

데이터 큐브 모델과 SVM을 이용한 철도 선로전환기의 교체시기 탐지

Replacement Condition Detection of Railway Point Machines Using Data Cube and SVM

최용주*, 오지영**, 박대희***, 정용화***, 김희영****

(Yongju Choi, Jeeyoung Oh, Daihee Park, Yongwha Chung, Hee-Young Kim)

요약

철도 선로전환기는 열차의 진로를 현재의 궤도에서 다른 궤도로 제어하는 장치이다. 선로전환기의 노후화로 인한 이상 상황은 탈선 등과 같은 심각한 문제를 발생할 수 있기 때문에, 선로전환기의 적절한 교체시기를 결정하는 것은 매우 중요하다. 본 논문에서는 국내 철도 현장에서 획득한 선로전환기의 전류신호로부터 다차원 데이터 큐브를 구성하고 OLAP(On-Line Analytical Processing) 분석을 통하여 체계적으로 “교체가 필요한 데이터”와 “교체 시점이 아닌 데이터” 집합을 정제하여 분류하였다. 또한 선로전환기의 교체시기 탐지 문제를 이진 분류 문제로 해석하여 이진 분류기의 대표적 모델인 SVM(Support Vector Machine)을 탐지기로 설계함으로써 선로전환기의 노후화에 따른 적절한 교체시기를 탐지하는 시스템을 제안한다. 이때, 입력되는 전류 신호를 DWT(Discrete Wavelet Transform)와 PCA(Principal Components Analysis) 기법으로 고차원의 특징벡터 신호를 정보의 손실을 최소화하면서 저차원의 특징벡터로 변환한다. 실제 국내에서 운행 중인 선로전환기의 이상상황 정보가 포함된 대규모의 전류 신호를 이용하여 제안하는 시스템의 성능을 실험적으로 검증한 바 98%를 넘는 탐지 정확도를 확인하였다.

■ 중심어 : 철도 선로전환기; 교체시기 탐지; 데이터 큐브; OLAP; SVM;

Abstract

Railway point machines act as actuators that provide different routes to trains by driving switchblades from the current position to the opposite one. Since point failure caused by the aging effect can significantly affect railway operations with potentially disastrous consequences, replacement detection of point machine at an appropriate time is critical. In this paper, we propose a replacement condition detection method of point machine in railway condition monitoring systems using electrical current signals, after analyzing and relabeling domestic in-field replacement data by means of OLAP(On-Line Analytical Processing) operations in the multidimensional data cube into “does-not-need-to-be-replaced” and “needs-to-be-replaced” data. The system enables extracting suitable feature vectors from the incoming electrical current signals by DWT(Discrete Wavelet Transform) with reduced feature dimensions using PCA(Principal Components Analysis), and employs SVM(Support Vector Machine) for the real-time replacement detection of point machine. Experimental results with in-field replacement data including points anomalies show that the system could detect the replacement conditions of railway point machines with accuracy exceeding 98%.

■ keywords : Railway Point Machines; Replacement Condition Detection; Data Cube; OLAP; SVM;

I. 서론

열차 사고는 발생 시 막대한 경제적 손실 및 인명피해를 야기

함에 따라 열차 운행의 안정성은 매우 중요하다[1]. 특히, 열차의 진로를 변경시키는 핵심 장치인 선로전환기(그림 1 참조)의 결함은 열차의 탈선 및 충돌과 직결되는 중요한 요소 중 하나로, 철도 시스템에서 가장 중요한 설비 중 하나이다[2]. 철로

* 준회원, 고려대학교 컴퓨터융합소프트웨어학과

** 정회원, (주)세화 부설연구소 연구개발팀

*** 정회원, 고려대학교 컴퓨터융합소프트웨어학과

**** 정회원, 고려대학교 공공정책대학 국가통계전공

본 연구는 2015년도 중소기업청 기술혁신개발사업 혁신형기업과제(S2312692)의 지원을 받아 수행된 연구결과임.

접수일자 : 2017년 06월 07일

게재확정일 : 2017년 06월 29일

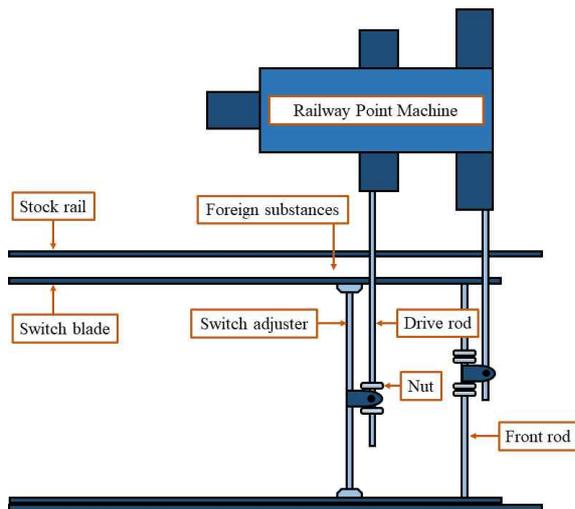
수정일자 : 2017년 06월 23일

교신저자 : 박대희 e-mail : dhpark@korea.ac.kr

를 달리는 열차의 방향을 한 선로에서 다른 선로로 분기하기 위하여 설치한 설비를 분기기라 하며, 이 분기기를 작동시키는 것을 선로전환기라고 한다. 외부 환경에 노출되어 작동하는 선로전환기의 특성 상, 선로전환기의 부품들은 외부 충격을 자주 받게 된다. 이러한 외부 충격들로 인해 선로전환기는 지속적인 물리적 손상을 받게 되며 이는 선로전환기의 미세한 균열이나 누적 손상을 야기하게 되어 결국 선로전환기의 결함으로 이어진다[3]. 현재, 국내에서는 관리자가 현장에 직접 나가서 선로전환기의 유지보수를 주기적으로 수행하고 있으나, 향후 ICT기술에 의한 선로전환기의 이상상황을 조기에 탐지하고, 유지보수 및 점검이 필요한 선로전환기에 대해서만 선택적으로 적절한 대처를 수행하는 지능화 시스템이 요구된다[4].



