

계층적 군집분석 기반의 Continuous Risk Profile을 이용한 고속도로 사고취약구간 선정

이서영¹ · 김철순² · 김동규³ · 이청원^{2*}

¹ LG CNS 스마트교통사업부, ² 서울대학교 건설환경공학부, ³ 서울대학교 건설환경종합연구소

Identifying Hotspots on Freeways Using the Continuous Risk Profile With Hierarchical Clustering Analysis

LEE, Seoyoung¹ · KIM, Cheolsun² · KIM, Dong-Kyu³ · LEE, Chungwon^{2*}

¹ Smart Transportation Business Division, LG CNS, Seoul 100-741, Korea

² Department of Civil and Environmental Engineering, Seoul National University, Seoul 151-744, Korea

³ Integrated Research Institute of Construction and Environmental Engineering, Seoul National University, Seoul 151-744, Korea

Abstract

The Continuous Risk Profile (CRP) has been well known to be the most accurate and efficient among existing network screening methods. However, the classical CRP uses safety performance functions (SPFs) which require a huge investment to construct a database system. This study aims to suggest a new CRP method using average crash frequencies of homogeneous groups, instead of SPFs, as rescaling factors. Hierarchical clustering analysis is performed to classify freeway segments into homogeneous groups based on the data of AADT and number of lanes. Using the data from I-880 in California, the proposed method is compared to other several network screening methods. The results show that the proposed method decrease false positive rates while it does not produce any false negatives. The method developed in this study can be easily applied to screen freeway networks without any additional complex database systems, and contribute to the improvement of freeway safety management systems.

Continuous Risk Profile(CRP)은 고속도로의 사고취약구간을 선정하는 방법론 중에서 정확성과 효율성이 뛰어난 것으로 알려져 있다. 그러나 전통적인 CRP는 데이터베이스 구축을 위한 대규모 투자를 필요로 하는 안전성능함수를 이용한다. 본 연구는 안전성능함수 대신 동질 그룹들의 평균사고건수를 규모조정계수로 이용하는 CRP를 제안하는 것을 목적으로 한다. 고속도로 구간들을 동질 그룹으로 분류하기 위하여 각 구간의 AADT와 차로 수 자료를 기반으로 하는 계층적 군집분석이 수행된다. 제안된 모형은 캘리포니아의 I-880 자료를 이용하여 다른 여러 가지 사고취약 구간 선정방법들과 비교된다. 분석 결과에 따르면, 제안된 모형은 false negative를 발생시키지 않으며 false positive rate를 감소시킨다. 본 연구에서 개발된 방법론은 추가적인 복잡한 데이터베이스 없이 고속도로 사고취약구간을 선정하는 데에 활용될 수 있으며, 또한 고속도로 안전관리시스템을 개선하는 데에 기여할 수 있다.

Key Words

Average Crash Frequency, Continuous Risk Profile, Hierarchical Clustering Analysis, Hotspot, Safety Performance Function
평균사고건수, Continuous Risk Profile, 계층적 군집분석, 사고취약구간, 안전성능함수

* : Corresponding Author
chungwon@snu.ac.kr, Phone: +82-2-880-7368, Fax: +82-2-873-2684

Received 13 April 2013, Accepted 17 July 2013

I. 서론

고속도로에서 발생하는 대부분의 사고는 일반도로에서 발생하는 사고와 비교했을 때, 상대적으로 사고의 규모가 크며 심각한 인명피해를 유발할 가능성이 크다. 따라서 고속도로 구간 중 사고에 취약한 구간을 선정하여 그에 따른 개선책을 수립함으로써, 고속도로의 안전성을 향상시킬 필요가 있다. 또한 사고취약구간(hotspot)을 선정하는 과정에서 효과적인 방법론을 적용한다면, 도로의 실제 위험구간(true hotspot)을 탐색하기 위해 소요되는 전문가의 시간 및 노동력을 절약할 수 있다. 이것은 고속도로 안전 개선사업의 효율적 예산분배 문제와 직결되어 있다.

