



을 위해 표적정보를 자동으로 전파 및 공유하는 기능을 제공한다<sup>(2)</sup>. 1999년 1차 사업을 시작으로, 전담급 이상 육상지휘소와 전담감시대, 지원부대 및 함정에 설치되었다. 체계의 개략적인 구성도는 <그림 2>와 같으며, 구성장비 중 핵심장비에 대한 설명은 <표 1>과 같다.

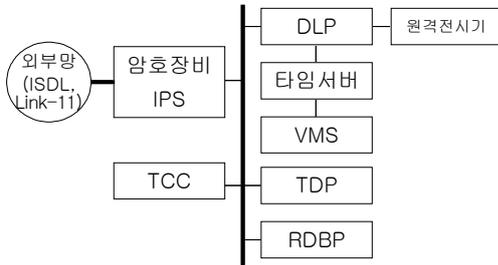


Fig. 2. Composition of KNTDS

Table 1. Function of Core Parts

장비	주요 기능
데이터링크처리기 (DLP)	· 데이터링크와의 연동 · 데이터 포맷 변환
전술자료처리기 (TDP)	· 전술정보 실시간 처리 · TCC 정보 요구에 대응
관계형데이터베이스 처리기(RDBP)	· 전술/일반자료 저장 관리 · 전술자료 백업
운용자콘솔(TCC)	· 처리된 전술정보 전시 · 운용자 편의기능 제공

지휘통제체계 유지보수는 체계의 운영 과정에서 체계 운영 전담조직이 해결할 수 없는 작업(장애 복구, 요구사항 및 문제점 해결 등)을 유지보수 기관이나 용역업체에서 처리하는 것을 의미한다. 유지보수는 체계의 수명주기 중에서 가장 긴 기간을 차지하고 있고, 그렇기 때문에 체계의 신뢰성 보장을 위한 매우 중요한 활동이라고 할 수 있다<sup>(3)</sup>. 유지보수 기간 중에 발생한 장애에 대한 데이터를 효과적으로 활용하면 유지보수의 연속성과 용이성을 향상시킬 수 있고, 이는 비용적인 측면에서도 절감 효과를 가져 올 수 있다.

2.2. Feature Selection(속성선택)

데이터 마이닝은 방대한 데이터로부터 유용한 정보를 추출하고, 데이터 간 상관관계를 발견하여 의사 결정에 이용하는 과정을 의미한다. 현재 데이터 마이닝을 지원하는 도구가 많이 개발되어 있으며, 정보를 추출하기 위한 데이터셋(Dataset)이 확보되어 있으면 도구를 이용하여 손쉽게 정보를 추출할 수 있다.

지휘통제체계 장애에 관한 데이터셋을 작성하여 데이터 마이닝 도구를 통해 분석하면 체계의 장애와 가장 연관이 깊은 속성이 도출되고, 장애가 발생하기 전에 장애 발생 가능성을 어느 정도 예측 가능한 수준까지 도달할 수 있다. 또한 도출된 속성을 앞서 언급한 체계 유지보수에 활용한다면 선제적이고 신속한 장애 복구에 도 크게 기여할 수 있다.

Feature Selection(속성선택)은 데이터 마이닝 기법 중 하나로, 방대한 데이터셋에서 특정 데이터로부터 가장 좋은 성능을 나타내는(가장 연관이 깊은) 데이터의 속성을 찾아내는 것이다<sup>(4)</sup>. Feature Selection을 지원하

는 도구의 대표적인 작동 알고리즘은 Search Method(탐색 방법)와 Feature Evaluator(특징 평가)가 있으며 종류는 <그림 3>과 같다.

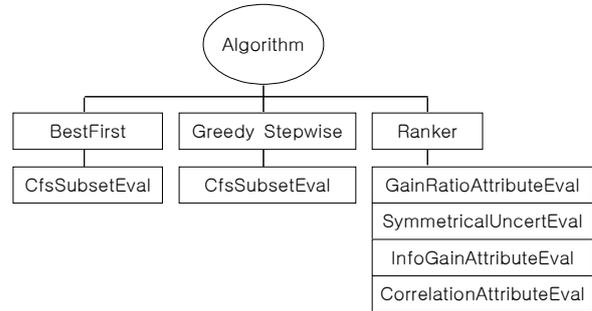


Fig. 3. Search Method and Feature Evaluator Algorithm

Search Method의 하위 알고리즘에는 BestFirst, Greedy Stepwise, Ranker가 있고, Feature Evaluator의 하위 알고리즘에는 CfsSubsetEval, GainRatioAttributeEval, SymmetricalUncertEval, InfoGainAttributeEval, CorrelationAttributeEval이 있다. 각 알고리즘별 기능은 <표 2>와 같다.