(a) 실제 현장에 설치된 선로전환기



(b) 선로전환기의 구성 요소(NS-AM타입)

그림 1. 철도 선로전환기

선로전환기에 관한 최근 연구동향을 살펴보면, 선로전환기의 결함 탐지 및 진단을 위한 많은 노력들이 철도 선진국을 중심으로 진행되고 있음을 알 수 있다. 그러나 선로전환기의 노후화에

따른 적절한 교체시기를 결정하는 현장중심의 보고서나 학술적 차원의 연구결과는 매우 제한적이다. 선로전환기의 교체시점은 철도 관리지침에 의하면 20만 회 이상 전환되거나 10년 이상 운행하였을 때 교체하는 것으로 알려져 있으나, 현장에서는 예산상의 문제 등으로 인하여 잘 지켜지지 않는 실정이다. 반면, 선로전환기의 교체시점에 관한 학술적 연구로는 Sa 등[5]의 연구가 유일하게 발견된다. Sa 등[5]은 실제 국내에서 운행 중인 선로전환기로부터 획득한 대규모의 전류 신호를 이용하여 shapelet 알고리즘으로 선로전환기의 교체시기를 탐지하는 연구결과를 최근 발표하였다.

본 연구는 Sa 등[5]의 연구의 연장선으로, 대규모의 실제 현장에서 획득한 선로전환기의 전류신호로 다차원 데이터 큐브를 구성하고 OLAP(On-Line Analytical Processing)분석을 통하여 체계적으로 “교체가 필요한 데이터”와 “교체 시점이 아닌 데이터” 집합을 정제하여 분류한다. 또한 선로전환기의 교체시기 탐지 문제를 이진 분류 문제로 해석하여 이진 분류기의 대표적 모델인 SVM(Support Vector Machine)을 탐지기로 설계함으로써 선로전환기의 노후화에 따른 적절한 교체시기를 탐지하는 시스템을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 선로전환기에 관련한 최근 연구방법론의 문헌고찰을 간략하게 소개하고, 3장에서는 본 연구에서 제안하는 다차원 데이터 큐브 모델을 이용한 선로전환기의 교체시기 탐지 시스템에 대해 상세히 기술한다. 4장에서는 실험 결과 및 성능 분석을, 마지막으로 5장에서는 결론 및 향후 연구과제에 대해 논한다.

II. 관련 연구

선로전환기에 관한 학술적·실천적 연구는 크게 선로전환기의 결함 탐지·진단과 선로전환기의 노후화에 따른 교체시점에 관한 연구로 나눌 수 있다. 또한, 사용하는 센서의 종류에 따라, 전기적 신호기반 혹은 선로전환기가 실제 구동 시 발생하는 소리 정보에 기반을 둔 연구로 구분된다.

Vileiniskis 등[6]은 선로전환기를 구동시키는 모터의 전류값의 변화를 이용하여, 미리 설정해 둔 임계값을 벗어나는지의 여부를 근거로 선로전환기의 이상상황을 판단하였다. Eker 등[7]은 모터 전류측정 센서 등을 이용하여 취득한 특징 정보들을 PCA(Principal Components Analysis)와 SVM을 이용하여 drive rod가 비정상적으로 작동하는 상태인 불일치(out-of-adjustment) 상태 여부를 탐지하였다. Asada 등[8, 9]은 선로전환기로 유입되는 전기적 신호(전류와 전압)를 이용하여 선로전환기의 상태를 모니터링 하는 시스템을 제안하였다. 또한 Kim 등[10]은 전류신호에 기반을 둔 DTW(Dynamic Time

Warping) 기법을 이용하여 학습과정 없이 실제 철도 현장에 바로 적용이 가능한 선로전환기의 이상상황을 탐지하는 실천적 차원의 고장탐지시스템을 소개하였다.

이상에서 살펴본 바와 같이, 최근 선로전환기에 관한 연구들은 대부분 선로전환기의 모터를 동작시키기 위해 입력되는 전기적 신호를 단순한 임계값 수치로 판단하거나 혹은 이상상황 탐지를 분류의 문제로 해석하여 기계학습기반의 분류기로 문제를 해결하고 있음을 알 수 있다. 반면, Lee 등[4]은 선로전환기가 실제 구동 시 발생하는 소리 정보를 대상으로, MFCC (Mel-Frequency Cepstrum Coefficients)특징 벡터를 추출하고 SVM을 이용하여 선로전환기의 이상상황을 탐지 및 분류하는 연구결과를 최근 발표하였다. 이는 전기적 신호를 주요 입력으로 하는 선로전환기의 현재까지의 연구영역에서 소리분석 기반의 연구가 독립적으로 혹은 전기적 신호 시스템에 보조적 역할을 담당할 수 있다는 가능성을 보여준 최초의 연구결과이다.

