

공공데이터 기반 화재 발생 요인 분석

김다현^o, 백재순^{*}, 김성진(교신저자)^{*}

^o명지전문대학교 ICT융합공학과,

^{*}명지전문대학교 ICT융합공학과

e-mail: dh6613@mjc.ac.kr^o, {hisoon99, ict214548}@mjc.ac.kr^{*}

Analysis of Fire Occurrence Factors based on Public Data

Da-hyun Kim^o, Jai-Soon Baek^{*}, Sung-Jin Kim(Corresponding Author)^{*}

^oDept. of ICT Convergence Engineering, Myongji College,

^{*}Dept. of ICT Convergence Engineering, Myongji College

● 요약 ●

최근 통계자료에 따르면, 매년 화재 발생 건수와 이로 인한 피해가 증가하는 것으로 나타나고 있다. 이에 따라 본 연구에서는 여러 공공 데이터를 분석하여 각 지역별 특성이나 환경적 요인에 따른 화재의 원인을 분석해고보가 한다. 분석을 위하여 화재 발생 건수, 시도별 공장 수와 인구수, 계절별 화재 발생수, 장소별 화재 요인에 대한 각각의 공공 데이터를 활용하였다. 분석 결과 지역별 공장 수와 비례하지않은 화재 발생 건수나, 대중적으로 생각하는 겨울철 가장 많은 화재의 발생보다 봄철에 더 많은 화재가 발생되었던 통계 결과와 같이 각 지역적인 특성이나 발화요인에 따라 화재 발생률에 미치는 영향은 다르게 나타난다는 것을 알 수 있었다.

키워드: 빅데이터(BigData), 분석(Analysis), 화재(Fire), 발화요인(Ignition Cause), 시각화(Visualization)

I. Introduction

2023년 04월 24일 행정안전부의 통계에 따르면 최근 10년간 연평균 41,257개의 화재 건수와 2,286명의 인명피해가 발생되었다. 화재 발생 건수는 전년도인 2022년부터 점점 증가하는 추세로 나타난다[1]. 또한, 올해 1월-6월 경기북부에서 발생한 화재는 총 1387건으로 지난해 같은 기간 1292건보다 7.4% 증가한 것으로 조사 되었고, 화재로 인한 인명피해 규모는 지난해 95명에서 올해 101명으로 6.3% 소폭 늘었다고 한다[2].

본 논문에서는 공공데이터 포털, 소방청에서의 관련 자료를 활용하며, R studio를 사용한 자료 분석을 했다. 각 지역별 특성이나 환경적 요인에 따른 원인을 파악하고자 화재발생 건수, 시도별 공장 수와 인구 수, 계절별 화재 발생수, 장소별 화재 요인에 대한 각각의 공공 데이터를 활용했으며 데이터 시각화를 통한 분석 결과를 제시함으로써 예측 불가한 재난사고에대한 대응방안과 경각심을 주고자 한다.

II. Preliminaries

서민승(2021)의 연구에서는 머신러닝을 이용하여 경기도의 화재위험요인에 대한 예측 분석하였다. 화재 발생에 영향을 주는 요인들의 위험순서를 파악하여 제시하고 있다[3].

류정우(2015)의 연구에서는 데이터마이닝을 이용하여 기상정보에 따른 화재 위험을 평가하였고, 이를 통해 기상정보에 따른 화재위험평기등급과 특정 기상조건에서 화재 요인에 따른 화재위험도를 제시하고 있다[4].

노주형(2022)의 연구에서는 합성곱 신경망 기반 분류 모델을 통해 화재 예측 성능을 분석하고 있다.

이 3가지 연구들은 모두 특정 기술을 기반으로 화재 예측을 진행하고 이에 따른 성능을 살펴보고 있는 연구들이다. 본 연구에서는 공공 데이터를 기반으로 화재 관련 데이터를 지역별 특성이나 환경적 요인에 기반하여 화재와의 연관성을 살펴보고자 한다. 이를 위하여 본 연구에서는 3가지 가설을 수립하여 분석을 실시하였고, 3가지 가설은 아래와 같다.

첫째, 경기도지역의 비교적 많은 제조업체나 공장의 분포가 높은 화재 발생율의 원인이 된다.

둘째, 겨울철이 다른 계절에 비해 화재 발생률이 더 높을 것이다 셋째, 인구 밀도가 높은 지역에서의 화재 발생률은 더 높다

III. The Proposed Scheme

1. 공장 분포에 따른 화재 발생률

첫 번째 가설은 ‘경기도지역의 비교적 많은 제조업체나 공장의 분포가 높은 화재 발생률의 원인이 된다’ 이며 가설의 성립 여부를 파악하고자 시도별 공장 수와 시도별 인구수를 대조하여 지역적 특성이 가설에 영향을 미치는지 살펴보았다.

