축소된 주성분을 통해 자료를 해석하여 Raw data에 나타나지 않은 새로운 관계를 찾아낼 수 있다. 이를 통해 선박엔진 결합진단시스템에서 여러 변수(센서)들 중에서 주요 변수(센서)를 추출한다. Correlation 분석에서는 Auto-Correlation과 Cross-Correlation 두 가지 분석이 가능하다. Auto-Correlation 분석으로 각 센서 내의 신호에서 특정한 주기성이 발생하는지 확인하고, Cross-Correlation 분석으로 여러 센서들 간의 주기성을 확인한다. Correlation 분석을 통해서 여러 변수들 중에 주기성을 가지는 변수들을 추출하여 변수의 차원을 줄일 수 있다.

본 연구에서는 이전 연구에서 수행된 통계적 분석 기법 이외에 패턴 분석 기법인 데이터 이산화를 통해 기존의 Raw data를 특정 이벤트 로그로 변화하여 데이터 필터링하는 방법을 제안하며, 필터링 된 이벤트 로그를 선박엔진 결합진단시스템에 적용하여 결합의 원인을 효율적으로 추출해낼 수 있도록 한다.

3.2 데이터 이산화

데이터 이산화란 수치형 속성의 도메인을 여러 개의 분절된 부분으로 나누는 것을 의미한다[13]. 이산화를 하는데 있어 중요한 기준은 크게 정보의 손실을 최소화 하는 것, 그리고 미지의 데이터에 대한 범주 값의 일반성을 최대화 하는 것의 두 가지로 나뉜다. 그러나 이 두 가지는 서로 상충되는데, 그것은 이산화 구간의 수가 많아질수록 정보의 손실은 적어지지만, 각 구간의 학습 집합(Training set)에 대한 종속성이 증가(Over-fitting)하여, 그 일반성이 감소하기 때문이

다. 따라서 원래 데이터의 정보를 유지하면서 동시에 일반화된 대표 값을 산출할 수 있는 적절한 수준의 이산화 구간을 찾는 것이 이산화 알고리즘의 주된 목적이다[13].

데이터 이산화는 데이터의 특징에 따라 <표1>과 같이 분류된다. 도메인 구간의 일부만 사용하거나 전체 도메인 구간을 사용하는지의 여부에 따라 전자는 지역적(Local) 후자는 전역적(Global) 방법으로 분류될 수 있으며, 목적속성의 사용 여부에 따라 교사(Supervised)와 비교사(Unsupervised)로 구별 가능하다. 또한 이산화 과정 중 정지 시점(Stop point)을 직접 사용자가 입력하는 Direct 방법과 Indirect 방법으로 분류가능하며, 이산화를 통해 각 데이터 속성을 합치는지(Merge) 분리(Split)하는지와 훈련 과정에서 이산화 될 필요가 있는 속성이 발견될 경우에만 분할하는 동적(Dynamic)인 방법과 데이터 집합의 모든 수치형 속성을 이산화 시키는 정적(Static)인 방법으로 분류할 수 있다. 본 논문에서는 <표1>과 같은 다양한 이산화의 방법 중 대표적으로 사용되는 Equal-width, Equal-frequency, Entropy, Zeta 알고리즘에 대해 간략하게 소개한다.

'Bin'이란 현재 주어진 데이터 집합에서 인스턴스가 최소한 한번 이상 나타난 좌표의 값으로 정의한다[18].

'Binning'이란 Raw data에서 일정 한 크기의 Bin이라는 공간으로 분할하는 것이며, Equal-width 방법과 Equal-frequency 방법이 이에 속한다. Equal-width 방법은 각각 속성의 도메인 구간을 일정한 간격, 일정한 인스턴스 수를 갖는 구간으로 분할하며, 이와 비슷하게 Equal-frequency 방법은 전체 데이터의 속성 값에서 같은 데이터의 개수로 일정구간의 Bin으로 분류하는 것을 말한다[13].

Entropy를 이용한 이산화 과정은 목적 속성을 고려하며, 데이터의 불확실성에 대한 정보를 기반으로 한다. Entropy란 확률 정보(Probability)를 불확실성 정보(Uncertainty)로 변환해 주는 일종의 정보 변환 함수이며[19], 그 값이 높을수록 불확실한 정보이다. 또한 모든 목적 속성의 분포가 균일할수록 Entropy는 큰 값을 가지고, 어떤 하나의 목적 속성의 비율이 다른 목적 속성의 비율보다 높아질수록 작은 값을 가진다. 이와 같이 Entropy는 Binning 방법과 다르게 목적 속성 값과 연관되어 있는 정도를 측정할 수 있다.

Entropy와 같이 많이 사용되는 Zeta 알고리즘은 각 Class(변수)와 특징(목적 속성)과의 연관성 정도를 이용한다. Entropy에서는 Entropy 값(불확실성의 정도)이 낮은 것을 채택하는 반면에 Zeta 알고리즘에서는 각 Class와 특징과의 관계 정도가 높은 것을 채택

<표 1> 데이터의 특징에 따른 이산화 방법[14]

Methods	Global/local	Supervised/unsupervised	Direct/incremental	Splitting/merging	Static/dynamic
Equal-width	Global	Unsupervised	Direct	Splitting	Static
Equal-frequency	Global	Unsupervised	Direct	Splitting	Static
1R	Global	Supervised	Direct	Splitting	Static
D2	Local	Supervised	Indirect	Splitting	Static
Entropy	Local	Supervised	Indirect	Splitting	Static
Mantaras	Local	Supervised	Indirect	Splitting	Static
ID3	Local	Supervised	Indirect	Splitting	Dynamic
Zeta	Global	Supervised	Direct	Splitting	Static
Accuracy	Global	Supervised	Direct	Splitting	Static
ChiMerge	Global	Supervised	Indirect	Merging	Static
Chi2	Global	Supervised	Indirect	Merging	Static
ConMerge	Global	Supervised	Indirect	Merging	Static

하여 Cut-point로 결정한다[20].