반면, 본 연구의 주제인 선로전환기의 노후화에 따른 교체시점에 관한 연구로써, Choi 등[11]은 선로전환기의 노후화와 직결되는 선로전환기의 스트레스 탐지를 위해 소리 신호를 이용하여 적절한 특징벡터를 선택한 후, SVM을 적용하여 선로전환기의 스트레스 여부를 탐지하였다. Sa 등[5]은 실제 국내에서 운행 중인 선로전환기로부터 획득한 대규모의 전류 신호를 이용하여 shapelet 알고리즘으로 선로전환기의 교체시기를 탐지하였다. 그러나 위 연구에서는 철로 전문가의 개인적 직관에 의존한 교체 전과 후로 레이블된 전류 신호 데이터를 사용하여 shapelet 알고리즘의 학습 및 검증 실험을 했기에 인간에 의한 분류 오류들이 시스템에 내재될 수밖에 없다는 한계점에서 자유롭지 못하다.

III 데이터 큐브 모델과 SVM을 이용한 선로전환기의 교체시기 탐지 시스템

본 논문에서 제안하는 데이터 큐브 모델과 SVM을 이용한 선로전환기의 교체시기 탐지 시스템은 아래 그림 2와 같은 구조를 갖는다. 제안하는 시스템은 데이터 전처리 모듈(data preprocessing module), 특징 선택 모듈(feature selection module), 교체시기 탐지 모듈(replacement detection module)로 구성된다.

3.1 데이터 전처리 모듈

수년간 축적된 RDW(Railway Data Warehouse) 정보를 이용한 데이터 큐브의 구축을 위하여 Oracle AWM(Analytic Workspace Manager)[12]을 이용한다. 선로전환기의 이상상황 분석을 위한 데이터 큐브 모델은 가장 보편적인 스타 스키마(star schema)(그림 3)를 다차원 분석 모델로 사용한다[13]. 데이터 큐브는 데이터가 여러 차원으로 모델링되며, 차원(dimension)과 사실(fact)로 정의된다. 본 연구에서는 4개의 차원(시간, 선로전환기 이상상황 발생 수, 선로전환기 설치 장소, 선로전환기 운행 라인)을 정의하고 각 차원의 추상화 정도에 따라 다양한 OLAP 연산을 통하여 다차원적인 분석을 수행한다. 개념계층(concept hierarchy)은 차원 테이블에서의 하위 개념 집합으로부터 보다 상위의 일반적인 개념들로의 사상의 연속(a sequence of mapping)을 의미하며 추상화정도를 조절하여 다차원적 분석을 수행하는데 활용된다[14].