Table 2. Function of Algorithm

알고리즘	기능
BestFirst	· 모든 속성을 순서화하여 깊이 우선 탐색을 통해 값을 도출
Greedy Stepwise	· 속성의 하위 데이터의 임의 공간에서 시작하여 정방향/역방향 탐색 수행
Ranker	· 속성별 개별 평가를 통해 속성의 순위를 계산
CfsSubset Eval	· 각 개별 속성의 예측 가능성과 데이터 중복 정도를 고려하여 속성의 가치를 평가
GainRatio AttributeEval	· 속성에 대한 이득 비율을 산정
Symmetrical UncertEval	· 속성값의 대칭 정도를 측정하여 속성 평가
InfoGain AttributeEval	· 속성과 관련된 정보 획득을 측정
Correlation AttributeEval	· 속성 간의 상관관계를 측정하여 평가

3. 해군 지휘통제체계 장애 분석

3.1. 데이터 전처리와 실험 환경

지휘통제체계의 장애 분석을 위한 도구를 사용하기에 앞서 우선 도구가 입력받을 수 있는 CSV(Comma Separated Values) 파일 형태의 데이터셋이 필요하다.

본 연구에서는 2016년 1월부터 2018년 12월까지 3년 동안의 0함대사 지휘통제체계 장애현황 데이터를 기반으로 데이터셋을 작성하였으며, 그 속성은 <표 3>과 같다. 1, 2, 3번 속성은 3년간의 날짜를 1일 단위로 넘버링하고, 4번 속성은 운용부대, 5번 속성은 설치장비 그리고 6번 속성은 정상 혹은 장애 여부를 나타낸

다. 데이터셋을 작성하기 위한 프로그램으로는 MS Excel을 이용하며, Excel에서는 CSV 파일 형태를 지원한다. <표 4>는 분석을 위한 환경 구성을 나타낸다.

Table 3. Failure Data Attributes

1	2	3
year	month	date
4	5	6
unit	component	result

Table 4. Analysis Environment

구분		내용
장비	CPU	Intel(R) Core(TM) i5-3550 CPU @ 3.3GHz
	RAM	8GB
	HDD	465GB
	OS	Windows 7 Ultimate K
SW	Dataset	Microsoft Excel 2010
분석 도구		Weka 3.8.3

3.2. 분석 결과

데이터 전처리 과정에서 선정한 속성 중 장애와 연관 깊은 속성을 추출하기 위해 WEKA를 이용하여 Feature Selection 기법을 적용한다(5). WEKA는 뉴질랜드 Waikato University에서 개발한 데이터 마이닝 도구로, 데이터셋을 입력받아 처리하여 특정 속성과 가장 연관이 있고 유의미한 속성을 도출한다.

본 연구에서는 장애 발생과 연관 있는 속성이 무엇인지가 핵심이므로, 기존 속성이 result가 되고, 나머지 5개의 속성(year, month, date, unit, component)은 변수 속성이 된다. <그림 4>는 위에서 언급한 Feature Selection 알고리즘을 적용하여 result 속성과 가장 연관이 깊은 속성을 순서대로 도출한 것이고, <그림 5>부터 <그림 9>까지는 KNCCS를 대상으로, result 속성을 기준으로 각각의 변수 속성과의 연관관계를 도식화한 것이다. 그래프의 X축은 변수 속성을, Y축은 장애 여부를 나타내었고, 장애가 발생한 데이터는 적색으로 표시하였다.

```

=== Attribute Selection on all input data ===
Search Method:
  Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 6 result):
  Gain Ratio feature evaluator

Ranked attributes:
0.0008265 4 unit
0.000272 5 component
0.0001385 2 month
0.0000391 3 date
0 1 year

Selected attributes: 4,5,2,3,1 : 5
    
```

Fig. 4. Attribute Selection Output

<그림 4>에서는 result 속성과 가장 연관이 깊은 속성을 순서대로 도출한 결과이다. unit, component, month, date, year 순으로 연관 깊이가 깊었고, year 속성은 result 속성과는 연관성이 전혀 나타나지 않았다. 다시 말해, 체계의 장애와 가장 연관이 깊은 속성은 unit으로, 이는 특정 부대에만 체계 장애가 집중적으로 발생한다는 것을 의미한다. unit 속성의 연관도는 다른 속성들에 비해 압도적인 수치로 높았다.

