



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0013017
(43) 공개일자 2019년02월11일

(51) 국제특허분류(Int. Cl.)
G06Q 50/10 (2012.01) G06F 17/00 (2019.01)
G06Q 10/04 (2012.01)
(52) CPC특허분류
G06Q 50/10 (2015.01)
G06F 17/00 (2019.01)
(21) 출원번호 10-2017-0096912
(22) 출원일자 2017년07월31일
심사청구일자 2017년07월31일

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
백준걸
서울특별시 강동구 아리수로50길 50, 305동 2803호
유영지
경기도 군포시 수리산로 33, 842동 1702호 (산본동, 계룡아파트)
(74) 대리인
김등용, 김홍석

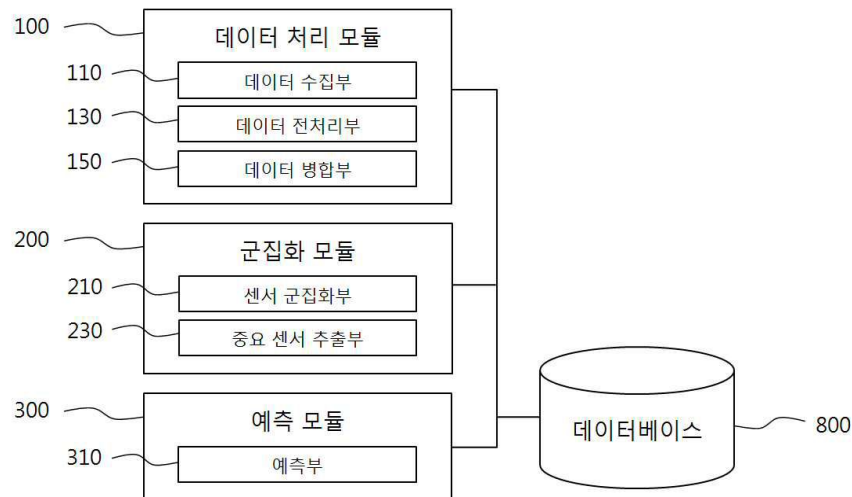
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 센서 군집화 기반의 설비 건강 모니터링 방법 및 장치

(57) 요약

데이터 처리 모듈, 군집화 모듈 및 예측 모듈을 포함하는 설비 건강 모니터링 장치에서 수행되는 센서 군집화 기반 설비 건강 모니터링 방법이 개시된다. 상기 설비 건강 모니터링 방법은 설비에 부착된 센서로부터 실시간으로 수집되는 FDC 데이터를 수집하는 단계, 설비 수명 지표 생성 단계, 선택 단계, 표준화 단계, 특질 추출 단계, 이상치 센서 제거 단계, 센서 대표 패턴을 정의하는 단계, 유사한 특성을 가진 센서들을 군집화하는 단계, 중요 센서 추출 단계, 중요 센서들의 특질들을 이용하여 설비 건강 모니터링 모델을 학습하는 단계 및 새로운 센서 데이터가 수집되면 학습된 모델을 기반으로 설비의 건강을 진단하고 예측하는 단계를 포함한다.

대 표 도 - 도1



(52) CPC특허분류

G06Q 10/04 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 2016R1A2B4013678

부처명 미래창조과학부

연구관리전문기관 한국연구재단

연구사업명 중견연구

연구과제명 빅 데이터 기반 스마트팩토리 구현을 위한 프리딕티브 에널리틱스 기법 개발

기 여 율 1/1

주관기관 고려대학교

연구기간 2016.06.01 ~ 2019.05.31

명세서

청구범위

청구항 1

데이터 처리 모듈, 군집화 모듈 및 예측 모듈을 포함하는 설비 건강 모니터링 장치에서 수행되는 센서 군집화 기반 설비 건강 모니터링 방법에 있어서,

상기 데이터 처리 모듈이 설비에 부착된 복수 개의 센서로부터 실시간으로 수집되는 FDC(Fault Detection & Classification) 데이터를 수집하는 단계;

상기 데이터 처리 모듈이 설비 수명 지표를 생성하는 설비 수명 지표 생성 단계;

상기 데이터 처리 모듈이 각 센서에서 수집된 FDC 데이터 변수 중에서 유의하지 않은 변수를 제거하는 변수 선택 단계;

상기 데이터 처리 모듈이 모든 센서 데이터가 동일한 평균, 표준편차를 가지도록 각 센서 데이터의 표준화를 수행하는 표준화 단계;

상기 데이터 처리 모듈이 각 센서 데이터의 평균과 표준 편차를 계산하여 특징을 추출하는 특징 추출 단계;

상기 데이터 처리 모듈이 이상치 센서를 제거하는 이상치 센서 제거 단계;

상기 군집화 모듈이 센서 대표 패턴을 정의하는 단계;

상기 군집화 모듈이 추출된 특징을 기반으로 군집화 알고리즘을 적용하여 유사한 특성을 가진 센서들을 군집화 하는 군집화 단계;

상기 군집화 모듈이 각각의 군집의 특성을 분석하여 설비의 상태와 상관관계가 높은 군집에 포함된 센서들을 중요 센서로 추출하는 단계;

상기 예측 모듈이 중요 센서들의 특징들을 이용하여 설비 건강 모니터링 모델을 학습하는 단계; 및

상기 예측 모듈이 새로운 센서 데이터가 수집되면 학습된 모델을 기반으로 설비의 건강을 진단하고 예측하는 단계를 포함하는 센서 군집화 기반의 설비 건강 모니터링 방법.

