

선박내부 이상감지의 정확도 향상을 위한 실시간 신뢰구간 추정

김영주*, 허유경*, 정민아*

*목포대학교 컴퓨터공학과

e-mail:xfile7@mokpo.ac.kr, alice2020@hanmail.net, majung@mokpo.ac.kr

Real-time Confidence interval estimation for Improved accuracy of Ship-inside the anomaly detection

Yeong-Ju Kim*, Heo-yu Kyung*, Majung-A Jeong*

*Dept of Computer Engineering, Mok-po University

요 약

본 논문은 선박내부의 센서데이터 이상감지를 위해 실시간 신뢰구간을 설정하고 신뢰구간을 초과하거나 미만이 되면 경보를 통해 관리자에게 알려주는 모니터링을 위한 신뢰구간 추정이다. 여기서, 이상감지 예측의 정확도 향상을 위해 단순지수평활법과 이동평균법의 평균제곱오차를 비교 평가 하였다. 실험결과, 이동평균법의 평균제곱오차가 단순지수평활법 보다 적게 나와 선박 내부 모니터링을 위한 신뢰구간은 이동평균법을 적용하였다.

1. 서론

현재 해상에 운용되는 선박이 점차 대형화 되고 있고, 해상운송을 하는 대형 컨테이너선과 유조선 그리고 여객만을 수송하는 호화객선(크루즈선)처럼 대형화된 선박의 내부를 실시간으로 모니터링 하는 것은 불의의 사고나 위기상황 속에서 신속히 대처하기 위함이다.[1]

본 논문에서는 센서데이터의 이상 감지를 위해 신뢰구간을 설정하고 신뢰구간을 초과하거나 미만이 되면 경보를 통해 관리자에게 알려주는 모니터링을 위해 시계열 데이터 예측 방법인 단순지수평활법과 이동평균법의 평균제곱오차를 비교 평가 하였다. 실험결과, 이동평균법의 평균제곱오차가 단순지수평활법 보다 작게 나와 선박 내부 모니터링을 위한 신뢰구간은 이동평균법을 적용하였다. 시계열 예측값 실험은 SPSS를 사용하였다. 관리자는 실시간마다 변하는 신뢰구간에 따른 모니터링 정보를 활용하여 선박 내부에서 발생한 위급한 상황에서 신속하고 정확하게 의사결정을 할 수 있다.

2. 관련연구

2.1 단순지수평활법

지수평활법은 가장 가까운 과거에 가장 큰 가중치를 부여하는 것으로 모형의 정확성이 높고, 쉽게 이해할 수 있고, 필요한 계산이 많지 않다. 또한 컴퓨터 기억장소를 많이 요구하지 않는다. 이러한 이유로 시계열분석 방법 중에서 단기예측을 실시하는 데에 가장 많이 이용된다.[2]

단순지수평활법을 사용한 예측값은 아래 식(1)과 같다.

$$F_t = (1 - \alpha)F_{t-1} + \alpha D_{t-1} \dots \dots \dots \text{식(1)}$$

F_t : 기간 t의 예측값

F_{t-1} : 기간 (t-1)의 예측값

D_{t-1} : 기간 (t-1)의 실제값

α : 지수평활계수($0 < \alpha < 1$)

2.2 이동평균법

이동평균법은 시계열 데이터의 일정 기간의 관측치를 활용하여 시계열 관측치의 추세 변화를 파악하는 기법으로 일정구간을 시간에 따라 구간 이동을 수행하면서 관측치들의 추세를 생성한다.[3]

이동평균법을 사용한 예측값은 아래 식(2)와 같다.

$$MA = \frac{X_{t-(N-1)} + X_{t-(N-2)} + X_{t-(N-3)} + \dots + X_{t-(i+1)} + X_{t-1}}{N}$$

$$+ \frac{X_{t-(i-1)} + X_{t-(i-2)} + \dots + X_{t-(N-N-1)} + X_{t-(N-N)}}{N} \dots \dots \dots \text{식(2)}$$

2.3 평균제곱오차

평균제곱오차(mean squared errors;MSE)는 예측오차를 제공하는 것으로 정확한 예측을 가능하게 하기 위해서는 오차의 크기를 측정하는 것이 중요하며 오차의 크기는 작을수록 현실과 부합된다.[4]

평균제곱오차는 아래 식(3)과 같다.