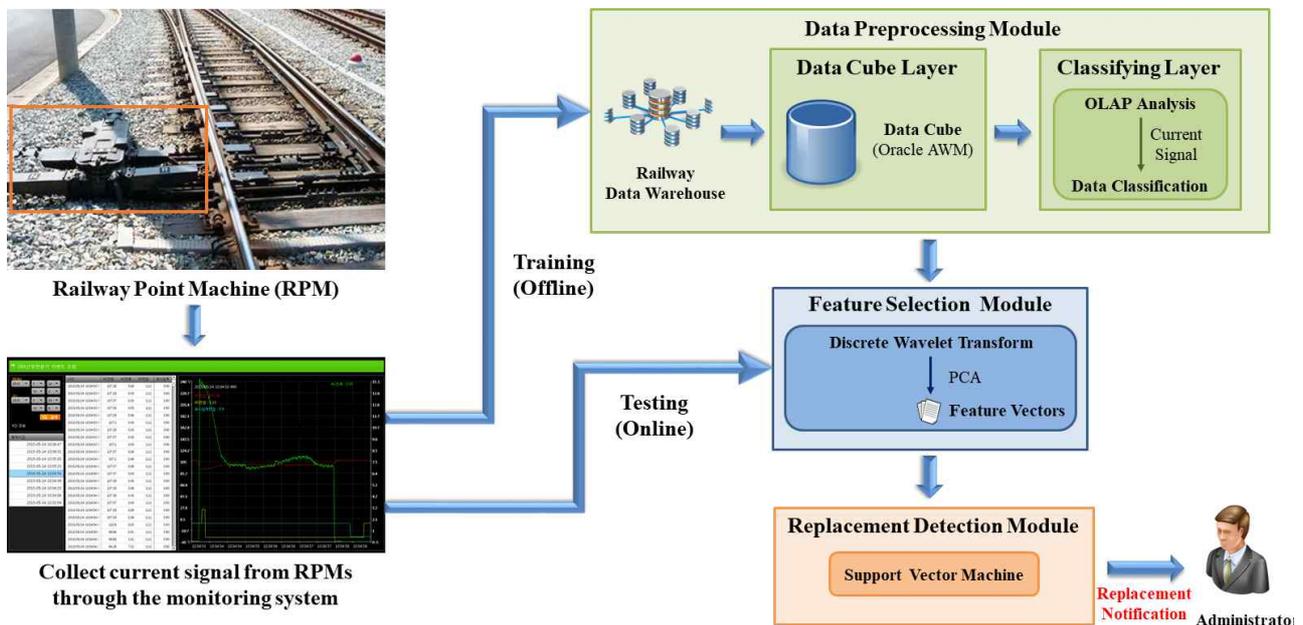


그림 2. 데이터 큐브 모델과 SVM을 이용한 선로전환기의 교체시기 탐지 시스템

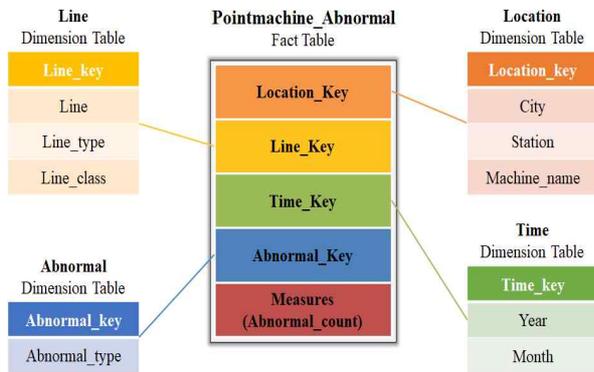


그림 3. 선로전환기의 데이터 큐브 모델의 스타 스키마

OLAP은 방대한 양의 이력 데이터를 관리하며, 요약(summarization)과 집계(aggregation)를 위한 도구를 제공한다. 본 연구에서는 선로전환기의 이상상황 빈도수를 중심으로 롤업(roll-up), 드릴다운(drill-down), 슬라이스(slice)와 다이스(dice) 등의 OLAP 연산을 이용한 OLAP 분석을 통하여 체계적으로 “교체가 필요한 데이터”와 “교체 시점이 아닌 데이터” 집합을 정제하여 분류한다.

3.2 특징 선택 모듈

실제 국내에서 운행 중인 선로전환기의 모니터링 시스템으로부터 획득한 전류 신호를 DWT(Discrete Wavelet Transform, 이산 웨이블릿 변환)을 이용하여 아날로그 신호를 이산화한 후, PCA(Principal Component Analysis) 기법으로 고차원의 특징벡터 신호를 정보의 손실을 최소화하면서 저차원의 특징벡터로 변환한다. DWT는 입력신호를 웨이블릿 함수와 스케일 함수에 각각 곱하여 고주파와 저주파로 주파수 대역을 나누어 원하는 주파수 대역에서의 변환된 신호를 분석할 수 있으며 수식 1과 같이 표현된다[15, 16]. 본 연구에서는 웨이블릿 변환 중에서 가장 일반적이고 간결한 형태로, 계산속도가 빠른 haar 웨이블릿을 이용하여 DWT를 수행한다. 반면, PCA는 고차원의 정보를 최대한 유지하면서 차원 축소를 통하여 데이터의 특징을 표현하는 기법으로 신호처리, 특징추출, 자료의 차원 감소, 영상처리 등 많은 분야에서 사용되고 있다[17, 18].

$$DWT(i, j) = \sum_i \sum_j x(t) 2^{-i/2} \Psi(2^{-i}t - k) \quad (1)$$

위 수식에서 $x(t)$ 는 시간 t 의 함수이며, Ψ 는 웨이블릿 기저 함수를 나타내고, i 와 j 는 주파수 해상도와 시간 축에 대한 이동 시간을 의미한다.

3.3 교체시기 탐지 모듈

교체시기 탐지 모듈에서는 노후화에 따른 선로전환기의 교체시기 탐지를 위해 미리 훈련을 마친 SVM을 기반으로, 실시간

으로 유입되는 전류 신호의 교체시기 여부를 탐지한다. 이진 분류 문제에서 많이 활용되는 SVM은 지도학습(supervised learning) 기반으로 두 클래스 간의 최적의 선형 결정 평면을 찾음으로써 두 개의 클래스를 분류하는 분류기이다[13]. 이때 결정 평면은 학습 원소들의 가중화된 조합이며, 이러한 학습 원소들을 서포트 벡터(support vector)라 부르며 이들은 두 클래스간의 경계면에 존재한다. 예를 들면, 선형 분리가 가능한 데이터를 가정할 때 최대 여백 분류의 목표는 서포트 벡터들의 거리가 최대가 되는 초평면(hyperplane)에 의해 두 개의 클래스를 분리하는 것이다. 이러한 초평면은 최적 경계 초평면(optimal separating hyperplane)이라고 불리며, 2차 프로그래밍 문제의 해를 구함으로써 서포트 벡터를 구한다. 또한 선형 분리가 불가능한 데이터인 경우에는, 입력 벡터를 고차원의 특징 공간으로 비선형 매핑한다(그림 4 참조). 이때 목표 함수와 결정 함수가 벡터의 내적으로 표현됨에 따라, 계산적으로 복잡한 매핑을 명시적으로 계산할 필요 없이, 벡터 연산 대신에 Mercer 조건을 만족하는 커널 함수로 치환할 수 있다[19]. 본 연구에서는 SVM의 커널 함수로 방사 기저 함수(Radial Basis Function, RBF)를 사용한다.