청구항 2

제1항에 있어서,

상기 군집화 단계에서 상기 군집화 알고리즘은 Hierarchical algorithm, K-means algorithm, K-medhoids algorithm, DBSCAN(Density Based Spatial Clustering of Applications with Noise), 및 OPTICS(Ordering Points To Identify the Clustering Structure)를 포함하는,

센서 군집화 기반의 설비 건강 모니터링 방법.

청구항 3

제1항에 있어서,

상기 센서 대표 패턴을 정의하는 단계에서 상기 센서 대표 패턴은 시간에 따라 증가하는 제1 패턴, 시간에 따라 감소하는 제2 패턴, 시간에 독립적이며 바뀌지 않는 패턴인 제3 패턴, 및 시간에 독립적인 기타 패턴인 제4 패턴을 포함하는,

센서 군집화 기반의 설비 건강 모니터링 방법.

청구항 4

제1항에 있어서,

상기 변수 선택 단계에서,

다중 공선성(multicollinearity)을 제거하기 위해 다음의 식을 이용하여 계산한 VIF(variance inflation factor)를 사용하여 변수를 선택하는,

센서 군집화 기반의 설비 건강 모니터링 방법.

$$VIF_i = (1 - R_i^2)^{-1}$$

여기에서, R_i^2 는 i 번째 변수를 종속 변수로 하고 남은 i-1 번째 변수를 독립 변수로 사용하여 회귀(regression)를 결정하는 계수

청구항 5

설비에 부착된 복수 개의 센서로부터 실시간으로 수집되는 FDC(Fault Detection & Classification) 데이터를 수집하여, 설비 수명 지표를 생성하고, 각 센서에서 수집된 FDC 데이터 변수 중에서 유의하지 않은 변수를 제거하며, 모든 센서 데이터가 동일한 평균, 표준편차를 가지도록 각 센서 데이터의 표준화를 수행하고, 각 센서 데이터의 평균과 표준 편차를 계산하여 특질을 추출하며, 이상치 센서를 제거하는 데이터 처리 모듈;

센서 대표 패턴을 정의하여, 추출된 특질을 기반으로 군집화 알고리즘을 적용하여 유사한 특성을 가진 센서들을 군집화하고, 각각의 군집의 특성을 분석하여 설비의 상태와 상관관계가 높은 군집에 포함된 센서들을 중요 센서로 추출하는 군집화 모듈; 및

중요 센서들의 특질들을 이용하여 설비 건강 모니터링 모델을 학습하고, 새로운 센서 데이터가 수집되면 학습된 모델을 기반으로 설비의 건강을 진단하고 예측하는 예측 모듈을 포함하는 센서 군집화 기반의 설비 건강 모니터링 장치.

청구항 6

제5항에 있어서,

상기 군집화 알고리즘은 Hierarchical algorithm, K-means algorithm, K-medhoids algorithm, DBSCAN(Density Based Spatial Clustering of Applications with Noise), 및 OPTICS(Ordering Points To Identify the Clustering Structure)를 포함하는 센서 군집화 기반의 설비 건강 모니터링 장치.

청구항 7

제5항에 있어서,

상기 센서 대표 패턴은 시간에 따라 증가하는 제1 패턴, 시간에 따라 감소하는 제2 패턴, 시간에 독립적이며 바뀌지 않는 패턴인 제3 패턴, 및 시간에 독립적인 기타 패턴인 제4 패턴을 포함하는,

센서 군집화 기반의 설비 건강 모니터링 장치.

청구항 8

제5항에 있어서,

다중 공선성(multicollinearity)을 제거하기 위해 다음의 식을 이용하여 계산한 VIF(variance inflation

actor)를 사용하여 각 센서에서 수집된 FDC 데이터 변수 중에서 유의하지 않은 변수를 제거하는, 센서 군집화 기반의 설비 건강 모니터링 장치.

$$VIF_i = (1 - R_i^2)^{-1}$$

여기에서, R_i^2 는 i 번째 변수를 종속 변수로 하고 남은 $i-1$ 번째 변수를 독립 변수로 사용하여 회귀 (regression)를 결정하는 계수

발명의 설명

기술 분야

[0001] 본 발명은 반도체 공정과 같은 첨단 공정에서 사용되는 설비의 유지보수 주기 및 시점을 결정하기 위한 설비 건강 모니터링 방법 및 장치에 관한 것으로, 더욱 상세하게는, 설비에 부착된 센서로부터 실시간으로 수집되는 데이터 기반의 설비 건강 지표를 제시하고 상태 기반의 유지보수가 가능하도록 하는 센서 군집화 기반의 설비 건강 모니터링 방법 및 장치에 관한 것이다.

배경 기술

[0003] 본 발명은 첨단 공정 설비에서 수집되는 센서 데이터의 군집화 및 센서 기반의 설비 건강 모니터링 기법에 관한 것으로, 설비의 건강과 상관성이 높은 중요 센서를 추출하고 설비의 건강도를 정확하게 모니터링 및 예측하는 기법에 관한 것이다.

