

정박 중 준해양사고 원인에 대한 빅데이터 분석 연구

노범석* · 김태훈** · † 강석용

*,**,† 한국해양수산연수원 교수

요약 : 준해양사고를 줄이기 위하여 준해양사고 등을 분석하여 사고 예방에 활용하였다. 하지만 준해양사고 건수가 많은 대신 주내용이 정성적이기 때문에 다양한 정량적 데이터로 분석하기에는 현실적 어려움이 있었다. 이리 장단점을 고려하여 준해양사고에 대해서 그동안 단순한 내용 검토 방식에서 통계적 분석과 이를 통한 객관적 결과 도출이 가능한 빅데이터 기법을 적용한 연구가 필요하다. 이를 위해 10,000여 건의 준해양사고 보고서를 전처리 작업을 통해 통일된 양식으로 정리하였다. 이 데이터를 기반으로 1차로 텍스트마이닝 분석을 통해 정박 중 준해양사고 발생 원인에 대한 주요 키워드를 도출하였다. 주요 키워드에 대해 2차로 시계열 및 클러스터 분석을 통해 발생할 수 있는 준해양사고 상황에 대한 경향 예측을 도출하였다. 이번 연구에서는 정성적 자료인 준해양사고 보고서를 빅데이터 기법을 활용하여 정량화된 데이터로 전환할 수 있고 이를 통해 통계적 분석이 가능함을 확인하였다. 또한 빅데이터 기법을 통해 차 후 발생할 수 있는 준해양사고 객관적인 경향을 파악함으로써 예방 대책에 대한 정보 제공이 가능함을 확인할 수 있었다.

핵심용어 : 준해양사고, 통계 분석, 빅데이터, 텍스트마이닝, 시계열 및 클러스터 분석

Contents

- 1 연구배경
- 2 연구내용 및 방법
- 3 데이터 분석 절차
- 4 데이터 분석 내용
- 5 데이터 분석 결과 및 제언

2. 연구내용 및 방법

1. 연구배경

- 대형 해양사고의 경우, 사고가 발생하기 전 수많은 경미한 사고(준해양사고)와 징후들이 존재하고 이들의 반복되는 과정에서 발생
- 각 선사에서 단순히 제공받은 준해양사고 보고서를 유용한 정보로 활용할 수 있도록 빅데이터 구축 작업이 요구됨
- 이러한 빅데이터에서 유용한 정보를 도출할 수 있는 통계 및 분석 기법이 필요함