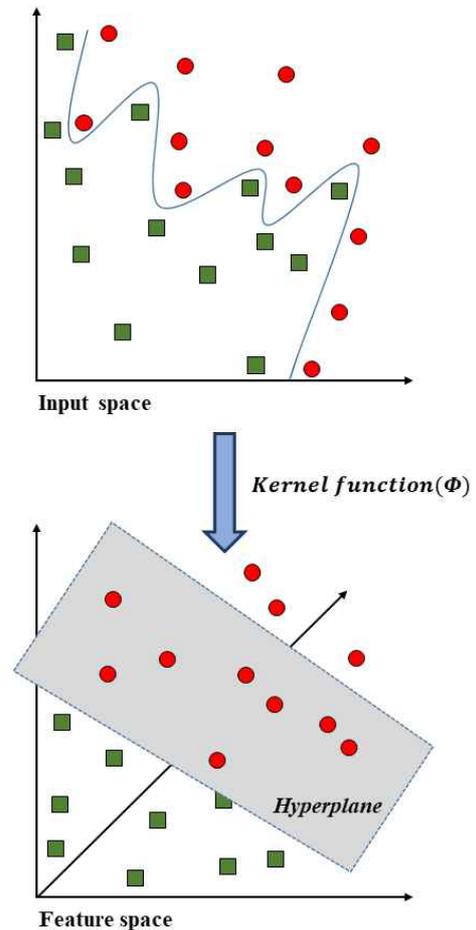


그림 4. 비선형 SVM의 커널 트릭

IV. 실험 및 결과 분석

4.1 실험 데이터 획득

실험을 위한 데이터는 국내 7개 역사에서 취득한 선로전환기의 전류 신호를 이용하였다. Oracle database를 이용하여 2013년 02월부터 2016년 01월까지 약 3년간의 선로전환기의 누적 전류 신호를 데이터 큐브로 구축하였다. 표 1은 국내 7개 역사에서 취득한 선로전환기의 실제 교체에 관한 기초 통계표이다. 선로전환기는 관리지침에 의하면 20만 회 이상 전환되거나 10년 이상 운행하였을 때 교체하는 것으로 알려져 있다[20]. 그러나 표 1의 기초 통계를 보면, 모든 선로전환기들이 10년 이상 사용되었으나 전환 횟수가 20만회 미만인 것들이 대부분임을 알 수 있다. 결국, 아직 운행이 가능한 선로전환기임에도 불구하고 관리자들의 주관적 판단에 의해 많은 선로전환기들이 현장에서 교체되고 있음이 확인된다.

표 1. 선로전환기의 교체에 관한 기초 통계표

역	선로전환기 개수	교체 전 전환횟수	교체 전 운행기간
A	15	1,284~33,272	12년
B	13	653~19,391	12~14년
C	17	12,875~107,927	12~14년
D	2	11,442~137,370	10~13년
E	7	5,778~391,141	12~16년
F	5	5,209~82,795	13~14년
G	8	436~108,600	14~17년

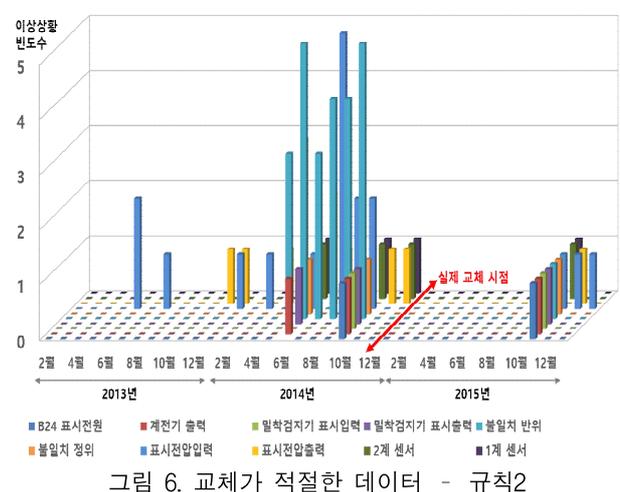
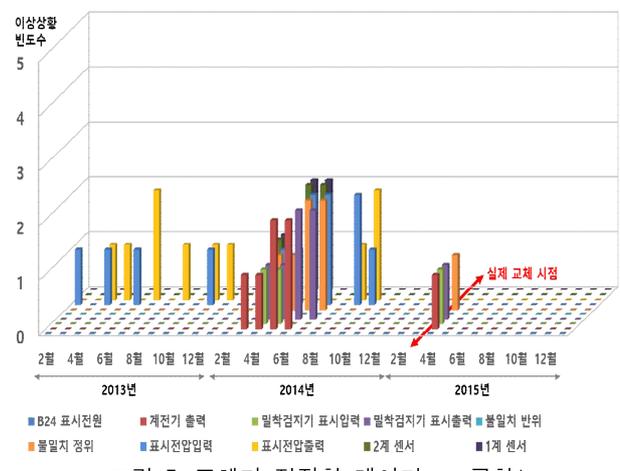
4.2 분류 규칙의 정의

Oracle AWM을 이용하여 선로전환기가 작동하였던 시간의 흐름에 따라 이상상황 발생 빈도수를 확인하는 OLAP 연산을 수행함으로써, 표 2와 같은 선로전환기의 교체시점에 관한 분류 규칙을 새롭게 정의하였다. 선로전환기의 이상상황으로 인한 결함 및 장애는 인명피해 및 경제적인 손실로 이어지기 때문에 선로전환기의 안전성에 비중을 두어 비교적 보수적으로 규칙들을 설정하였다.

표 2. 데이터 분류 규칙

규칙	교체가 필요한 데이터 (교체가 적절했던 상태)	교체 시점이 아닌 데이터 (조기에 교체한 상태)
1	1달에 3번 이상, 이상상황이 발생한 경우	이상상황이 없는 경우
2	2달 연속으로 2번 이상, 이상상황이 발생한 경우	이상상황은 있지만, 1달에 1~2번 간헐적으로 발생한 경우
3	3달 연속으로 1번 이상, 이상상황이 발생한 경우	

그림 5~7은 표 2의 분류 규칙에 따라 교체가 필요한 데이터 (교체가 적절했던 상태)로 분류된 예시, 그림 8~9는 교체 시점이 아닌 데이터(아직 운행이 가능한 상태)로 분류한 예시를 보여준다. 가로축은 월 단위로 구분된 시간 축이며, 세로축은 이상상황의 종류별 발생 수를 의미한다. 선로전환기의 이상상황 중 대부분을 차지하는 AS전압은 선로전환기의 제어 전압 장애이며 이는 관리자가 설정한 임계값에 의해 이상상황 여부가 판단된다. 임계값은 선로전환기의 노후화 정도와 조정·기계실과 현장의 설치 위치 등에 따라 다르기 때문에 유동적인 값을 갖는다. 결국, AS전압은 실제 선로전환기의 결함이나 이상상황 발생과는 큰 연관성이 없다[20]. 따라서 본 논문에서는 AS전압의 이상상황은 고려하지 않았다. 표 2의 규칙을 기반으로 실제 현장에서 교체가 이루어진 데이터들 대상으로 본 연구에서 새롭게 분류한 결과, 교체가 필요한 데이터(교체가 적절했던 상태)는 327개, 교체 시점이 아닌 데이터(아직 운행이 가능한데 교체한 상태)는 825개로 분류되었다.