앞서 설명한 데이터 이산화의 3가지 방법인 Equal-frequency, Entropy, Zeta 방법 중 연산 시간의 경우 Equal-frequency 방법과 Zeta 알고리즘이 이산화를 수행하는 속도가 매우 빨랐으며, Entropy는 수행하는 속도가 비교적 느린 것으로 분석된다. 그러나 에러율(Error rate)의 경우 Entropy가 가장 낮으며, Equal-frequency와 Zeta 알고리즘의 경우 높게 나타났다[14].

결함분석을 위해 수행되는 데이터 필터링 과정에서 효율성을 위해 연산 속도 및 저장 공간적인 측면을 고려해야 하지만 원본 데이터의 정보를 손실하지 않는 정확성 및 신뢰성 또한 고려되어야 한다. 따라서 본 논문에서는 데이터의 특징 및 연산 속도, 그리고 정확도의 측면을 고려해 이산화 방법 중에서 Equal-frequency와 Entropy 방법을 적용한다. 이를 통해 기존의 데이터 수치 값을 속성 값으로 대체하여, 방대한 양의 Raw data에서 중요 정보만을 추출한다. 뿐만 아니라 필터링 된 이벤트 로그를 기반으로 결함의 원인을 파악할 수 있는 이벤트 기반 선박엔진 결함진단시스템을 개발한다.

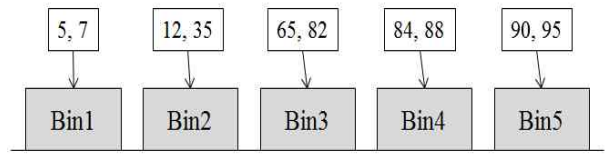
4. 선박엔진 결함진단시스템

4.1 이산화 기반 Raw data 변환 (이벤트 로그 생성)

결함분석을 위해 수집된 원본 데이터는 분석의 시간적 공간적 효율성을 위해 필터링 과정이 필요하다. 즉 Raw data에서 중요 정보만을 추출하는 과정이 필요하며, 본 논문에서는 원본 데이터 정보를 유지하면서 동시에 일반화된 대표 값으로 산출할 수 있는 이산화 과정을 적용하여 이벤트 로그로 변환한다.

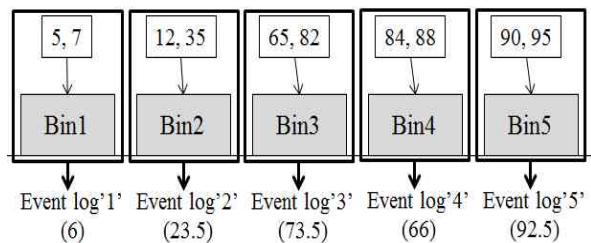
<그림 4>는 Equal-frequency binning 방법을 설명한 예로 원본 데이터를 정렬한 후 똑같은 개수의 인스턴스(각 데이터)로 묶어 'Bin'이라는 박스로 묶는다. 묶는 인스턴스의 수는 사용자가 직접 지정하는 형식이며, 묶은 인스턴스들의 대표 값은 데이터 변환의 특성 중 'Smoothing'이라는 특성을 고려해야 한다. 'Smoothing'이란 데이터로부터 잡음을 제거하기 위해 데이터 추세에 벗어나는 데이터를 추세에 맞게 변환

※Sorted data(Sensor 'n')
=[5 7 12 35 65 82 84 88 90 95]



<그림 4> Equal-frequency binning 방법

하는 방법으로 Bin에 속해있는 인스턴스들을 'Mean value', 'Median value', Boundaries value' 등으로 대체하는 방법을 사용한다. 예를 들어 <그림4>와 같이 기존의 데이터를 정렬한 후 '2'라는 고정 크기로 인스턴스 2개씩 Bin으로 묶는다. 여기서 '2'라는 Frequency는 본 시스템에서는 '윈도우'라는 개념으로 '2'라는 윈도우로 묶는다는 의미와 같다. 총 10개의 데이터를 '2'개씩 묶으면 원본 데이터는 5개의 Bin으로 축소된다. 각 Bin 내의 데이터 Mean value를 대표 값으로 하면, 기존의 10개의 데이터는 <그림 5>와 같이 [6, 23.5, 73.5, 66, 92.5]의 5개의 이벤트 로그로 축소된다.



<그림 5> Equal-frequency binning 방법을 통한 이벤트 로그 변환

실제 센서 데이터를 이용할 경우 각 센서 데이터를 사용자가 지정한 고정된 윈도우 크기로 묶어 Mean value를 이용하여 데이터 축소가 가능하다. Binning을 통한 이산화의 경우 적절한 윈도우 크기를 결정해야 한다. 윈도우 크기를 작게 할 경우 세분화하여 정확성이 증가하지만 Overfitting 현상이 발생하여 데이터가 한쪽으로 편중되는 현상이 일어나고, 윈도우 크기를 크게 할 경우 데이터의 양이 축소되지만 정확도적인 측면에서 떨어진다. 따라서 적절한 윈도우 크기를 고려해야 한다.

또 다른 방법인 Entropy를 이용한 이산화 방법은

각 센서 데이터들의 불확실성 정도를 측정하여 Cut-point를 결정하고 속성 값을 그 구간의 대표 값으로 설정한다.