[0004] 과거에는 엔지니어들의 경험적 지식을 기준으로 일정 주기의 예방 유지보수를 수행했기 때문에 설비의 상태와 상관없이 유지보수를 수행하는 경우가 많았다. 그러나 하나의 설비에 백여 개의 센서가 장착되면서 실시간으로 온도나 압력 등을 측정하여 설비의 상태를 반영하여 유지보수를 수행하고자 한다.

[0005] 센서 데이터를 기반으로 하는 베어링이나 모터 등과 같은 단순 부품의 수명 예측 기법은 다양한 방법론을 통해 알고리즘이 개발되고 있지만, 첨단 공정에서 사용되는 설비의 수명은 수많은 변수의 영향을 동시에 받기 때문에 기존의 알고리즘을 그대로 적용하기 어렵다.

[0006] 백여 개의 센서가 부착된 설비로부터 실시간으로 생성되는 방대한 양의 데이터를 핸들링하여 건강 모니터링 알고리즘을 학습 시에는 복잡한 데이터 구조로 인하여 핸들링 비용이 증가하고 노이즈 센서가 포함되어 설비 수명 예측 정확도가 떨어질 수 있다.

[0007] 따라서 설비의 노후화와 같은 상태 변화와 상관성이 높은 중요 센서들을 추출하고 해당 센서들의 교호작용을 고려하여 설비의 상태를 모니터링하고 수명을 예측하는 기법이 필요하다.

선행기술문헌

특허문헌

[0009] (특허문헌 0001) JP 4978789 B2
(특허문헌 0002) KR 1686338 B1
(특허문헌 0003) JP 5301717 B2
(특허문헌 0004) KR 2014-0011064 A

발명의 내용

해결하려는 과제

[0010] 본 발명이 이루고자 하는 기술적인 과제는 설비의 상태 변화와 상관성이 높은 센서를 추출하고 각 센서들의 교호작용을 고려한 센서 군집화 기반의 설비 건강 모니터링 방법 및 장치를 제공하는 것이다.

과제의 해결 수단

[0012] 본 발명의 일 실시 예에 따른 센서 군집화 기반의 설비 건강 모니터링 방법은 설비 건강 모니터링 장치에서 수행되고, 데이터 처리 모듈이 설비에 부착된 복수 개의 센서로부터 실시간으로 수집되는 FDC(Fault Detection & Classification) 데이터를 수집하는 단계, 데이터 처리 모듈이 설비 수명 지표를 생성하는 설비 수명 지표 생성 단계, 데이터 처리 모듈이 각 센서에서 수집된 FDC 데이터 변수 중에서 유의하지 않은 변수를 제거하는 변수 선택 단계, 데이터 처리 모듈이 모든 센서 데이터가 동일한 평균, 표준편차를 가지도록 각 센서 데이터의 표준화를 수행하는 표준화 단계, 데이터 처리 모듈이 각 센서 데이터의 평균과 표준 편차를 계산하여 특질을 추출하는 특질 추출 단계, 데이터 처리 모듈이 이상치 센서를 제거하는 이상치 센서 제거 단계, 군집화 모듈이 센서 대표 패턴을 정의하는 단계, 군집화 모듈이 추출된 특질을 기반으로 군집화 알고리즘을 적용하여 유사한 특성을 가진 센서들을 군집화하는 군집화 단계, 군집화 모듈이 각각의 군집의 특성을 분석하여 설비의 상태와 상관관계가 높은 군집에 포함된 센서들을 중요 센서로 추출하는 단계, 예측 모듈이 중요 센서들의 특질들을 이용하여 설비 건강 모니터링 모델을 학습하는 단계, 및 예측 모듈이 새로운 센서 데이터가 수집되면 학습된 모델을 기반으로 설비의 건강을 진단하고 예측하는 단계를 포함한다.

[0013] 또한, 본 발명의 일 실시 예에 따른 센서 군집화 기반의 설비 건강 모니터링 장치는 설비에 부착된 복수 개의 센서로부터 실시간으로 수집되는 FDC(Fault Detection & Classification) 데이터를 수집하여, 설비 수명 지표를 생성하고, 각 센서에서 수집된 FDC 데이터 변수 중에서 유의하지 않은 변수를 제거하며, 모든 센서 데이터가 동일한 평균, 표준편차를 가지도록 각 센서 데이터의 표준화를 수행하고, 각 센서 데이터의 평균과 표준 편차를 계산하여 특질을 추출하며, 이상치 센서를 제거하는 데이터 처리 모듈, 센서 대표 패턴을 정의하여, 추출된 특질을 기반으로 군집화 알고리즘을 적용하여 유사한 특성을 가진 센서들을 군집화하고, 각각의 군집의 특성을 분석하여 설비의 상태와 상관관계가 높은 군집에 포함된 센서들을 중요 센서로 추출하는 군집화 모듈 및 중요 센서들의 특질들을 이용하여 설비 건강 모니터링 모델을 학습하고, 새로운 센서 데이터가 수집되면 학습된 모델을 기반으로 설비의 건강을 진단하고 예측하는 예측 모듈을 포함한다.