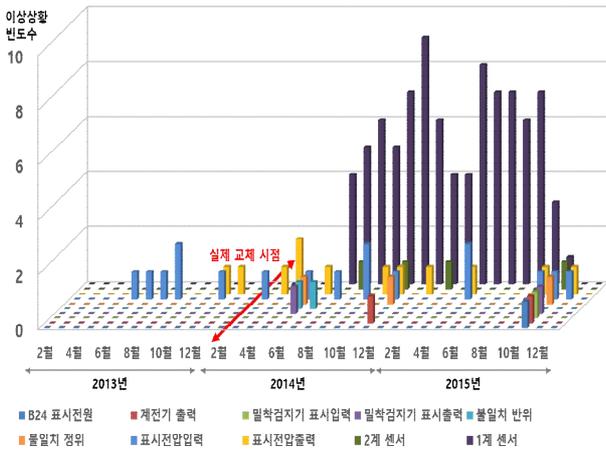


그림 7. 교체가 적절한 데이터 - 규칙3

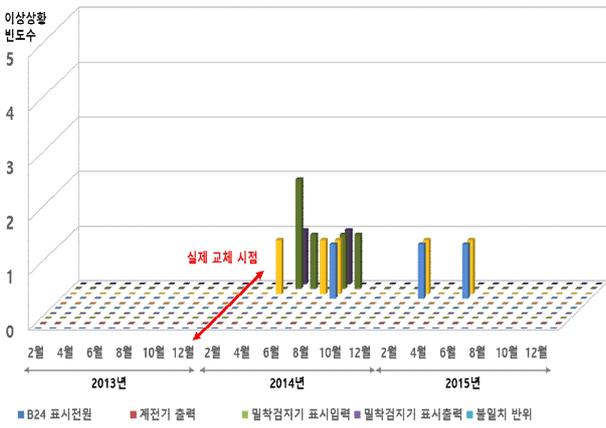


그림 8. 조기 교체한 데이터 - 규칙1

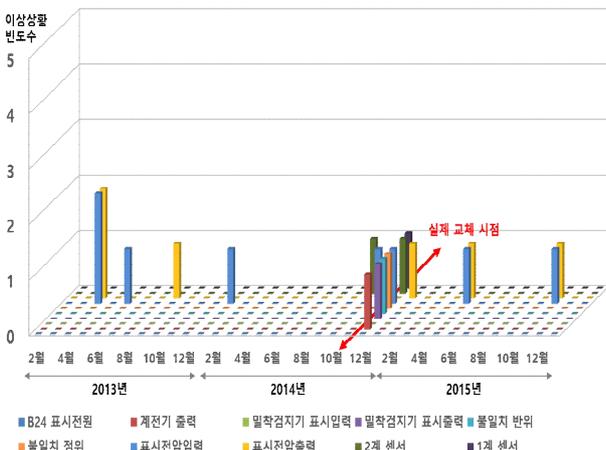


그림 9. 조기 교체한 데이터 - 규칙2

4.3 특징 추출 및 특징 선택

전류 신호로부터 이산화 특징벡터를 추출하기 위해 Matlab(2015a)을 이용하여 이산 웨이블릿 변환을 수행하였다. 선로전환기로부터 수집한 전류 신호는 0.01초 단위로 샘플링 되

었으며, 각각의 전류 신호들은 354~582(3.54초~5.82초)의 다양한 길이를 가진다. 이산 웨이블릿 변환을 통해 292의 차원을 갖는 특징벡터가 생성 되었으며, PCA 수행으로 전체 292개의 특징 중 14개의 주성분이 선택되어 전체 분산의 97.013%의 설명력을 갖는 95.20%의 차원감소 효과를 확인하였다.

4.4 선로전환기의 교체시기 탐지

전류 신호를 이용한 선로전환기의 교체시기 탐지 실험은 교체가 필요한 데이터 327개, 교체 시점이 아닌 데이터 825개의 전류 신호를 대상으로 R 3.2.3의 SVM 라이브러리를 이용하여 교체시기 탐지 실험을 수행하였다. 전체 데이터 셋의 50%를 무작위로 선택한 데이터로부터 그리드 검색(grid search) 방법을 이용하여 SVM의 상대적 중요성을 조정하는 상수(trade-off constant) c 값은 8.293으로, RBF kernel의 계수 값인 γ 는 0.204로 설정하였고, 학습에 사용하지 않은 나머지 50%의 데이터 셋에 대하여 교체시기를 판별하는 성능 실험을 수행하였다.