3. 데이터 분석 절차

1. 데이터 분석 과정

† Corresponding Author : 종신회원, sykang53@seaman.or.kr
* 정회원, bsro@seaman.or.kr

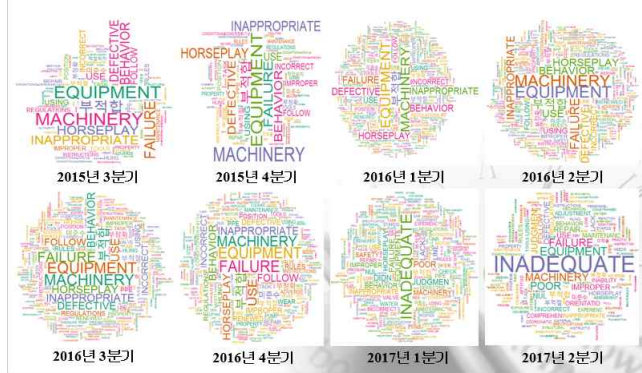
3. 데이터 분석 절차

2. 각 선사에서 보내준 준해양사고 보고서에 대한 데이터 전처리

- 보고서 내 정형적인 사항과 비정형적 사항을 정리하여 분석이 용이한 형태의 데이터로 변환

Ship's type	Date(local)	Place	Time	Details of event	Cause of accident	Prevention Measures
컨테이너선	2017-01-19	인양	1. 육상 작업 중 작업 중 2.인양 차 1. 인양 작업으로 인해 육상 작업 2. 육상작업 작업을 미흡하고 육상			
컨테이너선	2017-01-24	오일	1. OS WORKED OUT ON DECK	1. HE DIDNT WEAR PROPER	1. HOSE 연결 시 철저한	
컨테이너선	2017-02-10	인양	1. 2017-02-10일 중 인양하는 도중	1. 불규칙 고압에 의한 파손부식	1. 새로운 PPE 제작	
컨테이너선	2017-02-23	인양	1. 부양장치 검사 중 인양하는 도중	1. 승선로 측방을 침범하여 파손	1. 추후 인양 배후에의 안전장치	
양치선	2017-03-14	인양	1. YEORU에서 GONGGAD로 항해	1. 본선의 인양용이 운용 후 운용기 1. 본선의 인양용용량에 운용기구를		
양치선	2017-03-24	인양	1. 2017-03-24 항해	1. 인양 중 인양용 NO2 GLE 1. 발진기 운전 중 고압 진동에 의해 1. 인양 중 PATROL 장치의 시정		
양치선	2017-03-27	인양	1. 항해 중 KAKOR INSPECTION을	1. 배 중 전동기 배터리에 인양함 1. 기어용 정밀검사를 실시함		
양치선	2017-04-08	인양	1. 육상 HYDROTEST를 위해	1. 작업이 없는 다른 작업의 사용 1. 작업 시 올바른 안전장비 및 도		
양치선	2017-04-28	인양	1. LUG는 높은 직경의 경우 위해	1. 고압을 견디지 못하여 파손 1. 작업 시 올바른 수단은 물체에		
양치선	2017-04-29	인양	1. LUMA승용 중 NO2 GLE FID	1. LO-RING 손상	1. 인양용 LUMA 승용 철저한 시정	
양치선	2017-04-29	인양	1. DAESAN항 집안용 배기 펌프기 1. MUSEAL 의 불완전으로 인한 누출	1. 인양 중 승용 철저한 시정		
양치선	2017-01-02	인양	1. NO2 MISSO UNIT CLR LIQUID CHUALL		1. LIQUID CHANGING VV CROSS	
양치선	2017-01-20	인양	1. 4/1 DOME SAMPLING VALVE	1. 작업 중 작업자 부주의로 인해	1. 작업담당자에게 계동 SAMPLING	
양치선	2017-01-20	인양	1. NO2 BOILER WATER LEVEL	1. 위험한 LEVEL GAUGE 부품 불량으로	1. 승선 시 위험한 LEVEL GAUC	
양치선	2017-01-21	인양	1. 인양 1# MOORING LINE 인양	1. BRAKE TIGHT 하지 않은 상태에서 1. 계동 작업 중 시 시	1. 승선시 안전 도	
양치선	2017-01-23	인양	1. DECK FRESH WATER LINE 누출	1. Gasket 불량	1. 통풍장치 중시 시	
양치선	2017-01-24	인양	1. NO2 DVG LO. LUMP TANK	1. LO LUMP TANK LEVEL SWITCH	1. 인양 기어에 대한 SPARE 보유	
양치선	2017-01-23	인양	1. WINDY WINDLASS REMOTE CON Tank	1. 인양 Drain Hole 손상	1. 인양용 인양용 인양용 인양용	
양치선	2017-01-24	인양	1. ELEVATOR 4TH DK LANDING	1. 각 DK 정장 클러 4TH DK	1. ELEVATOR 관련 SAFETY INTER	
양치선	2017-01-30	인양	1. SLECC-A CHEMICAL 주입TANK	1. SLECC-A CHEMICAL 주입TANK	1. 인양 인양 시 TANK LEVEL 및 누	
양치선	2017-01-04	인양	1. NO2 MISSO UNIT COOLER 불량	1. NO2 MISSO UNIT COOLER 불량	1. 인양 인양 시 인양용 인양용	
양치선	2017-01-09	인양	1. LIFTING DEVICE 인양 중 WIRE SWAP	1. 인양 인양 불	1. LIFTING DEVICE 인양 불	
양치선	2017-01-09	인양	1. MONITORIAL LASHING 작업 중	1. 작업 중 인양용 인양용 인양용	1. 인양 인양 불	
양치선	2017-01-16	인양	1. NO2 FWG BBRNE P.P	1. 4. 작업중, BBRNE PUMP	1. 인양 인양 불	
양치선	2017-01-17	인양	1. CRANE 인양 중 인양용 인양	1. BOLT 인양 도중 SPANNER가	1. 인양 인양 불	
양치선	2017-01-23	인양	1. 인양 인양 시 인양용 인양용 인양용	1. 인양 인양 불	1. 인양 인양 불	
양치선	2017-01-02	인양	1. NO2 LD LUMP TANK	1. GLOBE VALVE BODY 파손	1. 인양 인양 불	
양치선	2017-01-12	인양	1. NO2 FWG TEMPERATURE	1. TEMPERATURE SENSITIVITY	1. 인양 인양 불	

