

항로표지 장비용품의 고장예측 알고리즘 개발

김 환* · † 임성수

*충남대학교 컴퓨터공학과 석박사통합과정, † 충남대학교 컴퓨터공학과 교수

요약 : 다양한 소스로부터 수집되고 연동되는 데이터를 모델링하는 기술로 그래프 데이터베이스를 활용한 분석 기법이 각광받고 있다. 이 연구에서는 항로표지에서 관측되는 상태 및 주변 정보를 모델링하고, 고장진단 및 예측에 적용할 수 있는 기계학습 기법을 소개한다.

핵심용어 : 그래프 데이터베이스, 그래프 신경망, 시계열 분석, 패턴 인식, 이상 탐지

연구목표 및 내용

항로표지 고장 예측

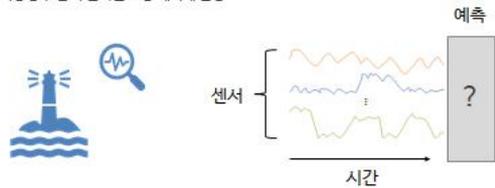
- 항로표지 데이터의 이해 및 모델링
 - 항로표지 관련 데이터 획득, 탐색적 데이터 분석, 특징 추출 등
 - 센서 간 관계를 고려한 그래프 모델링 → 그래프 기계학습 적용
- 항로표지 장비 이상 탐지 및 고장 예측 기술 개발
 - 벡터 표현 학습을 통한 시각화 → 중심 패턴과 얼마나 다른지 확인 가능
 - 이상치 점수 계산 → 다른 경도를 수치화하고, 높은 점수에 대한 근거 분석



연구목표 및 내용

항로표지 고장 예측

- 항로표지 상태 데이터 탐색
 - 항로표지의 센서를 통해 시간에 따라 변화하는 관측값 수집
 - 데이터 전처리 과정을 거쳐서 다변수 시계열 데이터로 변환
 - 시계열 데이터의 패턴을 학습하고 모니터링하여 이상탐지
 - 기존 패턴과 다른 데이터를 이상치 분류
 - 이상징후 탐지 결과를 고장 예측에 활용



데이터 엔지니어링

샘플 데이터 전처리 및 특징 추출

- 항로표지 상태 데이터 탐색
 - 항로표지의 센서를 통해 시간에 따라 변화하는 관측값 수집
 - 항로표지 상태 데이터의 이상탐지를 위한 샘플 데이터 활용
- 샘플 데이터: 월별 500만~700만 건
 - 2021년 1월: 6,029,648건
 - 2021년 2월: 5,465,276건
 - 2021년 3월: 5,861,851건
- 데이터 속성
 - 모든 속성값이 결측치인 경우 제거
 - 전압 및 전류와 관련된 속성 활용
 - 62개 중 8개 속성 위주 분석

	2017	2018	2019	2020	2021
1월	X	O	O	O	O
2월	X	O	O	O	O
3월	X	O	O	O	O
4월	X	O	O	O	O
5월	X	O	O	O	O
6월	X	O	O	O	O
7월	X	O	O	O	O
8월	X	O	O	O	O
9월	O	O	O	O	X
10월	O	O	O	O	X
11월	O	O	O	O	X
12월	O	O	O	O	X

데이터 엔지니어링

샘플 데이터 전처리 및 특징 추출

- 실험에 사용한 속성
 - ① 62개의 속성 중, 결측이 있는 속성은 모두 제외하고, 16가지의 속성으로 1차 선택
 - ② 실무자의 조언을 참고하여 항로 표지의 고장과 관련성이 높은 8가지의 속성 선택
 - Main volt, Solar volt, Battery volt, AIS current, Lantern current, Data logger current, Charger current, Discharger current
 - 전압 및 전류와 관련된 속성들로 선택 후 AI 모델을 통한 분석 시도