현재 널리 알려진 사고취약구간 선정 방법론은 크게 세 가지로, Sliding Moving Window (SMW), Peak Searching (PS), Continuous Risk Profile (CRP)이다. 이 중 2007년에 개발된 CRP(Chung and Ragland, 2007)는 다른 사고취약구간 선정 방법론에 비해 사고취약구간의 연도별 변화와 흐름을 면밀하게 검토하며, 사고취약구간으로 선정되었음에도 실제로는 도로상의 특별한 결함이 존재하지 않는 지점의 비율(false positive rate)이 작다(Kwon et al., 2012). 또한, 교통량에 따른 사고율의 변화를 반영하지 못하는 SMW와 PS와 달리 규모조정계수(rescaling factor)로 이를 반영하게 된다. 이와 같은 장점으로 인해 CRP가 사고취약구간을 효과적으로 선정할 수 있으나, CRP를 사용하기 위해서는 분석과정 중 규모조정계수로 사용되는 안전성능함수(SPF: safety performance function)가 필요하다. 안전성능함수를 구축하는 과정 자체가 간단치 않으며, AADT, 차로 수 등의 기본적인 자료뿐만 아니라 도로운영과 관련된 기하구조 DB 등의 방대한 고속도로 자료가 필요하다(Alluri and Ogle, 2012). 이것은 고속도로 DB가 제대로 구축되어 있지 않은 국가나 지역에서는 CRP의 활용이 제한적이라는 것을 의미한다.

따라서 본 연구에서는 안전성능함수의 구축이 어려운 국가나 지역에서도 활용할 수 있는 CRP 방법론을 제시하고자 한다. 기존 방법론에서 활용되는 안전성능함수 구축을 위해서는 고속도로 DB와 숙련된 연구진이 필요하다. 이에 비해 교통량(AADT)과 차로 수 정보를 활용하는 새로운 방법론은 분석이 상대적으로 간단하고, 이해가 쉬우므로 실무활용성이 높다는 장점이 있다.

본 연구의 II장에서는 사고취약구간을 선정하는 방법

론에 대한 이해를 위해 기존문헌 고찰을 수행하였다. III장에서는 분석에 필요한 자료를 정리하고, 기존의 CRP와 본 연구에서 제시하는 새로운 방법의 CRP 분석을 동시에 수행하였다. 가장 기본적인 고속도로 자료인 AADT와 차로 수를 이용하여 계층적 군집분석(hierarchical clustering analysis)을 시행하였으며, 그 결과로 산출된 그룹별 평균사고건수를 기존 CRP 분석에서 활용되는 안전성능함수 대신 사용하였다. IV장에서는 분석 결과를 통해 새로운 방법론의 적용 가능성을 검토하였다. 먼저 기존의 CRP와 본 연구에서 제시하는 CRP 분석을 통해 선정된 사고취약구간을 비교하였다. 최종적으로 다른 사고취약구간 선정 방법론(SMW, PS)과의 효율성 분석을 시행하였다. 이를 통하여 군집분석을 이용한 CRP도 기존의 CRP처럼 다른 사고취약구간 선정 방법론에 비해 효과적으로 사고취약구간을 선정함을 검증하였다. V장에서는 결론 및 향후 연구과제를 기술하였다.

II. 사고취약구간 선정 방법론 및 기존문헌 고찰

1. 사고취약구간 선정 방법론

1) Sliding Moving Window

SMW는 고정된 단위길이의 Window가 도로구간을 따라 정해진 길이만큼 이동하면서 각 Window의 평가지표 값을 산출하는 방법이다. 해당 도로구간 중 사고감소 잠재성이 가장 큰 Window의 평가지표 값이 해당 도로구간의 대푯값이 되어 다른 도로구간들과 비교정렬 되는 과정을 통해 사고취약구간을 선정한다. 이 방법의 특징은 분석대상이 되는 구간을 서브구간으로 나누는 동질구간 분할(segmentation) 과정이 필요하다는 것과 평가지표 값을 산출하는 기준이 되는 Window가 다음 Window와 중첩된다는 점이다.