발명의 효과

[0015] 본 발명의 일 실시 예에 따른 센서 군집화 기반의 설비 건강 모니터링 방법 및 장치의 센서 데이터 전처리 기법은 이상치(outlier) 센서를 제거하고, 시간의 흐름에 따른 센서 데이터의 패턴을 추출함으로써 데이터의 크기를 줄이고 노이즈를 제거할 수 있는 효과가 있다.

[0016] 또한, 센서 군집화를 통해 설비의 상태 변화와 상관성이 높은 센서 군집, 시간의 흐름과 상관없이 일정한 범위 내에서 안정적인 값이 수집되는 센서 군집, 무작위 패턴이 수집되어 상관성이 낮은 센서 군집 등으로 정확하게 분류할 수 있는 효과가 있다.

[0017] 또한, 설비 건강 모니터링 학습 시 설비의 상태 변화와 상관성이 높은 센서들을 추출하여 사용함으로써 설비의 수명 진단 및 예측의 성능을 향상시키는 효과가 있다.

도면의 간단한 설명

[0019] 본 발명의 상세한 설명에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 상세한 설명이 제공된다.

도 1은 본 발명의 일 실시 예에 따른 센서 군집화 기반의 설비 건강 모니터링 장치의 기능 블록도이다.

도 2는 도 1에 도시된 설비 건강 모니터링 장치를 이용한 센서 군집화 기반의 설비 건강 모니터링 방법을 설명하기 위한 흐름도이다.

도 3은 FDC 요약 변수의 VIF를 계산한 표의 일 예를 도시한다.

도 4의 (a)는 원데이터 그래프의 일 예를 도시하고, 도 4의 (b)는 도 4의 (a)를 표준화 및 데이터 병합한 결과

를 도시하는 그래프이다.

도 5는 이상치 센서 데이터의 일 예를 도시한다.

도 6은 센서 데이터의 다양한 패턴을 도시한다.

발명을 실시하기 위한 구체적인 내용

- [0020] 본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시 예들에 대해서 특정한 구조적 또는 기능적 설명은 단지 본 발명의 개념에 따른 실시 예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시 예들은 다양한 형태로 실시될 수 있으며 본 명세서에 설명된 실시 예들에 한정되지 않는다.
- [0021] 본 발명의 개념에 따른 실시 예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시 예들을 도면에 예시하고 본 명세서에서 상세하게 설명하고자 한다. 그러나 이는 본 발명의 개념에 따른 실시 예들을 특정한 개시 형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물, 또는 대체물을 포함한다.
- [0022] 본 명세서에서 사용한 용어는 단지 특정한 실시 예를 설명하기 위해 사용된 것으로서, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 본 명세서에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0023] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0025] 이하, 본 명세서에 첨부된 도면들을 참조하여 본 발명의 실시 예들을 상세히 설명한다.
- [0026] 먼저, 도 1을 참조하여, 본 발명의 일 실시 예에 따른 센서 군집화 기반의 설비 건강 모니터링 장치에 대해 상술한다.
- [0027] 도 1은 본 발명의 일 실시 예에 따른 센서 군집화 기반의 설비 건강 모니터링 장치의 기능 블록도이다. 도 1을 참조하면, 설비 건강 모니터링 장치는 데이터 처리 모듈(100), 군집화 모듈(200), 예측 모듈(300), 및 데이터베이스(800)를 포함한다.
- [0028] 본 명세서에서 사용되는 '-부' 또는 '모듈'이라 함은 본 발명의 기술적 사상을 수행하기 위한 하드웨어 및 상기 하드웨어를 구동하기 위한 소프트웨어의 기능적, 구조적 결합을 의미할 수 있다. 예컨대, 상기 '-부' 또는 '-모듈'은 소정의 코드와 상기 소정의 코드가 수행되기 위한 하드웨어 리소스의 논리적인 단위를 의미할 수 있으며, 반드시 물리적으로 연결된 코드를 의미하거나 한 종류의 하드웨어를 의미하는 것은 아니다.
- [0029] 상기 설비 건강 모니터링 장치의 데이터 처리 모듈(100)은 데이터 수집부(110), 데이터 전처리부(130) 및 데이터 병합부(150)를 포함하며, 데이터베이스(800)로부터 데이터를 수집하고 전처리 및 병합하여 특질을 추출할 수 있다.
- [0030] 상기 설비 건강 모니터링 장치의 군집화 모듈(200)은 센서 군집화부(210) 및 중요 센서 추출부(230)를 포함하며, 추출된 특질들을 바탕으로 유사 센서들끼리 군집화하고 군집들의 특성을 분석하여 중요 센서 군집을 추출할 수 있다.
- [0031] 상기 설비 건강 모니터링 장치의 예측 모듈(300)은 예측부(310)를 포함하며, 중요 센서 군집을 이용하여 설비 건강 모니터링을 적용하고 설비의 건강을 진단 및 예측할 수 있다.
- [0032] 상기 설비 건강 모니터링 장치의 데이터베이스(800)는 수집된 데이터, 전처리 및 병합 데이터, 군집화 결과, 중요 센서, 또는 모니터링 결과 등을 저장할 수 있다. 본 명세서에서 데이터베이스라 함은, 각각의 데이터베이스

에 대응되는 정보를 저장하는 소프트웨어 및 하드웨어의 기능적 구조적 결합을 의미할 수도 있다.