그 결과 시도별 공장 소유 현황의 데이터 분석을 통해서 경기도 > 경상남도 > 경상북도 순으로 많이 분포되어있음을 알 수 있었고, 시도별 화재 발생 현황 데이터 분석을 통해서 경기도 > 서울특별시 > 경상남도 순으로 화재가 많이 발생한다는 것을 알 수 있었다.

이 두 가지의 데이터 비교를 통해 경기도 지역에서는 공장 수와 화재 발생 수가 비례했음을 알 수 있었지만, 그 외에 다른 지역에서는 비례하지 않았기에, 일부의 가설만 성립한다는 결과가 도출되었다.

이 외에 지역별 특성보다 공장과 같은 산업시설들이 화재 발생률에 영향을 주는지 알아보기 위해 화재 장소에 따른 화재 발생률 분석을 해보았다. 그 결과, 경기도 지역은 산업시설에 대한 발화요인으로의 화재 발생률이 높지만 다른 지역에서는 부주의(주거, 산불)로 인한 화재 발생률이 높았기에 결과적으로 산업시설에 대한 특효요인이 아닌 지역적인 특성의 영향이 있음을 알 수 있었다.

위의 결과를 종합적으로 살펴보면, 경기도 지역에서는 제조업체나 공장의 분포가 화재 발생률과 연관이 있을 가능성이 크다는 가설이 성립한 것으로 나타났다. 따라서 향후 연구나 정책 수립 시에는 지역 특성을 고려하여 접근해야할 것으로 판단된다. 특히 경기도와 같이 공장이 많이 분포된 지역에서는 산업 안전에 대한 강화된 대책이 필요하며, 그 외의 지역에서는 부주의로 인한 화재 예방을 강조하는 정책이 효과적일 수 있다.

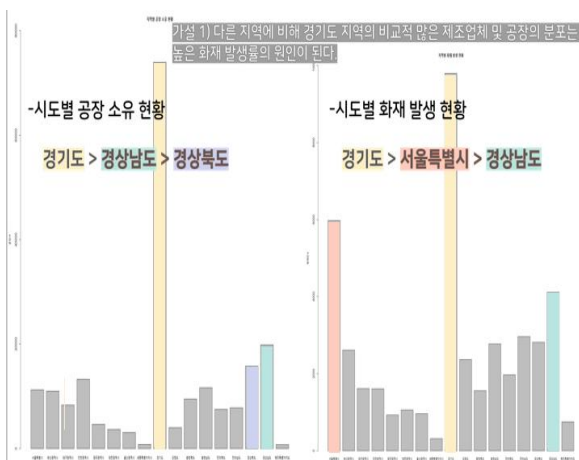


Fig. 1. Factory ownership status by city/province - data analysis on fire occurrence status by city/province

가설 1) 다른 지역에 비해 경기도 지역의 비교적 많은 제조업체 및 공장의 분포는 높은 화재 발생률의 원인이 된다.

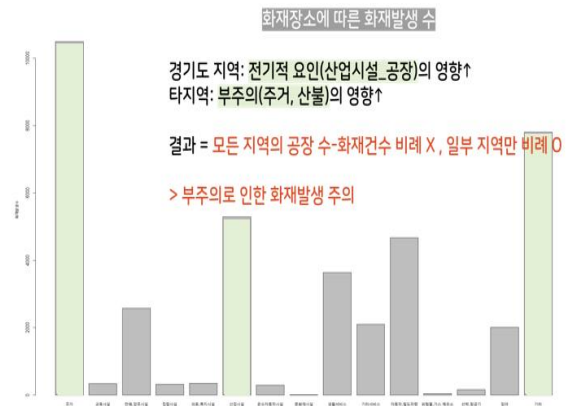


Fig. 2. Data analysis based on factors for each fire occurrence location



Fig. 3. Data analysis according to fire ignition factors

2. 계절별 화재 발생률

두 번째 가설은 ‘겨울철이 다른 계절에 비해 화재 발생률이 더 높을 것이다’ 이며, 가설의 성립 여부를 파악하고자 분기별로 데이터 전처리하여 나타낸 계절별 화재 요인과 화재 발생 건수의 두 가지 데이터로 분석해보았다. 최근 4년간 계절별로 나타난 화재 건수 데이터를 통해서 봄(3-5월) > 겨울(12-2월) > 가을(9-11월) > 여름(6-8월) 순으로 화재 발생률이 높았다. 이를 통해 결과적으로 겨울철보다는 봄철에 더 많은 화재가 발생하기에 이 가설은 성립하지 않는다는 것을 알 수 있었고, 겨울철보다는 봄철에 발생하는 부주의로 인한 화재를 더 주의해야한다는 것을 알 수 있었다.