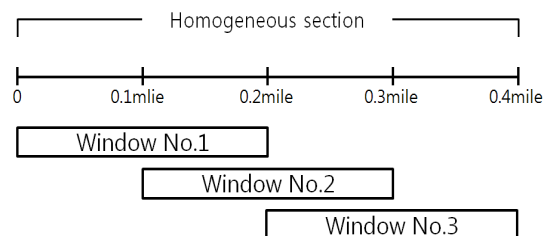


Figure 1. Concept of sliding moving window

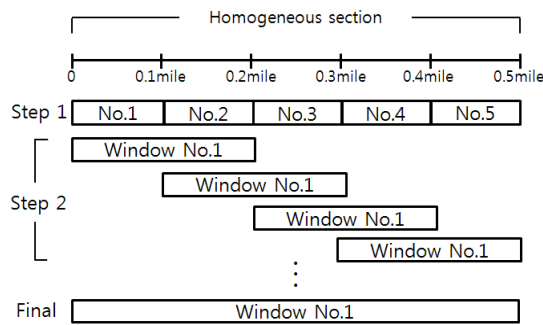


Figure 2. Concept of peak searching

2) Peak Searching

PS는 개별 도로구간을 일정 크기의 Window로 분할하고, 각 Window에 대한 평가지표 값을 계산하는 방법이다. 계산된 결과를 이용하여 정확도 테스트과정을 거치고, 만약 테스트 결과가 만족되지 않은 경우 Window의 크기를 점차적으로 증가시키면서 동일한 작업을 수행한다. 테스트 결과를 만족하는 평가지표의 최댓값을 도출하는 과정을 통해 최종적으로 사고취약구간을 선정한다. SMW와 마찬가지로 분석대상이 되는 구간을 서브구간으로 나누는 동질구간 분할 과정이 필요하다.

3) CRP

CRP는 도로를 따라 각 지점 혹은 구간에 내재된 위험도를 기준구간마다의 단위거리 당 사고밀도 값으로 산출하여 그래프로 도시하는 방법이다. 일정한 크기의 Window를 도로를 따라 조금씩 이동해가면서 Window 내부의 측정값을 구한 뒤 그래프를 도시한다. 다른 방법론과 다르게 동질구간 분할이 필요하지 않으며, 사고취약구간 전후의 상관관계를 고려할 수 있다는 점이 가장 큰 장점이다. 최고점의 위치와 형태를 통해 사고취약구간의 식별이 가능하며, 동일 지역에 대한 연도별 사고 위험도 변화 관측이 가능하다. 분석과정 중 규모조정계수로 사용되는 안전성능함수를 별도로 구축해야 한다.

SMW와 PS는 사고건수가 많지만 동시에 교통량도 많은 경우에 대해서는 고려하지 못한다는 단점이 있다. 즉, 사고의 원인이 교통량 증가에 의한 자연스러운 현상인지 아니면 특별한 원인이 있는 것인지 밝혀내기가 어렵다. 또한, 동질구간 분할 과정이 결과에 큰 영향을 미치게 되는데 사고취약구간이 동질구간의 경계에 위치하게 되면 실제 사고취약구간임에도 불구하고 사고 잦은

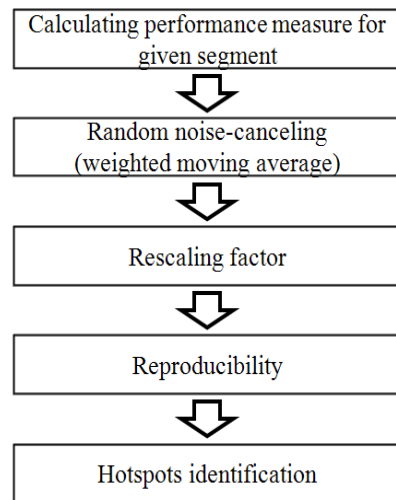


Figure 3. Analysis process of CRP

지점으로 검지되지 않는 구간(false negatives)이 발생할 가능성이 있다. CRP는 규모조정계수를 적용하는 과정에서 교통량을 반영하고, 연속된 그래프의 형태로써 동질구간 경계에서 발생할 수 있는 문제를 사전에 차단할 수 있다.