- [0033] 상기 설비 건강 모니터링 장치에 포함될 수 있는 제어모듈(미도시)은 상기 설비 건강 모니터링 장치의 전반적인 동작을 제어한다. 즉, 데이터 처리 모듈(100), 군집화 모듈(200), 예측 모듈(300) 및 데이터베이스(800)의 동작을 제어할 수 있다.
- [0034] 본 실시 예와는 다르게, 상기 설비 건강 모니터링 장치가 아닌 별도의 장치에 데이터 처리 모듈(100), 군집화 모듈(200), 예측 모듈(300) 및 데이터베이스(800)) 중 어느 하나 이상을 구비하는 것도 가능하다. 이때, 각각의 장치별로 제어모듈을 각각 구비할 수 있다.
- [0036] 이하, 도 2 내지 도 6을 참조하여, 본 발명의 일 실시 예에 따른 센서 군집화 기반의 설비 건강 모니터링 장치에서 수행되는 설비 건강 모니터링 방법에 대하여 자세히 살펴보도록 한다.
- [0037] 도 2는 도 1에 도시한 센서 군집화 기반의 설비 건강 모니터링 장치를 이용한 센서 군집화 기반의 설비 건강 모니터링 방법을 설명하기 위한 흐름도이다.
- [0038] 도 2를 참조하면, 센서 군집화 기반의 설비 건강 모니터링 방법은 데이터 수집 단계(S100), 데이터 처리 단계(S200), 센서 군집화 단계(S300), 및 설비 건강 모니터링 단계(S400)를 포함한다.
- [0039] 먼저, 센서 군집화 기반의 설비 건강 모니터링 장치의 데이터 처리 모듈(100)은 설비에 부착된 복수 개의 센서로부터 실시간으로 수집되는 FDC(Fault Detection & Classification) 데이터를 수집한다.
- [0040] 예를 들어, 80 개의 센서가 부착된 CVD(Chemical Vapor Deposition) 장비의 모든 센서에 대해 데이터 샘플링 속도가 1 Hz 인 경우, 하나의 장비로부터 매일 6,912,000개의 원데이터가 수집된다. 데이터 관리 및 활용의 효율성을 위해 수집된 원데이터는 평균, 표준 편차, 최소값 및 최대값 등과 같은 FDC 요약 변수(FDC summary variables)로 변환될 수 있다. 각 센서의 FDC 요약 데이터(FDC summary data)는 설비 건강 모니터링 모델에 대한 입력 변수로 사용될 수 있다.
- [0041] 다음, 센서 군집화 기반의 설비 건강 모니터링 장치의 데이터 처리 모듈(100)은 설비 수명 지표를 생성할 수 있다(S210). 설비를 교체하거나 유지보수 한 시점으로부터 설비 사용 시간을 측정하여 설비의 수명을 나타내는 지표로 생성한다. 이때, 설비를 교체하거나 유지보수 한 시점으로부터의 설비 사용 시간(설비 수명 지표)은 설비 건강 모니터링 모델의 종속 변수(Y변수)로 사용한다.
- [0042] 다음, 센서 군집화 기반의 설비 건강 모니터링 장치의 데이터 처리 모듈(100)은 변수 선택 및 표준화 단계(S220)를 수행한다. 구체적으로, 각 센서에서 수집된 FDC 데이터 변수(최소, 최대, 평균, 분산) 중에서 유의하지 않은 변수를 제거하고, 각각의 센서 데이터들을 표준화시켜 모든 센서 데이터가 동일한 평균, 표준편차를 가지도록 한다.
- [0043] 변수 선택을 위하여 분산 인플레이션 액터(variance inflation actor, VIF)를 사용할 수 있다. VIF는 공선성(collinearity)으로 인해 추정 회귀 계수의 분산(estimated regression coefficient)이 얼마나 증가하는지 측정하며, VIF는 아래의 수학식 1을 이용하여 계산한다.

수학식 1

[0044]
$$VIF_i = (1 - R_i^2)^{-1}$$

- [0045] 수학식 1에서, R_i^2 는 i 번째 변수를 종속 변수로 하고 남은 i-1 변수를 독립 변수로 사용하여 회귀(regression)를 결정하는 계수이다. VIF_i 가 10보다 크다면 i 번째 독립 변수가 다른 변수와 높은 상관 관계가 있음을 의미한다. 도 3은 FDC 요약 변수의 VIF를 계산한 표의 일 예를 도시한다. 도 3을 참고하면, 평균, 최소 및 최대 변수의 VIF가 10을 초과하므로 세 변수 사이에 강한 다중 공선성(multicollinearity)이 있음을 확인할 수 있으며, 다중 공선성을 제거하기 위해, 평균 및 표준편차 변수를 선택할 수 있다.

- [0046] 각 센서의 데이터는 아래의 수학식 2에 의해 z-스코어(z-score)로 정규화한다.

수학식 2

$$z_{ijk} = \frac{(x_{ijk} - \bar{x}_{ij})}{s_{ij}}$$

[0047]

[0048] 수학식 2에서, i 는 센서의 수, j 는 날짜, k 는 j 번째 날짜의 i 번째 센서에서 수집 한 데이터의 양, $\bar{x}_{ij}$ 및 s_{ij} 는 각각 데이터의 평균 및 표준 편차이다. z -스코어 스케일링 후 표준화된 데이터의 평균은 0이고 표준 편차는 1이다.