<표 2>와 같이 정렬된 원본 데이터에서 각 데이터 사이의 경계를 기준으로 Entropy를 계산한다. 이를 통해 <그림 6>과 같이 차례대로 Entropy를 측정하여 [66.5, 70.5, 73.5, 77.5, 80.5, 84.5]와 같은 Cut-point를 도출할 수 있다. 결정되어진 각 Cut-point 내의 속성 값을 각각의 이벤트 로그로 변환한다. 즉 <그림 6>에서 센서 값이 '64-65'에 속하는 데이터는 'Event log1'이 되고, '68-70'에 속하는 데이터는 'Event log2'로 정의한다. 이와 같이 다양한 데이터 범위의 값을 각 속성 값에 속하는 대표 속성 값으로 대체함으로써 데이터의 차원 축소가 가능하다.

Entropy의 경우 연산 수행 중에 연산을 중단 할 수 있는 Stop-point의 기준이 필요하다. Stop-point는 주로 Entropy가 어느 특정 임계치 값보다 작은 값이 나오면 정지하는 방법을 사용하며, 이 임계치는 사용자

수 있다. 결함 발생 이전에 수집된 데이터는 시간적 간격이 모두 다를 수 있고, 그 양이 상당하며, 속성 값 또한 다양하다. 따라서 위와 같은 이산화 과정을 통해 연속형 데이터에서 필요한 중요 정보를 추출하고, 기존의 데이터를 그 정보로 대체하여 이벤트 로그로 변환한다. 이를 위한 Binning을 방법과 Entropy를 이용한 방법은 그 내용이 상이하며 두 가지 경우에 대한 분석이 필요하다.

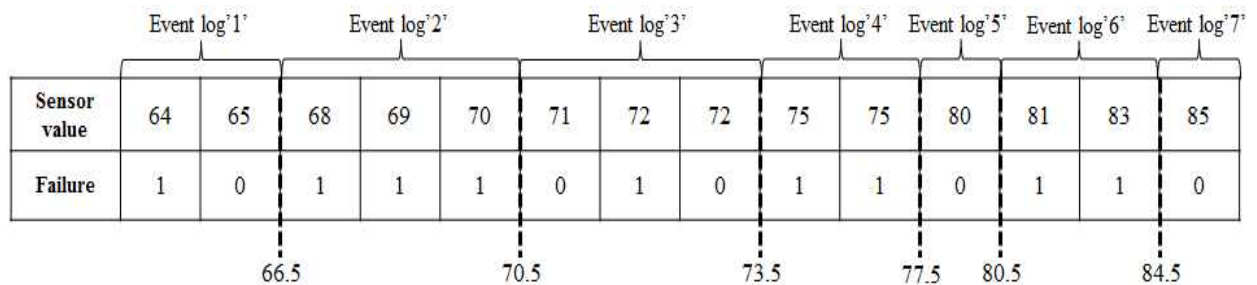
4.2 결함 패턴 분석

본 논문에서는 Matlab을 통해 선박엔진 데이터를 이용하여 다양한 정보를 얻을 수 있는 소프트웨어를 개발하였다. <그림 7>이 시스템의 메인화면으로 '센서정보', 'PCA 분석', 'Correlation 분석', 'Entropy' 그리고 마지막으로 결함정보를 추출할 수 있는 '패턴분석' 기능을 포함하고 있다.

<그림 8>은 센서 정보를 확인할 수 있는 기능으로

<표 2> Entropy 기반 이산화를 위해 정렬된 원본 데이터 (Failure가 1일 경우: 결함, Failure가 0일 경우: 결함이 아님)

Sensor value	64	65	68	69	70	71	72	72	75	75	80	81	83	85
Failure	1	0	1	1	1	0	1	0	1	1	0	1	1	0



<그림 6> Entropy 기반의 이산화를 통한 이벤트 로그 추출

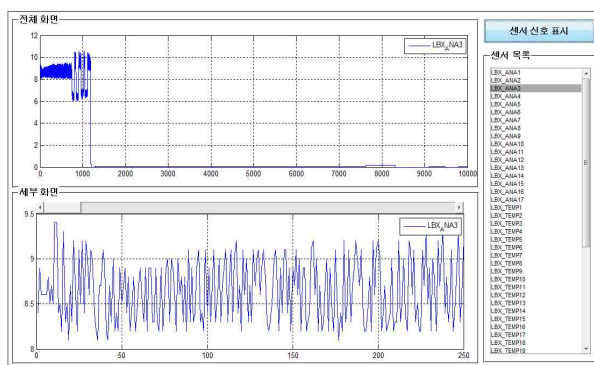
에 의해 입력되어진다. 이는 사용자의 주관적인 의견이 반영되며, Binning 방법의 윈도우 크기 결정과 같이 적절한 임계치의 Entropy 값 결정할 수 있어야 한다.

결함분석을 위해서는 결함이 발생하기 이전까지의 모든 데이터를 사용해야 진단에 대한 정확성을 높일

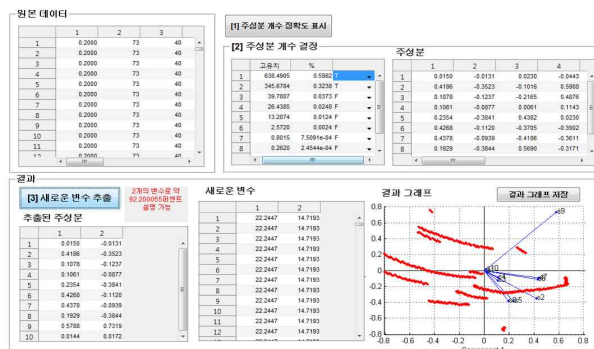
센서 목록에서 선택되어진 센서의 과형을 확인할 수 있으며 전체적 신호뿐만 아니라 세부적으로 신호의 형태를 확인할 수 있다. <그림 9>는 PCA 분석으로 앞서 언급한바와 같이 여러 종류의 센서 데이터들 중 중요 센서 추출이 가능하다. 고유치(Eigen value)를 사용자가 선택할 수 있게 하여, 각 정확도에 따른