CRP 분석과정은 Figure 3과 같다. 먼저 주어진 구간에 대한 평가지표를 산출하여 가중이동평균법(weighted moving average)을 이용하여 자료에 존재하는 임의의 잡음(random noise)을 제거함으로써 기본적인 그래프를 도시한다. 여기에 안전성능함수를 규모조정계수로 활용하여 실제 위험구간을 검토한다. 최종적으로 연도별 재현성(reproducibility)을 검토하여 사고취약구간을 선정하게 된다.

2. 기존문헌 고찰

AASHTO (2010)는 Highway Safety Manual을 통해 도로의 안전성 평가와 관련된 체계적인 분석과정을 제시하였다. 그 중 사고취약구간을 선정하는 세 가지 방법인 SMW, PS, simple ranking을 제시하고 있으며, 그에 관한 예를 수록하고 있다.

Chung and Ragland (2007)는 SMW를 보완하기 위한 새로운 방법론인 CRP를 개발했다. 단위거리 당 평균사고건수를 평가지표로 사용하였다. 연도별 CRP 그래프를 통해 연속적으로 Peak를 기록하는 지점을 사고취약구간으로 선정하게 된다. 이 연구는 규모조정계수로

분석 대상기간의 평균사고건수를 사용한 매우 간단한 형태이다.

Yu (2008)는 2007년 개발된 CRP를 국내 실정에 맞게 변수를 조정하고 곡선반경 및 교통량을 고려하여 국내 고속도로 4개 구간 207km에서 일어난 최근 10년간 사고 자료에 적용하였다. 이 연구를 통해 국내 고속도로의 연속된 위험도를 파악하고자 하였다.

Chung and Ragland (2009)는 CRP 분석을 통해 사고취약구간을 사전 검지하는 방법과 해당구간 및 그 이웃구간의 안전개선효과를 측정하는 방법에 대하여 연구하였다.

Oh et al. (2009)은 CRP 적용을 통해 젖은 노면상태에서 높은 사고율을 야기하는 지점의 특성에 대해 알아보았다. 적용 결과, 속도는 노면상태에 관계없이 사고에 영향을 주는 요소지만 젖은 노면 상태인 경우 더 중요한 영향을 미친다고 기술하였다.

Chung et al. (2010)은 CRP를 이용하여 사고취약구간으로 분류될만한 사고율을 가지는 않았지만 시스템적인 악화가 관측되는 장소를 사전 감지하는 알고리즘에 관하여 연구하였다. 이 연구를 통해 가중이동평균법이 적용된 CRP와 CUSUM algorithm을 사용하여 시스템적인 악화가 관측되는 구간을 선정하였다.

Kim (2011)은 3년간의 경부선, 서해안선 사고 자료에 CRP를 적용해보고, EPDO(equivalent property damage only) 가중치를 적용함에 따라 나타나는 차이를 비교하는 연구를 진행하였다. 적용 결과, EPDO 가중치를 이용하는 것이 상대적으로 안정적인 결과를 도출하는 것으로 나타났다.

마지막으로 동질구간 분할과정이 야기하는 분석의 민감도와 SMW, PS, CRP의 차이를 보다 잘 설명할 수 있는 Kwon et al. (2012)의 논문은 세 가지 방법론의 효율성을 평가하기 위해 작성되었다. 동질구간 분할에 상관없이 CRP에 의한 실제 위험구간을 포함하는 사고취약구간의 수가 가장 적으므로 CRP의 성능이 우수하다는 결론을 내렸다.