[0049] 다음, 센서 군집화 기반의 설비 건강 모니터링 장치의 데이터 처리 모듈(100)은 특질 추출 단계(S230)를 수행한다. 특질 추출 단계에서 각 센서별로 하루 동안 수집된 데이터들의 평균과 표준 편차를 계산하여 특질을 추출함으로써 데이터의 양을 줄이고 데이터가 수집된 패턴을 명확하게 표현할 수 있다.

[0050] 구체적으로, 아래의 수학식 3 내지 수학식 5를 이용하여 평균($Mean_i$) 및 표준편차(SD_i)를 계산한다.

수학식 3

$$\bar{z}_{ij} = \frac{1}{n} \sum_{k=1}^n z_{ijk}$$

[0051]

수학식 4

$$Mean_i = [\bar{z}_{i1} \bar{z}_{i2} \bar{z}_{i3} \dots \bar{z}_{ij}]$$

[0052]

수학식 5

$$SD_i = \sqrt{\frac{1}{m \cdot n} \sum_{j=1}^m \sum_{k=1}^n (z_{ijk} - \bar{z}_{ij})^2}$$

[0053]

[0054] 여기서 $\bar{z}_{ij}$ 는 j 번째 날짜에 i 번째 센서에서 수집된 표준화 된 데이터의 평균이고, m 은 설비를 교체하거나 유지보수 한 시점(또는 PM) 직전의 장비 사용 날짜이며, n 은 i 번째 센서에서 j 번째 날짜에 수집된 데이터의 총량을 의미한다.

[0055] 도 4는 원데이터 그래프 및 표준화 및 데이터 병합 결과 그래프의 일 예를 도시한다. 도 4의 (a)는 두 개의 센서에서 수집한 원 데이터(raw data)를 도시한다. 도 4의 (a)를 참고하면, 두 개의 센서로부터 수집된 두 데이터는 서로 다른 스케일(scales)과 값(values)을 가진다. 도 4의 (b)는 동일한 수의 데이터 및 균일한 스케일을 갖는 표준화 및 데이터 병합(aggregation)의 결과를 도시한다.

[0056] 다음, 센서 군집화 기반의 설비 건강 모니터링 장치의 데이터 처리 모듈(100)은 이상치 센서 제거 단계(S240)를 수행한다. 각 센서 별로 Y변수(설비의 수명 지표)와 X변수(단계 S230에서 추출된 특질) 간의 상관 계수를 계산하여 상관 계수가 0인 센서들을 이상치 센서(Outlier sensors)라고 정의하고 이상치 센서를 제거한다.

[0057] 즉, 각 센서의 노화, 평균 및 SD 변수의 회귀 모델(regression model)의 R^2 값이 0일 때, 해당 센서를 이상치

센서(또는 noise sensor)라고 정의하고, 이를 제거한다.