<그림 5>는 result 속성과 year 속성 간의 관계 그래프이며, year 속성의 값마다 대체적으로 균등하게 분포되어 있는 것을 확인할 수 있다. 이는 특정 year 값에 result가 종속되지 않는다는 것을 의미하고, 연관성이 낮다고 정의할 수 있다.

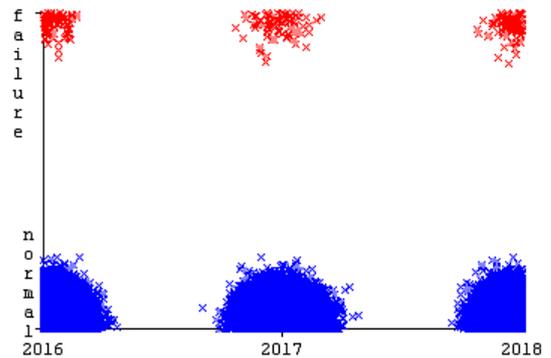


Fig. 5. Relation between result and year Attributes

<그림 6>은 result 속성과 month 속성 간의 관계 그래프이며, month의 특정 값에 result가 집중되어 있는 것을 볼 수 있다. month 값이 3, 4 혹은 8일 때에 result가 밀집되어 분포하고 있으며, KR/FE와 UFG 연습기간 등 지휘통제체계의 이용 빈도가 높아 그만큼 장애가 많이 발생하였다고 판단된다. 그리고 이러한 밀집도가 높을수록, 다시 말해 특정 속성 값에 종속되어 있을수록 두 속성간의 연관이 깊다고 할 수 있다.

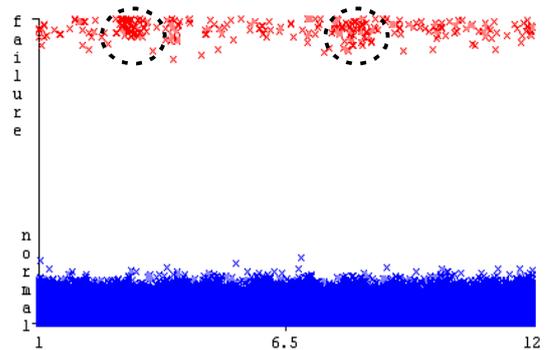


Fig. 6. Relation between result and month Attributes

<그림 7>은 result 속성과 date 속성 간의 관계 그래프이며, 앞서 언급한 year 속성과 마찬가지로 균등한 분포를 나타내고 있으므로 연관이 깊지 않다.

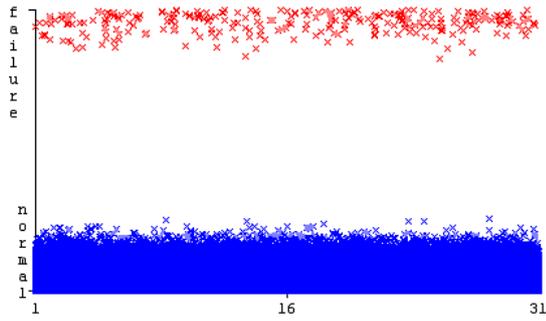


Fig. 7. Relation between result and date Attributes

<그림 8>은 result와 unit 속성 간의 관계 그래프이며, 사령부 본부(hq)가 압도적으로 많은 장애 현황을 나타내었고 이어서 00, 00함(CN, MS)이 장애가 많은 것으로 확인되었다. 이는 함정의 연식 및 체계 설치연도에도 관계가 깊고, 다시 말해 선령이 오래된 함정일수록 장애 발생 가능성이 높다는 것을 나타낸다.