CRP는 개발초기단계로, 안전성능함수를 규모조정계수로 이용한 CRP가 다른 사고취약구간 선정 방법론에 비해 효과적이라는 사실이 앞서 설명한 다양한 연구결과를 통해 입증되었다. 그러나 CRP는 분석하고자 하는 국가나 지역에 안전성능함수가 구축되어있느냐에 따라 사용가능여부가 달려있으며, 안전성능함수를 규모조정계수로 적용하지 않은 경우는 다른 방법론에 비해 사고취약

구간을 효과적으로 선정한다고 말할 수 없다. 따라서 본 연구에서는 안전성능함수가 구축되어 있지 않은 국가나 지역에서도 활용할 수 있는 새로운 CRP 분석 방법론을 개발하고자 한다.

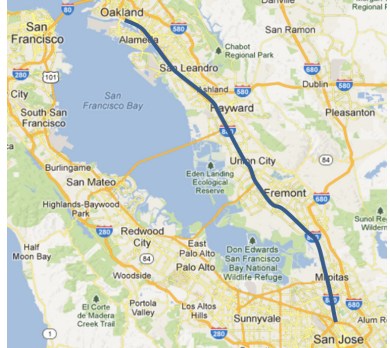
III. Continuous Risk Profile 구축

1. 기초자료의 구축

사고취약구간을 선정하는 방법론간의 효율성을 비교하기 위해서는 기준이 되는 도로의 실제 위험구간(true hotspot)에 대한 정보가 필요하다. 미국 캘리포니아 교통국(California Department of Transportation)에서는 교통전문가의 현장조사를 통해 선정된 도로의 실제 위험구간을 Table C라는 이름으로 체계적으로 관리하고 있다(Kwon et al., 2012). 우리나라의 경우에는 단순한 사고건수를 활용하여 고속도로 노선별로 사고 잦은 지점을 선정하고 개선사업을 시행하고 있지만 실제 위험구간에 대한 관리가 이루어지지 못하고 있는 실정이다. 이러한 이유로 본 연구에서는 국내의 자료를 이용하기에는 한계가 있으므로 미국 캘리포니아주의 고속도로 자료와 안전성능함수를 토대로 분석을 시행하고자 한다.

분석의 시간적·공간적 범위와 분석에 필요한 자료는 Table 1과 같다. 캘리포니아 교통국의 Statewide Integrated Traffic Records System (SWITRS) 자료를 활용하였다.

Table 1. Study scope and data

Division	Contents
Temporal scope	2004-2008 (for 5 years)
Data	Collision data, AADT, # of lanes
Spatial scope	I-880 in Alameda County California, United States 

2. 평가지표 산출 및 임의의 잡음 제거

분석대상 노선의 이정에 해당하는 사고건수의 단위거리 당 평가지표를 산출하는 과정이다. 사고취약구간 선정 을 위해 선택 가능한 평가지표는 평균사고건수, 사고율, 대물피해환산법을 적용한 평균사고건수, 안전성능함수를 이용한 초과 예측된 평균사고건수, EB Adjustment를 이용한 예측된 평균사고건수 등이 있다(AASHTO, 2010). 본 연구에서는 평균사고건수를 평가지표로 선택 하여 분석을 진행한다.

선택된 평가지표인 평균사고건수에 가중이동평균법을 적용한다. 이 과정은 평가지표에 포함된 임의의 잡음을 제거하기 위해 적용되는 단계로, Window의 중심에 가 중치를 주는 가중이동평균법을 이용하면 CRP 그래프의 Peak를 시각적으로 뚜렷하게 표현할 수 있다는 장점이 있다.