<그림 7> 선박엔진 결함진단시스템(메인화면)

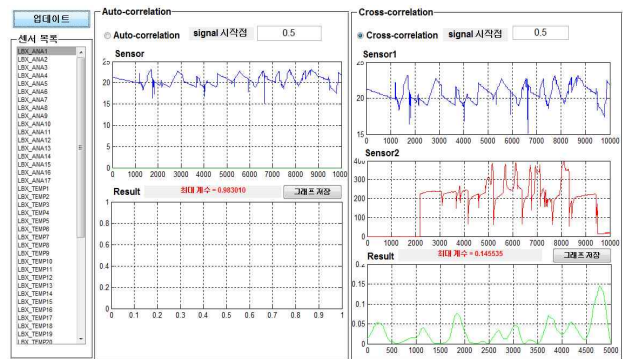


<그림 8> 선박엔진 결함진단시스템(센서정보)



<그림 9> 선박엔진 결함진단시스템(PCA분석)

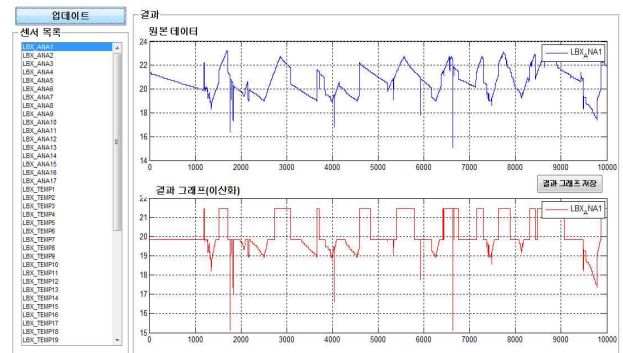
분석이 가능하다. 또한 <그림 10>은 Correlation 분석 기능으로 'Auto-correlation'과 'Cross-correlation' 기능을 제공한다. 'Auto-correlation' 분석으로 선택되어진 하나의 센서 내에 어떠한 주기성이 존재하는지 그래프와 그 값을 추출할 수 있고, 'Cross-correlation'을 통해 두 센서간의 주기성을 확인할 수 있다. <그림 11>은 Entropy 분석으로 기존의 원본 센



<그림 10> 선박엔진 결함진단시스템(Correlation 분석)

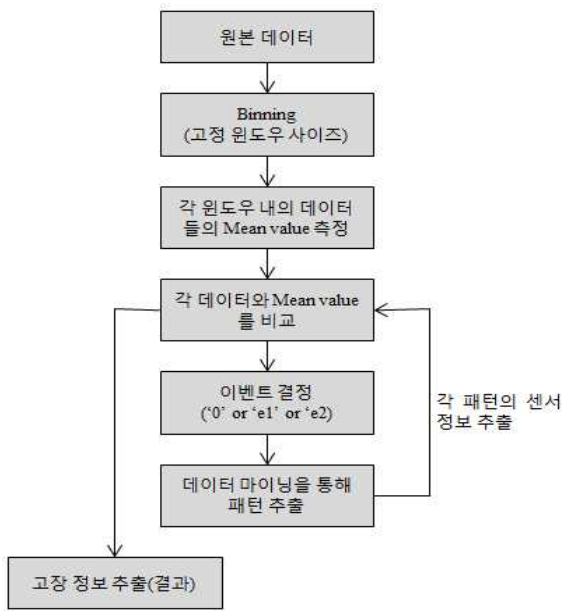
서 데이터를 이산화 형태의 그래프로 변환시켜 확인이 가능하다. 데이터베이스와의 연동을 통해 수행되며, 결과 그래프를 저장하여 보관할 수 있는 기능 또한 수행된다. 위와 같은 시스템의 기능을 통해 결함분석 이전에 수행되어야할 각 센서 데이터에 대한 정확한 해석이 가능하다.

이산화를 통한 이벤트 로그 변환 과정을 실제 테스트 중인 선박엔진 센서 데이터에 적용하였으며, 변환된 이벤트 로그를 통해 결함원인을 파악하기 위한 프로토콜을 개발하였다. 실험에 사용된 데이터는 56개의 센서로 구성되었으며, 각 센서에 약 4만5천개의 데이터(총56×4만5천)를 포함한다.



<그림 11> 선박엔진 결함진단시스템(Entropy 분석)

Equal-frequency를 이용한 결함진단시스템의 흐름도는 <그림 12>와 같으며 소프트웨어의 형태는 <그림 13>과 같다. 개발된 시스템에서는 결함원인 패턴을 추출하기 이전에 여러 종류의 센서 중에 결함을 결정할 기준을 사용자가 결정한 후 그 기준에 맞게 데이터를 가공한다. 또한 선택된 센서를 기준으로 결함 발



<그림 12> 이산화를 이용한 결함진단시스템의 흐름도

생 수에 대한 정보를 제공한다. 실험 데이터에서는 고온냉각수 온도 측정 센서를 결함 판별 센서로 선택하였을 경우 38개의 결함이 발생하였다. 가공된 데이터