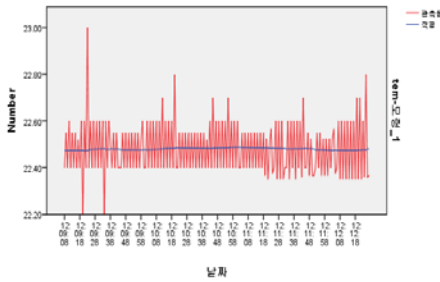
$$MSE = \frac{\sum_{t=1}^n (e_t)^2}{n} = \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n} \dots \dots \dots \text{식(3)}$$

여기서, t는 시간, Y_t 는 실측값, Y_t 의 예측값은 \hat{Y}_t , 예측값 \hat{Y}_t 와 실제값 Y_t 의 차이, 즉 예측오차 e_t 는 아래 식(4)와 같다.

$$e_t = Y_t - \hat{Y}_t \dots \dots \dots \text{식(4)}$$

3. 본문

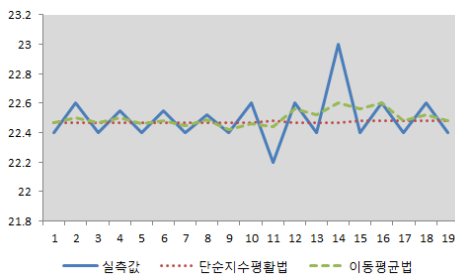
3.1 데이터분석



[그림1] 센서데이터(온도)

[그림1]의 센서데이터는 5분 동안 총3개의 센서로 데이터를 수집하였다. 센서데이터는 연속형 시계열 데이터로 각각 1초에 한개씩의 온도 정보를 ServerPC로 보냈다. 수집된 원시데이터는 1초 단위로 전처리하고 시계열 분석을 위한 데이터 정제 작업을 하여 총 200개의 온도 데이터로 실험을 하였다.

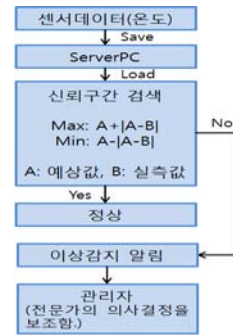
3.3 성능평가



[그림2] 실측값과 예측값 비교

[그림2]에서와 같이 단순지수평활법은 예측값 변화가 없이 일정하고 이동평균법(N=5)은 실측값에 따라 약간의 변동이 나타난다. 단순지수평활법은 시계열데이터의 단기예측을 실시하는데 가장 많이 사용하는 등의 실시간 모니터링의 장점이 있으나 [그림2]에서와 같이 실시간데이터의 변화를 반영하지는 못한다. 그에 반해 이동평균법은 실측값에 따른 추세를 반영한 결과가 나타났다. 본 논문에서 제안한 모니터링은 실시간 모니터링으로 최근 센서데이터의 값을 반영하여 신뢰구간을 설정하기 위해서는 이동평균법을 사용해야 한다. 더불어, 평균제곱오차가 단순지수평활법은 0.0132이고 이동평균법은 0.0095로 이동평균법이 오차 또한 적게 나타났다.

3.2 알고리즘



[그림3] 이상감지 알고리즘

위 [그림3]의 신뢰구간은 이동평균법으로 산출된 예측값 A와 절대오차(|A-B|)의 합이 최대값, 차가 최소값으로 정해지고 최대값과 최소값 사이의 신뢰구간을 정하고 실시간으로 입력된 센서데이터가 신뢰구간을 초과하거나 미만일 경우 이상감지를 관리자에게 알려주어 관리자의 의사결정을 보조하는 역할을 한다.

4. 결론

본 논문에서는 시계열데이터 예측 방법을 비교 평가하여 센서데이터의 이상감지 예측의 정확도 향상을 위한 신뢰구간을 추정 하였다. 이는, 연속형 시계열 데이터는 실시간마다 변하는 신뢰구간을 가진다는 특성을 반영하기 위한 실험으로 시계열데이터의 추세를 반영하는 이동평균법이 효과적임을 확인하였다. 추후, 같은 시간 간격으로 이산형 시계열 데이터를 수집하여 실험, 분석하여 연속형 시계열 데이터의 예측을 보정해야 한다. 또한, 다양한 센서 데이터의 모니터링 적용과 이상치와 결측치 그리고 오작동 등을 시계열데이터 분석 결과를 근거로 구분 예측할

수 있는 연구가 필요하다.

ACKNOWLEDGMENT

이 논문은 2013년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(NRF-2009-0093828) 또한 본 연구는 미래창조과학부 및 정보통신산업진흥원의 IT융합 고급인력과정 지원사업의 연구결과로 수행되었음(NIPA-2014-H0401-14-1009)

참고문헌

- [1] 박진관, 김영주, 박선, 이연우, 정민아, 이성로 “가중치 평균을 적용한 선박내부 이상 감지” 2013년도 한국정보처리학회 춘계학술발표대회 논문집 제20권 1호, 2013.5, 492-494.
- [2] 노형진, “SPSS/Excel에 의한 시계열분석”, 효산, 2008.11.
- [3] 김형일, “임베디드 센서를 위한 시계열 예측 기반 실시간 오류 검출 기법” 2011년도 한국컴퓨터정보학회논문지 제16권 12호, 2011.12, 11-21.
- [4] 김연형, “시계열예측”, 형설,2002.8.