식(1)은 CRP에 적용되는 가중이동평균법을 설명하 고 있다. $2L$ (Window)과 l (Increment)의 설정에 따 라 CRP의 결과는 영향을 받는다. $2L$ 의 길이를 증가시 킬수록 잡음을 효과적으로 제거할 수 있지만, 사고가 실 제로 많이 일어난 Peak의 효과가 감소하는 현상이 나타 난다. 따라서 Peak의 위치에 영향을 주지 않으면서 임 의잡음을 제거할 수 있는 최적의 $2L$ 길이인 0.2mile과 0.01mile의 l 을 이용한다(Chung et al., 2010).

$$M(d) = \frac{\sum_{i=-\min(L/l, (d-d_0)/l)}^{\min(L/l, (d_{end}-d)/l)} (L/l - |i| + 1) \times A(d + i \times l)}{\sum_{i=-\min(L/l, (d-d_0)/l)}^0 -i + \sum_{j=0}^{\min(L/l, (d-d_0)/l)+1} j + (L/l + 1)} \quad (1)$$

3. 규모조정계수 적용

1) 기존 CRP

기존의 CRP는 사고취약구간의 선정 과정에 적용되는 규모조정계수로 안전성능함수를 사용한다. 단위거리 당 평가지표에 가중이동평균법을 적용한 값을 $M(d)$ 라 하 고, 구간 당 구축된 안전성능함수가 해당 구간에서 발생 할 것이라고 예측되는 사고건수를 $B(d)$ 라 한다. 따라서 $M(d)$ 에서 규모조정계수인 안전성능함수, $B(d)$ 를 제외 한 부분($K(d) = M(d) - B(d)$)이 예측 값을 벗어난 초과사고건수로 검지가 되고, 이 초과사고건수가 발생하

Table 2. Highway rate groups for the I-880 freeway

Division	Description	Base rate	ADT factor
H55	Rural freeway 5-6 lanes	0.25	0.0050
H56	Rural freeway 7 lanes or more	0.20	0.0035
H61	Suburban freeway 5-6 lanes	0.20	0.0060
H62	Suburban freeway 7 lanes or more	0.25	0.0035
H64	Urban freeway 5-6 lanes	0.40	0.0055
H65	Urban freeway 7-8 lanes	0.40	0.0035
H66	Urban freeway 9-10 lanes	0.35	0.0030
H67	Urban freeway 11 lanes or more	0.35	0.0025

는 구간이 최종적으로 사고취약구간으로 선정된다.

분석에 필요한 규모조정계수인 안전성능함수를 얻기 위 해서는 차로 수, 교통량, 기하구조 등을 반영하여 분류한 Highway Rate Group이 필요하다. 미국 캘리포니아 교 통국에 의해 분류된 Highway Rate Group은 총 67개로, 분석 대상구간인 I-880 고속도로는 8개의 Highway Rate Group으로 구성되어 있다. Table 2는 I-880 고속도로의 안전성능함수에 의한 예측사고건수를 얻기 위해 필요한 Highway Rate Group의 내용 및 수치이다(California Department of Transportation, 2002).

I-880 고속도로의 Highway Rate Group에 해당하 는 Base Rate과 ADT Factor, 해당 구간의 AADT를 이용하여 예측사고건수를 산출한다. 식(2), (3), (4)는 안전성능함수를 통해 사고건수의 예측치를 구하는 과정 을 보여주고 있다.

$$R_E = Base\ Rate + ADT\ Factor \times Average\ ADT \quad (2)$$

$$N_E = \frac{R_E \times Travel}{10^6} \quad (3)$$

$$Travel = AADT \times 365\ Days \times Length \quad (4)$$

R_E : expected collision rate per Vehicle-Mile-Travel (VMT) determined for highway group E

N_E : expected number of collisions for highway group E

2) 계층적 군집분석을 활용한 CRP

본 연구의 차별성은 CRP 분석에서 규모조정계수로 방대한 자료가 필요한 안전성능함수 대신 상대적으로 필요한 자료가 적고, 계산이 용이한 계층적 군집분석을 사용한다는 것이다.

고속도로 사고와 관련한 통계 자료는 교통량(AADT), 차로 수, 기하구조, 안전시설물 설치정보 등이 포함될 수 있다. 이 중 AADT와 차로 수의 경우 구득 및 활용이 비교적 용이한 반면, 기하구조와 안전시설물의 경우 실제 설계도면 및 안전시설물 설치 사업의 디지털화 과정을 포함하고 있어 DB 구축에 많은 비용과 시간이 소요된다. 따라서 본 연구에서는 구득 및 활용이 용이한 AADT와 차로 수 정보를 기반으로 i) AADT+차로 수, ii) 차로 수, iii) AADT를 변수로 하는 계층적 군집분석을 수행하였다. SPSS를 이용한 계층적 군집분석을 통해 4가지 그룹으로 분류한 후, 그룹별 평균사고건수를 규모조정계수, $B(d)$ 로 설정하여 $M(d)$ 에 적용하였다.