를 사용자가 결정한 윈도우 크기로 Binning을 한 후 각 윈도우 내의 데이터들의 Mean value를 구한다. 측정된 Mean value를 Binning된 각 센서 데이터들의 정상 상태로 정의하고, 이 값과 기존의 데이터와의 관계 특징을 통해 이벤트를 결정한다. 현재의 시스템은 3가지의 이벤트 로그로 분류하였다. 각 데이터가 정상 상태보다 클 경우 'e1(Event log1)', 작을 경우를 'e2(Event log2)', 그리고 정상 상태 내의 데이터일 경우 '0' 값을 지정한다. 실험 데이터를 이용하여 윈도우 크기를 100으로 설정하고, 정상 범위를 70(%)로 설정하였을 경우 원본 데이터는 <표 3>과 같은 형태를 가지며, 윈도우로 묶어진 데이터와 정상 상태와의 비교를 통해 <표 4>와 같은 이벤트 로그 테이블로 변경된다. 표시된 데이터의 형태는 'e_센서번호_이벤트 특징 (Mean value와의 비교)'로 구성된다. 이렇게 이벤트 로그로 변환된 데이터는 데이터 마이닝에 적용하기 위해 각 윈도우 내의 센서들(총56개의 센서)의 데이터를 하나로 묶어 새로운 이벤트 로그로 변환한다. 이를 통해 38개의 결함에 대한 <표 5>와 같은 새로운 이벤트 로그가 완성되며, '----'은 데이터가 정상 범위 내의 데이터로 '0' 값을 나타낸다. <표 5>와 같은 이벤트 로그를 임계치 2로 설정하여 데이터 마이닝 기법인 FP-Tree 알고리즘에 적용하면[21] <표 6>과 같은

[1] 고장 기준 결정

Name	Description	Type	Normal(Min)	Normal(Max)	Data(Min)	Data(Max)	Alarm(>)	Failure
1	LBX_ana1	STARTING AL...	float	0	30 9000	0	30 9000	NaN F
2	LBX_ana2	L.O INLET PR...	float	4	5	0	8 6000	NaN F
3	LBX_ana3	F.O INLET PR...	float	0	16	0	16	NaN F
4	LBX_ana4	CHARGE AIR	float	0	6	0	6	NaN F
5	LBX_ana5	H.T WATER ...	float	70	85	0	87	NaN T
6	LBX_ana6	L.T INLET TE...	float	30	40	0	67	45 F
7	LBX_ana7	F.O INLET TE...	float	0	151	0	151	NaN F
8	LBX_ana8	L.O INLET TE...	float	60	70	0	76	80 F
9	LBX_ana9	ENGINE RPM	float	0	2000	0	2000	NaN F
10	LBX_ana10	T/C RPM	float	0	60000	0	60000	NaN F

[3] 고장 정보 추출 (결과)

이벤트 시퀀스	이벤트 번호	센서 이름	비교	기준값	정상(최소값)	정상(최대값)
#1	EVENT ID 28	CHARGE AIR PRESS	<	1.33	0	6
		T/C RPM	<	6954.86	0	60000
		LBX CYL. Temp A5	<	120.02	0	523
EVENT ID 24	CHARGE AIR PRESS	LBX CYL. Temp A7	<	127.43	0	535
		LBX CYL. Temp A7	<	127.43	0	535
EVENT ID 5	ENG RUN HOUR(MIN)	<	9.17	0	20	
EVENT ID 4	ENG RUN HOUR(MIN)	>	3.34	0	20	

[2] 이벤트 로그 수집

Event log 결정 기준: Mean Variance

Window 내의 데이터 수: 100

센서 정상상태 허용범위(%): 70

고급 설정 (window 반복 사용)

고장정보 추출 옵션 (default): Different, Same

임계치(반번수): 2

고장 정보 패턴

	1	2	3	4
1	28-2	[28 4]-2	[28 24 4]-2	[28 24 5 4]-2
2	58-2	[28 24]-2	[28 5 4]-2	
3	24-2	[28 5]-2	[28 24 5]-2	
4	48-3	[24 4]-2	[24 5 4]-2	
5	23-3	[24 5]-2		
6	5-3	[86 58]-2		
7	66-5	[61 49]-3		
8	4-5	[23 4]-2		
9	7-2	[5 4]-3		

<그림 13> 선박엔진 결함진단시스템(이벤트 로그 추출 및 결함분석)

<표 3> 선박엔진 센서 데이터(원본 데이터)

No	Sensor1	Sensor2	...	Sensor56
1	89.7	2	...	0.3
2	10.3	2	...	0.4
...
45000	35.1	3	...	0.2

<표 4> 원본 데이터의 이벤트 로그 변환

윈도우	Sensor1	Sensor2	...	Sensor56
w1	e12	e21	...	e561
w2	e11	e21	...	e562
...
w450	e12	e22	...	e562

<표 5> 결함(Failure)에 따른 이벤트 로그 변환

Failure	w1 (윈도우1)	w2	...	w450
F1	'A093'	'A003'	...	'----'
F2	'A005'	'A002'	...	'A0093'
...
F38	'A192'	'A002'	...	'----'

<표 6> FP-Tree를 이용한 패턴 분석

No	1-length	2-length	3-length	4-length
1	[28]-2	[28 4]-2	[28 24 4]-2	[28 24 5 4]-2
2	[58]-2	[28 24]-2	[28 5 4]-2	
3	[24]-2	[28 5]-2	[28 24 5]-2	
4	[49]-3	[24 4]-2	[24 5 4]-2	
...		
10	[61]-24			

패턴을 추출할 수 있다. 'Length'라는 의미는 패턴내의 이벤트 수를 말하며, '[패턴]-빈번수(Frequency)'와 같이 표현된다. 본 시스템에서는 각 센서들의 특징에 관한 연관관계에 의해 결함이 발생하는지에 목적이 있음으로 모든 패턴들 중에서 길이가 가장 긴 패턴을 중요 정보로 선택한다. 즉 [28 24 5 4]라는 이벤트 로그가 중요 정보로 선택되어지며, 이 패턴은 총 2번 발생한다는 것을 알 수 있다. 실질적으로 결함정보시스

템에서는 이러한 이벤트 로그뿐만 아니라 이벤트로그가 포함하고 있는 센서 정보를 제공한다. 결과는 <표 7>과 같으며 각 Event ID에 대한 센서 데이터의 특징 정보를 제공한다. 또한 결과를 저장하여 분석할 수 있는 기능을 제공한다.