실험 성능 측정을 위한 지표로는 RDR(Replacement Detection Rate: 교체 탐지율), FPR(False Positive Rate: 교체 시점이 아닌 데이터가 교체가 필요한 데이터로 잘못 탐지된 비율), FNR(False Negative Rate: 교체가 필요한 데이터가 교체 시점이 아닌 데이터로 잘못 탐지된 비율)을 사용하였다[13]. 각각의 수식은 다음과 같다.

$$RDR = \frac{\sum_{i=1}^m T_i}{\sum_{i=1}^n I_i} \tag{2}$$

$$FPR = \frac{\sum_{i=1}^k P_i}{\sum_{i=1}^l N_i} \tag{3}$$

$$FNR = \frac{\sum_{i=1}^{n-m} F_i}{\sum_{i=1}^n I_i} = 1 - \frac{\sum_{i=1}^m T_i}{\sum_{i=1}^n I_i} \tag{4}$$

위 식들에서 I_i 는 교체가 필요한 데이터의 개수, T_i 는 교체가 필요한 데이터를 교체가 필요한 데이터로 정상적으로 탐지한 데이터의 개수, N_i 는 교체 시점이 아닌 데이터의 개수, P_i 는 교체 시점이 아닌 데이터를 교체 시점인 데이터로 잘못 탐지한 데이터의 개수, F_i 는 교체가 필요한 데이터를 교체 시점이 아닌 데이터로 잘못 판단한 데이터의 개수를 의미한다.

본 논문에서 제안한 선로전환기의 교체시기 탐지 시스템의 성능 실험 결과는 표 3과 같으며, 선로전환기의 RDR은 98.78%, FPR과 FNR은 각각 1.70%, 1.22%라는 안정적인 실험결과를 확인하였다.

표 3. 교체 탐지 실험 결과

RDR	FPR	FNR
98.78%	1.70%	1.22%

선행 연구와의 비교 실험을 위해서 Sa 등[5]의 연구 방법론을 선택하여, 본 연구와 건설적 차원에서 비교 평가하였다. 선행연구 [5]는 본 연구에서 사용한 데이터 셋과 동일하기는 하나 철로 전문가의 개인적 직관에 의존하여 수작업으로 교체 전과 후로 레이블된 전류 신호 데이터를 사용하였다. 또한 [5]의 데이터 분류 방법론인 웨이프릿(shapelet) 알고리즘은, 시계열 데이터에 대하여 동일 클래스 내의 유사한 특징(동시에 다른 클래스 사이의 차별화되는 특징)을 가지는 모양(shapelet)을 학습을 통해 찾아내고, 의사 결정 트리를 이용하여 분류하는 기법[21, 22]으로, 전류신호의 정규화 전처리 작업 후 전류 신호에 대한 시계열 shapelet 정보만을 고려한다. 따라서 shapelet 알고리즘은 본 연구의 방법론과 달리 선로전환기의 노후화와 밀접한 상관관계를 갖는 선로전환기의 이상상황에 대한 중요한 정보를 판단 과정에서 고려하지 않는다. 본 연구에서 제안한 방법에 따라 새롭게 분류된 데이터 집합으로 선로전환기의 교체시기를 shapelet 알고리즘으로 실험하였다. 이번 shapelet 알고리즘의 실험에서는 두 클래스 간의 차이를 확연하게 구분하는 shapelet을 찾지 못했으며, 따라서 만족스러운 성능결과를 실험적으로 얻지 못했다(RDR=63.41%, FPR=17.48%, FNR=36.59%).

V. 결론 및 향후 연구

본 논문에서는 국내 7개 역사에서 약 3년간 취득한 선로전환기의 전류 신호로 다차원 데이터 큐브를 구성하고 OLAP분석을 통하여 체계적으로 “교체가 필요한 데이터”와 “교체 시점이 아닌 데이터” 집합을 정제하여 분류한 후, 전류 신호를 DWT와 PCA를 이용하여 최종적인 특징 벡터를 선택하였다. 마지막으로, 선로전환기의 교체시기 탐지 문제를 이진 분류 문제로 해석하여 기계학습의 대표적인 모델인 SVM을 이용하여 교체시기 여부를 판별하였다. 실제 국내에서 운행 중인 선로전환기의 이상상황 정보가 포함된 대규모의 전류 신호를 이용하여 제안된 시스템을 실험적으로 검증하였다.

향후 연구 과제로는 본 논문에서 제안된 프로토타입의 시스템을 실세계에서 구현·운용하기 위한 실용화 차원의 후속 연구

들이 요구되며, 선로전환기의 교체시기를 예측하기 위한 통계학적 예측 모델을 추후 설계하는 연구가 기대된다.

References

- [1] M. Guo, W. Wei, G. Liao, and F. Chu, “The Impact of Personality on Driving Safety among Chinese High-speed Railway Drivers,” *Accident Analysis & Prevention*, vol. 92, pp. 9–14, July. 2016.
- [2] J. H. Lee and Y. K. Kim, “A Study on Switching Power Measurement of an Electrical Point Machine Using a Sensor,” *Journal of the Korean Society for Railway*, vol. 18, no. 4, pp. 335–343, August. 2015.
- [3] C. S. Kim and G. H. Kang, “Fatigue Analysis of Reduction Gears Unit in Rolling Stock Considering Operating Characteristics,” *Journal of the Korea Academia Industrial Cooperation Society*, vol. 12, no. 3, pp. 1085–1090, March. 2011.
- [4] J. Lee, H. Choi, D. Park, Y. Chung, H.Y. Kim, and S. Yoon, “Fault Detection and Diagnosis of Railway Point Machines by Sound Analysis,” *Sensors*, vol. 16, no. 4, pp. 549, April. 2016.
- [5] J. Sa, Y. Choi, Y. Chung, H.Y. Kim, D. Park, and S. Yoon, “Replacement Condition Detection of Railway Point Machines Using an Electric Current Sensor”, *Sensors*, vol. 17, no. 2, pp. 263, January. 2017.
- [6] M. Vileiniskis, R. Remenyte-Priscott, and D. Rama, “A Fault Detection Method for Railway Point Systems,” *In Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 230, no. 3, pp. 852–865, February. 2016.
- [7] O. Eker, F. Camci, and U. Kumar, “SVM based Diagnostics on Railway Turnouts,” *International Journal of Perform Ability Engineering*, vol. 8, no. 3, pp. 289–398, May. 2012.
- [8] T. Asada and C. Roberts, “Improving the Dependability of DC Point Machines with a Novel Condition Monitoring System,” *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 227, no. 4, pp. 322–332, March. 2013.
- [9] T. Asada, C. Roberts, and T. Koseki, “An Algorithm for Improved Performance of Railway