† 교신저자 : 정희원, sungsu@cnu.ac.kr
* hwan.kim@o.cnu.ac.kr

데이터 엔지니어링

샘플 데이터 전처리 및 특징 추출

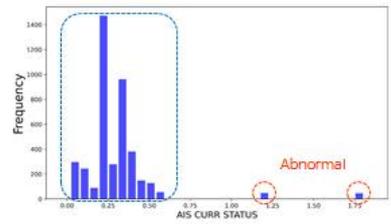
	MAIN VOLT	SOLAR VOLT	BATTERY VOLT	AIS CURR
평균	11.87	8.03	11.30	0.31
0을 제외한 평균	12.40	13.85	12.60	0.33
1보다 큰 값들의 평균	12.41	14.97	12.61	13.10

	LANTERN CURR	DATALOGGER CURR	CHARGER CURR	DISCHARGER CURR
평균	0.47	0.04	0.74	0.85
0을 제외한 평균	0.78	0.17	1.75	1.17
1보다 큰 값들의 평균	2.23	7.27	3.48	2.66

제한하는 알고리즘

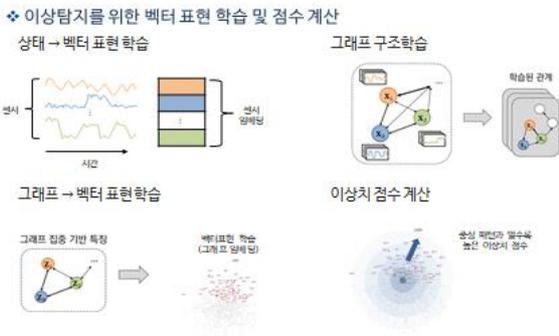
이상치 탐지 문제 및 제한 기법

- ❖ 이상치 (고장으로 예상되는 것)에 대한 정의
 - 비지도 학습 문제: 수집 가능한 데이터에서 실제 고장 상태 및 정상 상태에 대한 정보의 부재
 - 대다수의 데이터가 분포하는 부분에서 벗어난 포인트에서 고장 발생 확률이 높을 것이라 판단



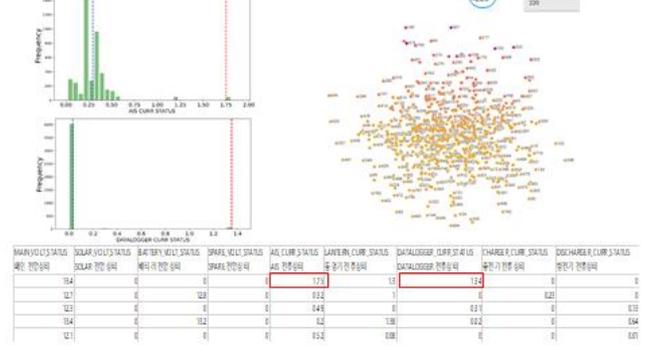
제한하는 알고리즘

이상치 탐지 문제 및 제한 기법



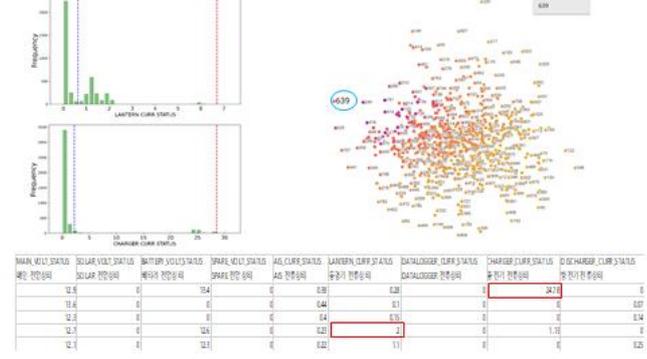
알고리즘 적용 결과

샘플 데이터 적용 및 결과 예시



알고리즘 적용 결과

샘플 데이터 적용 및 결과 예시



알고리즘 적용 결과

샘플 데이터 적용 및 결과 예시