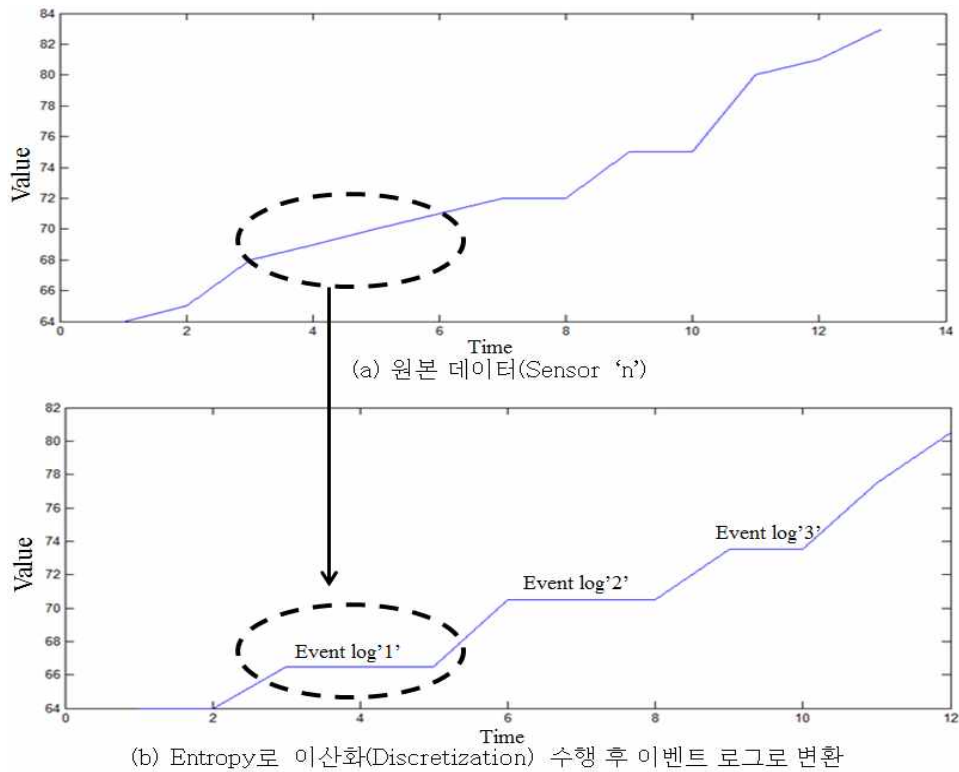
<표 7> 선박엔진 결함원인분석 결과

No	EventID	센서이름	비교	기준값
#1	Event ID 28	공기 센서 압력	<	6
		RPM	<	60000
		실린더 A 온도	<	523
		실린더 B 온도	<	535
	Event ID 24	공기 센서 압력	<	6
		실린더 B 온도	<	535
	Event ID 5	엔진 구동시간(min)	<	20
	Event ID 4	엔진 구동시간(min)	>	20

<그림 14>는 원본 센서 데이터(a)를 Entropy 과정을 적용한 결과(b)이다. 이와 같이 Entropy를 통해 연속형 데이터를 이산화 데이터로 변환할 수 있다. Entropy를 통해 변환되어진 데이터에서 4장에서 설명한 바와 같이 각 구간을 'Event log1' 등과 같은 형태의 이벤트 로그로 변환하여 데이터의 속성 값을 축소시킨다. 이를 통해 Binning 방법과 같이 이벤트 로그로 변환시키고 데이터 마이닝 기법을 통해서 패턴을 추출된다.

Entropy의 Stop-point를 사용자가 직접 지정해주는 것은 주관적으로 판단하여 본 시스템에서는 Entropy의 연산을 최종 데이터까지 모두 수행하도록 하여 원본 데이터를 이벤트 로그로 변환하였다. 그 결과 Binning을 이용한 이벤트 로그 변환 방법에서는 연산이 비교적 빠르게 수행될 수 있었지만 Entropy를 이용한 이벤트 로그 변환의 경우 최종 데이터까지 모두 수행함으로써 연산 속도가 상당히 느린 결과를 얻었다.

위와 같은 실험을 통해서 <표 8>과 같이 기존의 결함원인분석 방법과는 다른 효과를 얻을 수 있다. 즉 Equal-frequency binning 방법과 Entropy 방법을 통해 기존의 데이터를 이벤트 로그로 변환한 후 선박엔진 결함을 분석할 경우 데이터를 저장하는 공간적 측면과 연산을 위한 시간적 측면에서 효율적이며, 기존



<그림 14> Entropy를 이용한 이벤트 로그 변환

의 결함 분석 방법인 전문가 시스템 및 통계적 분석과는 달리 결함이 발생하기 이전에 각 센서들이 어떤

동작을 일으키는지에 대한 패턴들을 알려줌으로써 결함의 원인까지 추출 가능하다.

<표 8> 기존의 결함분석 방법과 이벤트 기반 결함분석 방법 비교

분석방법	특징
전문가 시스템 및 통계적 분석	<ul style="list-style-type: none"> 전문가로부터 얻어진 전문지식의 수준은 한계가 있기 때문에 범위를 벗어난 정보의 경우 해결하지 못하는 단점이 있음[6] 미리 정해놓은 센서 데이터의 범위를 넘길 경우 결함으로 판단하여, 결함의 유무를 판단하며 근본적인 원인은 추출할 수 없음
이벤트 기반 결함분석	<ul style="list-style-type: none"> 결함이 일어나기 이전에 각 센서로부터 수집된 데이터를 이용함으로써 결함 발생 이전에 각 센서들이 어떠한 동작 패턴이 발생하였는지 알 수 있음 이 패턴들을 분석함으로써 각 결함에 대한 원인이 추출 가능하며, 추출된 패턴을 바탕으로 결함 예측 및 진단이 가능함

5. 결론 및 향후 연구

본 논문에서는 이산화를 이용하여 데이터 필터링을 수행하였고, 변환된 이벤트 로그를 이용하여 결함진단 시스템을 개발하였다.