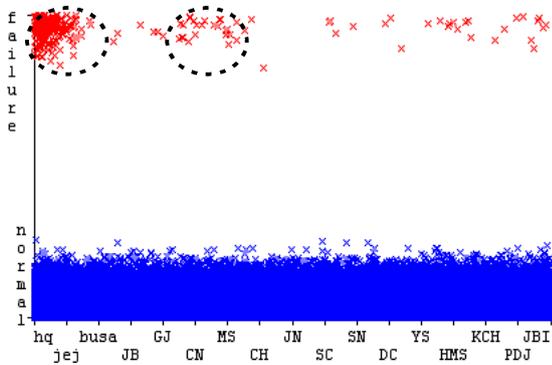


Fig. 8. Relation between result and unit Attributes

<그림 9>는 result와 component 속성 간의 관계이며, WAS서버(was), 단말장치(terminal)와 전시기(display)가 가장 빈번하게 장애가 발생하였다. 그래프 분석결과도 수치상과 동일하게 result 속성과 연관이 깊은 속성으로는 month, unit, component였고, 연관이 낮은 속성은 year, date였다.

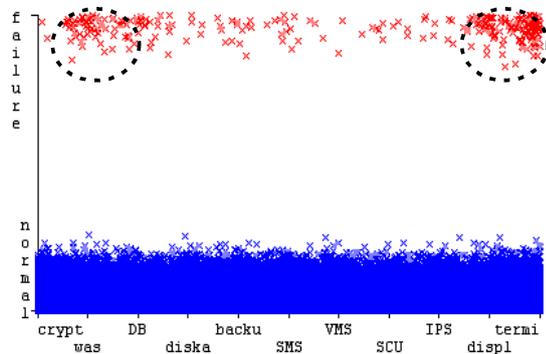


Fig. 9. Relation between result and component Attributes

KNTDS 또한 데이터셋을 활용하여 분석하였으며, KNCCS와 마찬가지로 year, date 속성은 result와의 연

관이 낮았고, month, unit, component 속성은 연관이 깊었다. unit 속성 측면에서 볼 때, KNTDS도 사령부 본부와 선령이 오래된 함정에 result 값이 집중적으로 분포하였고, component 측면에서는 사용자 콘솔인 TCC에 result 값이 밀집된 것을 확인하였다.

4. 결론 및 향후연구

본 연구에서는 지휘통제체계 장애 현황의 데이터셋을 활용하여 데이터 마이닝 도구를 통해 체계 장애와 가장 연관이 깊은 속성이 무엇인지를 도출하였다. 이를 실무에 적용하면, 장애가 발생했을 경우 장애 현황을 데이터셋에 추가하여 유지하고, 일정 주기(반기, 연간)마다 분석하여 장애를 예측할 수 있는 자료로 사용할 수 있다. 그리고 그 자료를 통해 유지보수 인력 및 자원이 집중적으로 배치되어야 할 시기와 위치를 산정할 수 있고, 나아가 유지보수 용역 계약 시에 가용 예산 산출에도 이용할 수 있다. 데이터셋에 사용되는 CSV 파일은 작성이 용이하기 때문에 데이터를 축적하는데 어려움은 없을 것이고, 따라서 실무에 적용하기에는 제한사항이 적을 것으로 판단된다. 또한 본 논문에서 제시한 속성 이외에도 체계 관리자에 따라 별도의 속성을 추가하여 다양한 분석이 가능하다.

향후에는 이번 실험에 적용했던 데이터셋의 속성 이외에 체계 장애와 관련이 있을 것으로 판단되는 다른 속성들을 추가적으로 연구 및 선별하여 분석하고, WEKA와는 다른 알고리즘을 사용하는 데이터 마이닝 도구들에도 데이터셋을 적용하여 유사도를 분석할 예정이다.

참고문헌

- 1) Navy Headquarters, "KNCCS Guidebook," 2017, p.3.
- 2) Navy Headquarters, "KNTDS Guidebook," 2011, pp.8~9.
- 3) KiWang Kim, Dongsu Kang, "C4I Maintenance Priority Decision using Maintenance Metric," *Journal of The Korea Society of Computer and Information*, Vol. 22, No. 12, 2017, pp. 145~153.
- 4) JiHee Nam, Dongsu Kang, "Analysis of ODT File Fuzzing Testcase in North Korea using Feature Selection Method," *The KIPS Spring Conference 2019*, Vol. 26, No. 1, 2019, pp.324~327.
- 5) Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H., "The WEKA Data Mining Software: An Update," *ACM SIGKDD Explorations newsletter*, Vol. 11, No. 1, 2009, pp.10~18.