계절 (character)	2022 년 (double)	2021 년 (double)	2020 년 (double)
봄(3-5월)	377	158	381
여름(6-8월)	29	16	39
가을(9-11월)	82	25	83
겨울(12,1-2월)	268	150	117

Fig. 4. Data preprocessing process for the number of fires by season

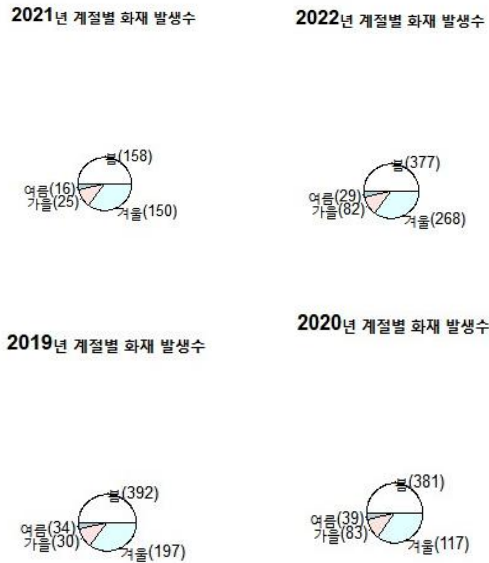


Fig. 5. Data analysis on number of fires by season

3. 인구밀도에 따른 화재 발생률

세 번째 가설은 ‘인구 밀도가 높은 지역에서의 화재 발생률은 더 높다’이며 가설의 성립 여부를 파악하고자 시도별 인구 수와 시도별 화재 발생 현황의 두 가지 데이터로 분석해보았다. 데이터 분석을 통해서 인구 수가 높은만큼 화재 발생률이 높다는 결과가 나왔으며 인구 수는 화재 발생 수와 비례하므로 이 가설은 성립한다는 것을 알 수 있었다. 인구 수가 많은 지역일수록 발생하는 화재도 많지만 지역 특성 상 더 발달되어있는 소방, 응급 시설을 통해 좀 더 빠른 조치가 이루어질 수 있다고 한다.

시도별	전체 인구 수
1 서울특별시	9417469
2 부산광역시	3295760
3 대구광역시	2366852
4 인천광역시	2989125
5 광주광역시	1468972
6 대전광역시	1473662
7 울산광역시	1110516
8 세종특별자치시	382589
9 경기도	13717827
10 강원도	1528037
11 충청북도	1624993
12 충청남도	2193214
13 전라북도	1774248
14 전라남도	1771431
15 경상북도	2620373
16 경상남도	3280829
17 제주특별자치도	676375

Fig. 6. pre-processing process for the total population by city and province

가설 3) 인구 밀도가 높은 지역에서의 화재 발생률은 더 높다.

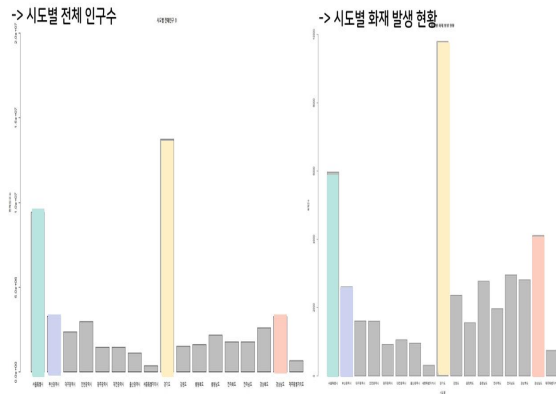


Fig. 7. Total population by city and city and fire occurrence rate by city data analysis

IV. Conclusions

지금까지 화재발생률에 영향을 미치는 특정요인에 대해 살펴봤다. 지역별 공장 수와 비례하지 않은 화재 발생 건수나, 대중적으로 생각하는 겨울철 가장 많은 화재의 발생보다 봄철에 더 많은 화재가 발생되었던 통계 결과와 같이 각 지역적인 특성이나 발화요인에 따라 화재 발생률에 미치는 영향은 다르게 나타난다는 것을 알 수 있었다. 따라서 어떤 특정 요인보다 부주의(주거, 산불, 담뱃불 등)로 인한 화재 발생을 가장 조심하며 언제 어디서나 경각심을 갖고 상시에 예방할 수 있어야 한다.

REFERENCES

- [1] Ministry of Public Administration and Security. Fire analysis in the last 10 years, average 41,257 cases per year, 2,286 casualties
- [2] Gyeonggi-do News Portal - 8,169 fires in Gyeonggi-do last year...8.4% Decline Compared to 2020
- [3] SSeo, M. S., Castillo Osorio, E. E., & Yoo, H. H. (2021). Predictive Analysis of Fire Risk Factors in Gyeonggi-do Using Machine Learning. *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, 39(6), 351-361.
- [4] Song, D. W. (2014). Prediction of the Risk of Fire Occurrence According to the Weather Information Using Statistics and Data Mining Techniques (Doctoral dissertation, Ph. D Thesis, Seoul National University of Science and Technology).
- [5] Joo-Hyung Roh, Se-Hong Min, & Min-Suk Kong. (2022). Analysis of Fire Prediction Performance of Image Classification Models based on Convolutional Neural Network. *Fire Science and Engineering*, 36(6), 70-77.