선박엔진의 결함뿐만 아니라 대부분의 시스템의 결함은 일반적으로 운용중의 기계적 결함에 의해 발생하는 것으로 순간의 일부 데이터가 아닌 결함 이전의 모든 데이터를 수집하여 수집된 데이터의 흐름을 분석해야 한다. 시스템의 운용은 단기간에 일어나는 것이 아닌 장시간동안 일어난 후 수집된 데이터이기 때문에 데이터의 양은 매우 방대하고 잡음, 중복정보가 포함되어 있으므로 직접 사용하기에는 정확도적인 측면과 데이터의 공간, 연산 시간 등의 측면에서 비효율적이다. 따라서 효율적인 결함분석시스템을 위해 데이

터 이산화의 개념을 적용하여 Raw data에서 중요 정보를 추출하여 특정 이벤트 로그로 변환한다. 실험에는 Equal-frequency binning과 Entropy를 통한 이산화 알고리즘을 적용하였다. Equal-frequency 방법은 고정된 윈도우(각 Bin 내의 데이터 수)로 데이터를 묶은 후 ‘Smoothing’ 처리를 위해 Mean value를 측정해 각 센서의 정상 범위로 정한다. 측정된 정상 범위와 실제 데이터를 비교하여 특징에 따라 이벤트 로그로 변환하고, 변환되어진 이벤트 로그를 각 결합에 대한 이벤트 로그 형태로 다시 한 번 변환한다. 변환되어진 이벤트 로그를 데이터 마이닝 기법인 FP-Tree 기법에 적용하여 특정 패턴 값을 얻고, 이 패턴에 대한 센서의 이름 및 특징 정보를 추출할 수 있으며, 추출된 정보를 결합의 원인으로 판단한다. 이와 같이 Binning을 통해 기존의 방대한 양의 센서 데이터를 이벤트 로그로 변환함으로써 데이터 필터링이 가능하였다. 빈번 패턴을 추출하기 위해 일반적으로 사용되어지고 있는 연관 마이닝 기법[22-23] 중 FP-Tree 기법을 적용하였을 때 원본 데이터(56개×4만5천개)의 경우 거의 시스템이 멈추어 버리는 것과 같은 연산 속도였으나 Binning을 통해 이벤트 로그로 변환한 데이터를 이용하였을 때에는 20초 내외의 아주 빠른 속도로 결과를 추출해 낼 수 있었다.

Entropy의 경우 기존의 연속형 데이터를 각 구간마다 Entropy를 계산하여 가장 낮은 Entropy 값을 가지는 구간을 선택하여 그 구간을 대표 속성 값으로 대체한다. 이 과정을 반복하여 축소된 속성 값을 가지는 이벤트 로그를 얻을 수 있다. Entropy를 통한 이산화를 적용할 때 보통 Stop-point는 보통 사용자에게 의해 정의된 수치 이하일 경우를 사용한다. 그러나 본 시스템에서는 Entropy 연산을 Stop-point 없이 최종 데이터까지 모두 수행하였다. 이로 인해 이벤트 로그의 종류가 많아지면서 연산 수행 결과가 상당히 늦게 도출되었다.

이번 연구를 기반으로 향후 원본 데이터를 이벤트 로그로 변환 시 정확성의 정도를 높이는 것에 초점을 맞추는 것을 목표로 한다. Binning 방법의 경우 사용자가 윈도우 크기와 정상 상태의 퍼센트(%) 정도를 결정하게 되는데 이는 주관적인 의견이 포함될 수 있다. 윈도우 크기를 크게 할수록 연산 속도는 빨라지지만 정확성이 떨어지고 반면에 윈도우 크기를 작게 할 경우 한쪽으로 편중되는 Overfitting 현상이 발생한다.

현재까지의 연구에서는 m-Cross validation과 같은 방법을 통해 수행되고 있으나 연산 속도가 늦어 쉽게 적절한 윈도우 크기와 정상 상태의 퍼센트(%) 정도를 결정할 수 있는 방법에 대한 연구가 필요하다. 또한 실제 현장에서 사용을 위해 이벤트 로그를 이용한 결과와 실제 결과와의 비교를 통한 정확한 효용성 평가를 위한 연구가 필요하며, 이 효용성을 높여 실제 현장에 적용될 수 있도록 해야 한다.