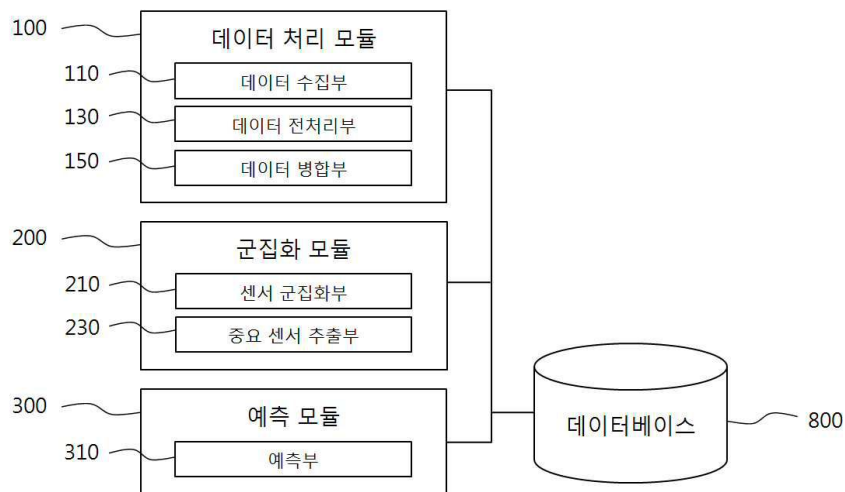
- [0058] 도 5는 이상치 센서 데이터의 일 예를 도시한다. 도 5를 참고하면, 도 5의 (a)의 센서 X26은 주어진 기간 동안 충분한 자료를 수집하지 못하였으며, 도 5의 (b)의 센서 X52의 경우 수집된 데이터에 일반적인 패턴이 없으며 다른 센서에 비해 상대적으로 크기가 큰 것을 확인할 수 있다. 이상치 센서 제거 단계(S240)에서 R^2 의 값이 0인 센서 X26 및 센서 X52는 제거될 수 있다.
- [0059] 데이터 전처리가 끝나면, 센서 군집화 기반의 설비 건강 모니터링 장치(10)의 군집화 모듈(200)은 추출된 특징을 기반으로 유사한 특성을 가진 센서의 군집화를 수행한다.
- [0060] 먼저, 센서 군집화 기반의 설비 건강 모니터링 장치(10)의 군집화 모듈(200)은 센서 대표 패턴을 정의한다(S310). 구체적으로, 센서 대표 패턴은 시간에 따라 증가하는 증가 패턴(increased patterns, 이하 '제1 패턴'이라 함), 시간에 따라 감소하는 감소 패턴(decreased patterns, 이하 '제2 패턴'이라 함), 시간에 독립적인 바뀌지 않는 패턴(unchanged patterns, 이하 '제3 패턴'이라 함) 및 시간에 독립적인 기타 패턴(other pattern, 이하 '제4 패턴'이라 함)을 포함한다.
- [0061] 도 6은 센서 데이터의 다양한 패턴을 도시한다. 패턴은 시간 경과에 따른 장비 상태의 변화를 나타내며, 도 6의 (a)는 시간에 따라 증가하는 패턴, 도 6의 (b)는 시간에 따라 감소하는 패턴을 도시한다. 도 6의 (c)는 시간에 독립적인 변하지 않는 패턴을 도시하며, 도 6의 (d)는 시간에 독립적인 기타 패턴을 도시한다. 도 6의 (a) 및 (b)의 패턴은 도 6의 (c) 및 (d)의 패턴에 비해 시간 경과에 따른 장비 상태 변화를 명확하게 반영한다.
- [0062] 다음, 센서 군집화 기반의 설비 건강 모니터링 장치의 군집화 모듈(200)은 군집화 알고리즘을 적용하여 센서들을 군집화할 수 있다(S320). 구체적으로, Hierarchical algorithm, K-means algorithm, K-medhoids algorithm, DBSCAN(Density Based Spatial Clustering of Applications with Noise), OPTICS(Ordering Points To Identify the Clustering Structure)의 총 5가지의 군집화 알고리즘을 적용하여 센서들을 군집화할 수 있다.
- [0063] 군집화는 주어진 데이터에 대해 클래스 정보가 없을 때 유사도에 근거하여 거리상 가까운 데이터끼리 여러 개의 부분 집합을 형성하는 방법으로, Hierarchical clustering은 거리 측정법을 기반으로 데이터 간의 유사도를 계산하여 군집화하는 방식이며, density-based clustering은 데이터 간의 밀도를 고려하여 높은 밀도로 밀집된 데이터들끼리 군집화하는 방식이다.
- [0064] 이때, 단계 S310에서 정의한 센서 대표 패턴의 개수에 따라 클러스터의 수를 설정한다. 예를 들어, 증가 된 패턴, 감소 된 패턴, 변하지 않은 패턴 및 다른 패턴을 분류하기 위해 클러스터의 수를 4로 설정할 수 있다. 계층적 클러스터링, K-means 알고리즘, K-medhoid 알고리즘의 경우, 클러스터 수를 4로 설정하고, DBSCAN의 경우, 엡실론 부근(epsilon neighborhood)의 크기와 엡실론 영역(epsilon region)의 최소 지점(minimum points)의 수는 모두 3으로 설정하며, OPTICS의 경우, 엡실론 부근의 크기의 상한은 60, 엡실론 영역의 최소 점(minimum points)의 수는 2, 클러스터를 식별하기 위한 임계 엡실론(threshold epsilon)은 3.2로 설정한다. DBSCAN 및 OPTICS 군집화 알고리즘을 통해 제3 패턴과 제4 패턴을 분류할 수 있다. 이때, 클러스터링 성능을 극대화하기 위해 그리드 검색을 통해 매개 변수를 예측하고 설정할 수 있다.
- [0065] 다음, 센서 군집화 기반의 설비 건강 모니터링 장치의 군집화 모듈(200)은 각각의 군집들의 특성을 분석하여 설비의 상태와 상관관계가 높은 군집을 찾아내고 해당 군집에 포함된 센서들을 중요 센서로 정의할 수 있다(S330).
- [0066] 다음, 설비 건강 모니터링 알고리즘 학습 단계(S410)에서 중요 센서들의 특징들을 이용하여 다중회귀모형을 학습한다. 이때, 센서들의 특징과 각 센서 종류의 교호작용을 반영한다.
- [0067] 새로운 데이터가 수집되면 학습된 모델을 기반으로 설비의 건강을 진단하고 예측한다(S420).
- [0068] 본 발명에 따른 센서 군집화 기반의 설비 건강 모니터링 방법을 통하여 설비 건강 모니터링 학습 시 설비의 상태 변화와 상관성이 높은 센서들을 추출하여 사용함으로써 보다 효율적으로 설비의 수명을 진단하고 예측할 수 있다.
- [0069] 본 발명은 도면에 도시된 실시 예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시 예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 등록청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