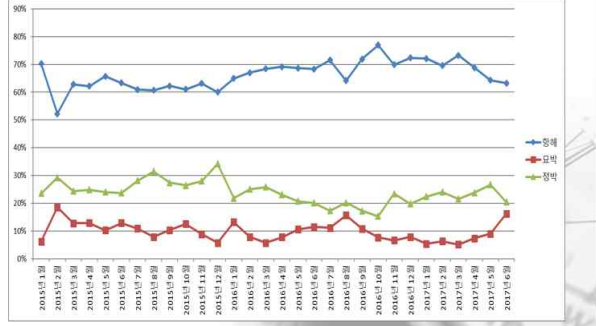
4. 데이터 분석 내용



4. 데이터 분석 내용

1. 전 처리된 데이터 자료에 대한 기초 분석 결과

(1) 항해 유형별 준해양사고의 년도별/월별 발생률 그래프



4. 데이터 분석 내용

2. 데이터 2차 분석 결과

(1) 정박 중 준해양사고 COA에 대한 분기별 텍스트마이닝 과정에서 도출되는 키워드 중에서 제일 많은 빈도수를 나타내는 10개 단어를 선정하여 동계 분석 실시

	15-01	15-02	15-03	15-04	16-01	16-02	16-03	16-04	17-01	17-02
EQUIPMENT	25%	24%	23%	26%	24%	28%	24%	23%	37%	36%
BEHAVIOR	17%	18%	17%	18%	15%	18%	16%	16%	32%	13%
HORSEPLAY	17%	18%	17%	18%	15%	18%	16%	16%	31%	13%
DEFECTIVE	18%	17%	16%	17%	16%	17%	16%	11%	0%	24%
RULES	6%	6%	7%	5%	6%	6%	8%	9%	0%	0%
PPE	5%	6%	5%	6%	5%	7%	7%	8%	0%	0%
POSITION	5%	5%	5%	6%	6%	0%	6%	8%	0%	0%
TOOLS	8%	6%	6%	5%	7%	0%	0%	0%	0%	0%
RENEWED	0%	0%	0%	0%	6%	6%	6%	8%	0%	0%
MAINTENANCE	0%	0%	5%	0%	0%	0%	0%	0%	0%	15%

4. 데이터 분석 내용

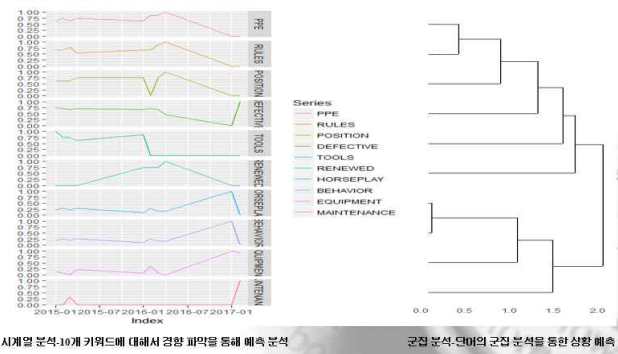
2. 데이터 1차 분석 결과

- 정박 중 준해양사고 사고 원인(Cause of Accident:COA)에 대한 분기별 텍스트마이닝 분석 결과



4. 데이터 분석 내용

(2) 10개 키워드에 대해 시계열 분석과 군집(클러스터링) 분석



5. 데이터 분석 결과 및 제언

1. 데이터 1차 분석 결과

	15:01	15:42	15:43	15:54	16:01	16:40	16:48	16:54	17:01	17:40
EQUIPMENT	25%	24%	23%	26%	24%	24%	24%	23%	27%	30%
BEHAVIOR	17%	16%	17%	18%	15%	18%	16%	16%	22%	13%
HORSEPLAY	17%	16%	17%	18%	15%	18%	16%	16%	20%	13%
DEFECTIVE	18%	17%	16%	17%	16%	17%	16%	11%	8%	24%
RULES	6%	6%	7%	5%	6%	6%	8%	9%	6%	6%
PPE	5%	6%	5%	6%	5%	7%	7%	8%	6%	6%
POSITION	5%	5%	5%	6%	6%	6%	6%	6%	6%	6%
TOOLS	8%	6%	6%	5%	7%	6%	6%	6%	6%	6%
RENEWED	2%	2%	2%	2%	2%	2%	2%	2%	2%	2%
MAINTENANCE	2%	2%	5%	2%	2%	2%	2%	2%	2%	15%