참 고 문 헌

- [1] 배정철, 강규홍, 하연철, 황경수, 이호상, 김성민, “IEEE 802.15.4 기반의 선박엔진 모니터링 시스템에 관한 연구”, 한국마린엔지니어링 학회, pp.383-385, 2010.
- [2] 김달현, “선박엔진용 실시간 고장진단시스템의 구현에 관한 연구”, 학위논문(석사), 2001.
- [3] 이순섭, 이동곤, “손상선박의 안정성 평가를 위한 통합 시스템 개발”, 한국 CAD/CAM 학회 논문집, vol.13, pp.227-234, 2005.
- [4] 이경호, 김화섭, 한선우, 박종현, 오준, “선박의 안정성 평가를 위한 네트워크 기반의 시뮬레이션 시스템 프레임 워크”, 한국 CAD/CAM 학회 논문집, vol.10, pp.356-364, 2005.
- [5] R.Isermann, “Fault-Diagnosis Systems: introduction from Fault Detection to Fault Tolerance”, Springer, Germany, pp.475, 2006.
- [6] 이기동, 정광교, 김원래, “소형선박용 고속디젤기관의 고장 예측 진단 시스템 연구”, 선박안전기술공단 선박안전, vol.26, pp.24-35, 2009.
- [7] 배정철, 전구양, 하연철, 류한성, “선박 엔진관리를 위한 무선 센서 네트워크 기초 특성 연구”, 선박안전기술공단 선박안전, vol.27, pp.53-62, 2009.
- [8] 김영일, 오현경, 천행춘, 유영호, “통계적 분석기법을 이용한 디젤기관의 고장진단 방법에 관한 연구”, 한국마린엔지니어링 학회지, vol.30, pp.247-252, 2006.
- [9] 이양민, 이광용, 배승현, 장휘, 이재기, “다채널 진동 센서를 이용한 선박 엔진의 진동 감지 및 고장 분류 시스템”, 정보처리학회지, vol.17, pp.81-92, 2010.
- [10] 천행춘, 유영호, “신경회로망을 이용한 디젤기관

- 의 데이터 이상감지 시스템에 관한 연구”, 한국마린엔지니어링 학회지, vol.26, pp. 493-500, 2007.
- [11] 김영일, 오현경, 유영호, “상관분석법에 의한 선박 기관실 고장진단 시스템 개발”, 한국마린엔지니어링 학회지, vol.30, pp.253-259, 2006.
- [12] Yu-Long, ZhuoBin Shi, Theingi Shwe, Xiao-Zhong Wang, “Fault Diagnosis of Marine Main Engine Cylinder Cover Based on Vibration Signal”, Proceeding of Machine Learning and Cybernetics, vol.2, pp.3018-3022, 2007.
- [13] 이상훈, “수치형 속성을 이산화 하기 위한 밀도 기반의 접근 방법”, 석사 학위 논문, pp.83-88, 2003.
- [14] Huan Liu, Farhad Hussain, Chew Lim Tan, Manoranjan Dash, “Discretization: An Enabling Technique”, Data Mining and Knowledge Discovery, vol.6, pp.393-423, 2002.
- [15] Bill Chiu, Eamonn Keogh, Stefano Lonardi, “Probabilistic discovery of time series motifs”, Proceedings of the international conference on Knowledge discovery and Data mining, pp.493-498, 2003.
- [16] Eamonn J. Keogh, Michael J. Pazzani, “A Simple Dimensionality Reduction Technique for Fast Similarity Search in Large Time Series Databases”, Knowledge discovery and Data mining, pp.122-133, 2000.
- [17] 이양지, 김덕영, 황민순, 정영수, “시스템 결함원인 분석을 위한 데이터 로그 전처리 기법 연구”, 한국 CAD/CAM 학회 논문집 vol.17, pp.97-110, 2012.
- [18] Tapio Elomaa. and Juho Rousul, “General and Efficient Multisplitting of Numerical Attributes”, Machine learning, vol. 36, pp. 201-244, 1999.
- [19] 전병환, 김재희, “연속적인 값을 갖는 속성의 이산화를 위한 상대적인 엔트로피 소개”, 대한전자공학회 학술대회 논문집, vol.9, pp. 1210-1213, 1995.
- [20] K.M Ho, P. D Scott, “Zeta: A Global Method for Discretization of Continuous Variables”, Knowledge Discovery and Data Mining, pp.191-194, 1997.
- [21] Jiawei Han, Jian Pei, Yien Yin, “Mining frequent patterns without candidate generation: A Frequent-Pattern Tree Approach”, Data mining and Knowledge discovery, vol.8, pp.53-81, 2004.
- [22] 서재학, 남인길, “연관마이닝에 의한 데이터베이스 캐시 설계”, 한국산업정보학회, vol.7, pp.16-32, 2002.
- [23] 오상현, 장중혁, “네트워크 패킷에 대한 연관 마이닝 기법을 적용한 네트워크 비정상 행위 탐지”, 한국산업정보학회, vol.4, pp.22-29, 2009.



이 양 지 (Yang Ji Lee)

- 한국해양대학교 제어자동화공학과 학사
- UNIST 디자인 및 인간공학부 석사과정

- 관심분야: Intelligent Failure Analysis, Signal Processing



정 영 수 (Young Soo Cheong)

- 연세대학교 기계공학과 학사
- 울산대학교 자동차선박기술대학원 석사
- 현대중공업 연구소

- 현대중공업 e-Business 사업본부, 조선 ERP 추진
- 현대중공업 통합전산실 부장
- 관심분야: Shipbuilding Logistics, Industrial IT Convergence



김 덕 영 (Duck Young Kim)

- 포스텍 산업공학과 학사
- 포스텍 산업공학과 석사
- 스위스 로잔 연방공과대학교(EPFL), 기계공학과 박사

- 고등기술연구원 생산기술센터 주임연구원
- 스위스 로잔 연방공과대학교, Post-doc
- 스위스 취리히 연방공과대학교 (ETHZ), Post-doc
- University of Warwick, Research Fellow
- UNIST 디자인 및 인간공학부 조교수
- 관심분야: Intelligent Failure Analysis, Smart Ship, Digital Manufacturing, Remote Laser Welding

논문접수일: 2012년 04월 13일
 1차수정완료일: 2012년 05월 25일
 2차수정완료일: 2012년 06월 10일
 게재확정일: 2012년 07월 04일



황 민 순 (Min Soon Hwang)

- 한양대학교 전자공학과 학사
- 현대정보기술 (울산 현대중공업 전산실)
- 현대중공업 통합전산실 차장

- 관심분야: e-Navigation, Smart Ship