- Condition Monitoring Equipment: Alternating Current Point Machine Case Study,” *Transportation Research Part C: Emerging Technologies*, vol. 30, pp. 81-92, May. 2013.
- [10] H. Kim, J. Sa, Y. Chung, D. Park, and S. Yoon, “Fault Diagnosis of Railway Point Machines Using Dynamic Time Warping,” *Electronics Letters*, vol. 52, no. 10, pp. 818-819, May. 2016.
- [11] Y. Choi, J. Lee, D. Park, J. Lee, Y. Chung, H.Y. Kim, and S. Yoon, “Stress Detection of Railway Point Machine Using Sound Analysis,” *KIPS Tr. Software and Data Eng*, vol. 5, no. 9, pp. 433-440, July. 2016.
- [12] AWM <http://www.oracle.com/technetwork/database/options/olap/index.html>
- [13] J. Han, M. Kamber, and J. Pei, *Data Mining Concepts and Techniques*, Third Edition, Morgan Kaufmann Publishers, 2012.
- [14] B. Noh, J. Lee, D. Park, and Y. Chung, “Crisis Management Analysis of Foot-and-Mouth Disease Using Multi-dimensional Data Cube,” *Journal of The Korea Contents Association*, vol. 17, no. 5, pp. 565-573, May. 2017.
- [15] I. Daubechies, “The Wavelet Transform, Time Frequency Localization and Signal Analysis,” *IEEE Trans. Information Theory*, vol. 36, no. 5, pp. 961-1005, 1990.
- [16] M.J. Shensa, “The Discrete Wavelet Transform: Wedding the a Trous and Mallat Algorithms,” *IEEE Transactions on Signal Processing*, vol. 40, no. 10, pp. 2464-2482, October. 1992.
- [17] I. Jolliffe, *Principal Component Analysis*, John Wiley and Sons, Ltd. 2002.
- [18] H. Moon, J. Park, and S. Pan, “Performance Analysis of Face Recognition by Face Image resolutions using CNN without Backpropergation and LDA,” *Smart Media Journal*, vol. 5, no. 1, pp. 24-29. March. 2016.
- [19] H. Kim, S. Lee, Y. Chung, D. Park, and H. Lee, “Multi Core Processor based Parallel SVM for Video Surveillance System,” *Journal of the Korea Institute of Information Security and Cryptology*, vol. 21, no. 6, pp. 161-169, December. 2011.
- [20] H. Choi, “Big Data Analysis of Railway Point Machine Framework Design and Implementation”, *Master’s Thesis of Graduate School Korea University*, February. 2017.
- [21] L. Ye and E. Keogh, “Time Series Shapelets: A New Primitive for Data Mining,” *In*

Proceedings of the ACM, pp. 947-956, June. 2009.

- [22] T. Rakthanmanon and E. Keogh, “Fast Shapelets : A Scalable Algorithm for Discovering Time Series Shapelets,” *In Proceedings of the 2013 SIAM International Conference on Data Mining, Society for Industrial and Applied Mathematics*, pp. 668-676, May. 2013.

저자 소개



최용주(준회원)

2017년 고려대학교 컴퓨터정보학과 학사 졸업.

2017년~현 재 고려대학교 컴퓨터정보학과 석사 과정.

<주관심분야 : 빅데이터, 데이터마이닝, 기계학습, 딥러닝, 인공지능>



오지영(정회원)

2001년 한밭대학교 전자계산학과 학사 졸업.

2003년 한밭대학교 컴퓨터공학과 석사 졸업.

2008년 한밭대학교 컴퓨터공학과 박사 졸업.

2016년~현 재 (주)세화부설연구소 연구개발팀 재직 중.

<주관심분야 : 빅데이터, 머신러닝, 데이터마이닝, Machine Care Prediction>



박대희(정회원)

1982년 고려대학교 수학과 학사 졸업.

1984년 고려대학교 수학과 석사 졸업.

1989년 플로리다 주립대학 전산학과 석사 졸업.

1992년 플로리다 주립대학 전산학과 박사 졸업.

1993년~현 재 고려대학교 컴퓨터융합소프트웨어학과 교수.

<주관심분야 : 빅데이터, 데이터마이닝, 인공지능, 융합 IT>



정용화(정회원)

1984년 한양대학교 전자통신공학과
학사 졸업.

1986년 한양대학교 전자통신공학과
석사 졸업.

1997년 U. of Southern California 박
사 졸업.

1986년~2008년 한국전자통신연구원
생체인식기술 연구팀 팀장.

2003년~현재 고려대학교 컴퓨터융합소프트웨어학과
교수.

<주관심분야 : 병렬처리, 영상처리, 융합IT>



김희영(정회원)

1991년 고려대학교 수학과 학사 졸업.

1995년 고려대학교 통계학과 석사 졸
업.

2002년 고려대학교 통계학과 박사 졸
업.

2013년~현재 고려대학교 공공정책
대학 국가통계전공 교수.

<주관심분야 : Martingales, Queueing theory, Markov
process, Financial time series analysis>