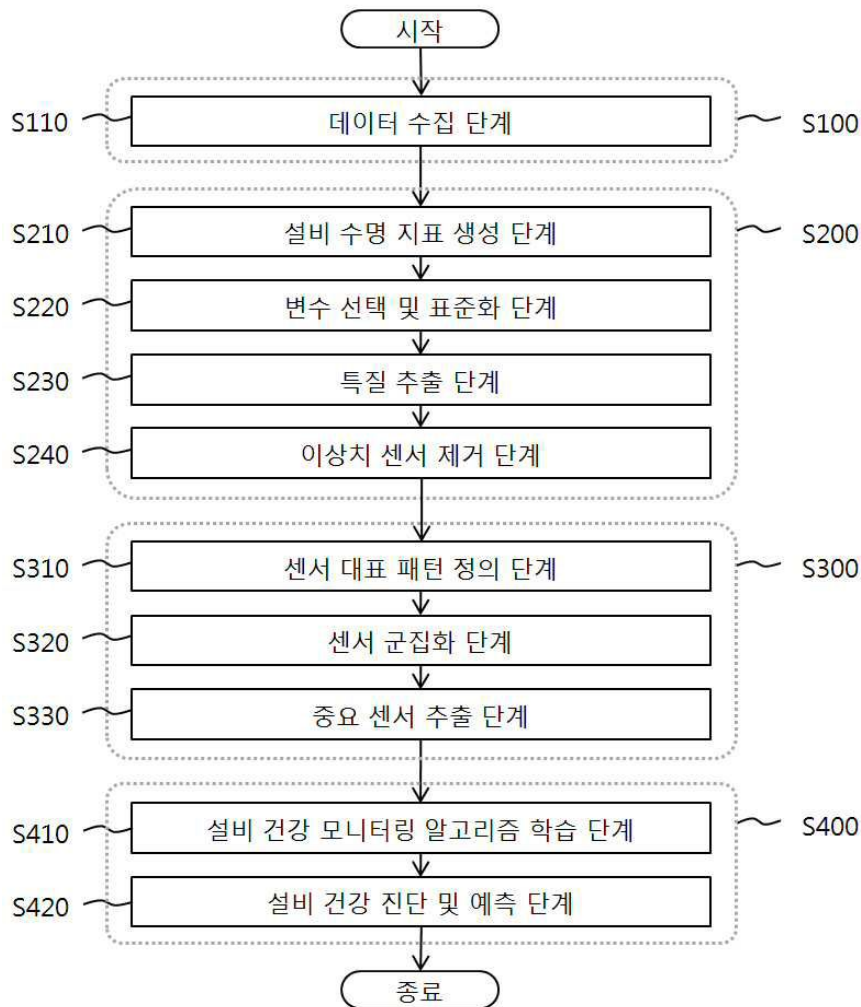
- [0071]
- 100 : 데이터 처리 모듈
 - 200 : 군집화 모듈
 - 300 : 예측 모듈
 - 800 : 데이터베이스

도면

도면1



도면2

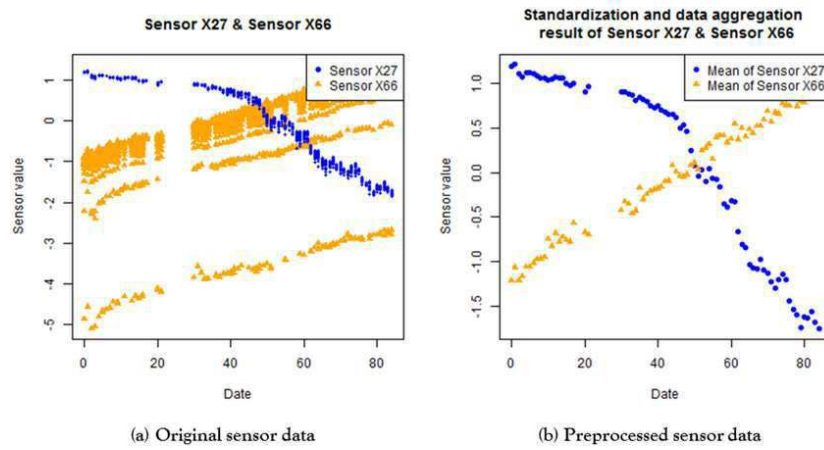


도면3

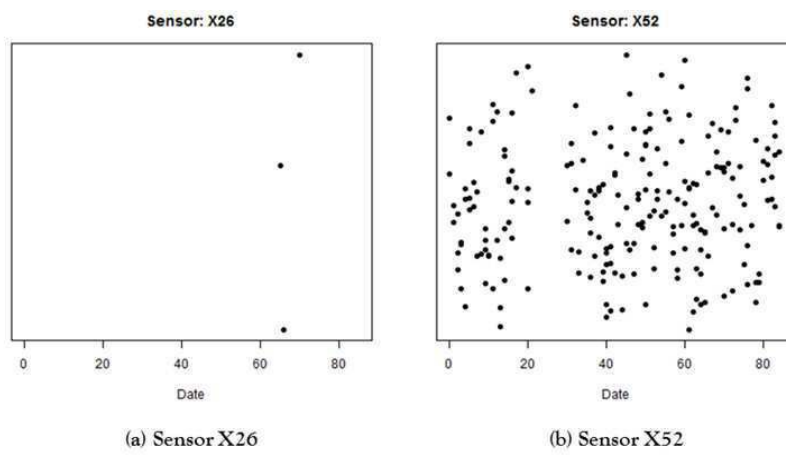
VIFs of FDC summary variables

Variable ^a	VIFs
Mean ^a	2864.23
SD ^a	8.313
MIN ^a	2029.44
MAX ^a	917.08

도면4



도면5



도면6