- (1) 데이터마이닝 분석에 의해 옆의 그림과 같이 빈도에 따라 키워드의 크기가 다르게 보이게 됨
- (2) 분석되는 키워드를 빈도수에 따라 정리한 후 의미 없는 단어는 삭제하고 상위 10개의 키워드에 대해서 통계 분석 실시
- 1) BEHAVIOR: 선박 운항 중 선원의 활동에 의한 것으로 해석 가능
- 2) DEFECTIVE: 선박 장비의 결함에 의한 것으로 해석 가능
- 3) EQUIPMENT: 선박 장비의 의한 것으로 해석 가능
- 4) HORSEPLAY: 선박 운항 중 선원의 부주의(집중력 결여)에 의한 것으로 해석 가능
- 5) MAINTENANCE: 선박 장비의 유보수에 의한 것으로 해석 가능
- 6) POSITION: 입출항 시 위치, 정비 시 위치 등에서 본인 임무의 소홀에 의한 것으로 해석 가능
- 7) PPE: 개인 보호 장구에 의해 발생된 것으로 해석
- 8) RENEWED: 선박 장비에 대해 부속품이나 선용품, 유류 교체 과정에서 발생된 것으로 해석 가능
- 9) RULES: 회사 규정, 국제 or 국내 법규 미준수에 의해 발생된 것으로 해석 가능
- 10) TOOLS: 공구, 도구, 장비 등을 사용하는 과정에서 발생된 것으로 해석 가능

5. 데이터 분석 결과 및 제언

(2) 군집(클러스터링) 분석을 덴드로그램으로 나타내면 키워드간의 상관관계와 밀접도 확인이 가능하며, 추가적으로 전문가 집단의 분석을 통해 높은 상관관계와 밀접도를 보이는 키워드를 가지고 발생할 수 있는 상황에 대해서 예측이 가능함

- 1) 군집 분석 프로그램의 특성상 각 키워드에 대해 골짜기 병합하는 과정이 발생하므로 여기에서는 2-3 단계까지의 상관관계와 밀접도에 대해서만 평가함
- 2) 군집 분석 그래프에서 보면 RULLS라는 키워드가 나오면 PPE 키워드가 많이 나오고 반대의 경우도 그러하므로 키워드간의 관계가 형성되는 것을 알 수 있으며, 이를 통해 예측 분석을 해보면 PPE 장비 등이 규정과 맞지 않아 준해양사고가 발생할 수 있음을 예측할 수 있음
- 3) 군집 분석 그래프에서 보면 1차적으로 RULLS와 PPE 키워드의 관계가 형성되고 2차적으로 POSITION 키워드도 관계가 형성되는 것을 알 수 있는데, 이것을 1, 2단계로 나누어 예측 분석을 해보면 1단계에서 PPE 장비의 규정 미준수에 의한 준해양사고를 예측이 가능하고 2단계에서는 PPE 장비의 규정 미준수 상황을 처리하는 과정에서 부주의한 위치 선정으로 준해양사고 발생을 예측할 수 있음

5. 데이터 분석 결과 및 제언

2. 데이터 2차 분석 결과

(1) 시계열 분석에 의한 트렌드 그래프를 보면 향후 예측이 가능함

- 1) EQUIPMENT 키워드의 경우 2015년에 낮은 분포율을 보이다가 2016년에는 증가세를 보였음. 2017년에 들어 약간의 감소세를 돌아섰음을 확인할 수 있음. 2016년에 어떤 대책이 실시되었는지 확인해볼 필요가 있음
- 2) POSITION, PPE, RULES 키워드의 경우 2015년에 비슷한 경향을 보이다가 2016년 초반 들어 증가세를 나타내고 중반기부터 감소되는 추세를 보이고 있으므로 향후 예측이 가능할 수 있음
- 3) TOOLS 키워드의 경우 2016년 초반까지 미세하게 증가세를 보이다 이후 급감하면서 2016년 중반 이후부터는 거의 나타나지 않으므로 향후 비슷한 경향을 보일 수 있다고 예측 가능함
- 4) RENEWED 키워드의 경우 2015년부터 2016년 중반기까지 상승했다가 그 이후 감소되는 경향을 보이므로 2016년 중반까지 어떤 사유에 의해 증가했는지 확인해보면 예방 방법으로 활용 가능할 수 있음
- 5) BEHAVIOR와 HOESEPLAY, DEFECTIVE와 MAINTENANCE 키워드의 경우 분석 기간동안 비슷한 증가 및 감소 분포율을 보이고 있으므로 서로간의 상관관계 분석 필요

5. 데이터 분석 결과 및 제언

3. 제언

- (1) 추가 연구를 통해 항해 유형별, 사고 유형별, 선종별 등으로 나누어 사고 원인, 사고 개요에 대해 세분화하여 빅데이터 기법으로 분석한다면 해양사고 예방에 필요한 정량적이고 구체적인 자료 추출이 가능할 것으로 판단됨
- (2) 추가적으로 중해심의 매년 발표하는 해양사고 통계자료 분석을 통해 준해양사고의 정성적인 자료와 해양사고의 정량적인 자료 사이의 상관 관계를 파악하는 연구가 필요할 것으로 보이며 구체적인 상관 관계와 연관성을 확인할 수 있다면 해양사고 예방에 도움이 될 것으로 판단됨