기존의 CRP에서 사용하는 안전성능함수의 구축과정은 동질구간 분할과 회귀분석 과정으로 이루어진다. 이 중 동질구간 분할 과정은 안전성능함수의 정확도에 큰 영향을 미치는 과정으로, 교통량뿐만 아니라 세부 기하구조를 모두 고려해야 한다. 이러한 자료는 단기간 내에 구축이 힘들뿐만 아니라 방대한 관련 자료가 갖춰져 있어야 한다. 따라서 본 연구에서는 안전성능함수가 구축되어 있지 않은 국가나 지역에서 사용할 수 있도록 그룹별 평균사고건수를 규모조정계수로 사용하여 사고취약구간을 선정하고자 한다.

i) AADT+차로 수, ii) 차로 수, iii) AADT를 변수로, 단위이정 당 평균사고건수를 분석 시나리오로 설정하여 계층적 군집분석을 수행한 후, 설정된 군집에 따

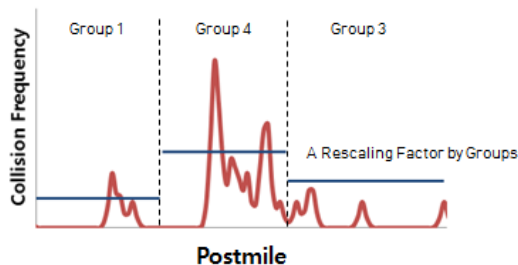


Figure 4. CRP with hierarchical clustering analysis

라 평균사고건수를 도출하여 이를 규모조정계수로 활용하였다.

군집분석은 관측 개체를 몇 개의 그룹으로 나눔으로써 대상 집단에 대한 이해 및 효율적인 활용을 가능하게 하기 위한 것으로서, 보통 케이스 수를 줄이는데 목적이 있다. 특히 계층적 군집분석은 내부적으로 동질적인 다수의 그룹으로 p 개 변수가 측정된 자료 n 개 개체를 단계적으로 묶어 군집을 형성하는 분석방법으로, 군집을 형성하는 방법에는 최단 연결법, 최장 연결법, 중심 연결법, 중위수 연결법, 군집 간 평균연결법, 군집 내 평균연결법, Ward 군집법이 있다(Berthouex and Linfield, 2002). 이 중에서 Ward 군집법은 각 그룹별 설명변수의 표준편차를 최소화하기 위한 방법이다(Hair et. al., 2010). 본 연구에서는 노출변수(exposure variable)의 분산을 최소화함으로써 각 그룹별 유사성이 정의된다고 가정하였다. 이러한 이유로 본 연구에서는 SPSS를 이용하여 Ward 군집법으로 계층적 군집분석을 시행하였다.

Table 3, 4, 5는 계층적 군집분석을 활용한 그룹별 평균사고건수의 결과를 보여주고 있다. 군집분석을 통하

Table 3. Rescaling factors by clustering (AADT + Lanes)

Group	Count	AADT	Number of lanes	Average crash frequency
1	1,108	72,735	3.04	0.22
2	986	93,267	3.00	0.24
3	2,119	113,040	4.69	0.51
4	340	55,562	7.00	0.21

Table 4. Rescaling factors by clustering (AADT)

Group	Count	AADT	Average crash frequency
1	2,042	78,034	0.25
2	1,991	108,045	0.45
3	398	132,048	0.53
4	122	32,434	0.11

Table 5. Rescaling factors by clustering (Lanes)

Group	Count	Number of lanes	Average crash frequency
1	2,062	3	0.23
2	677	4	0.57
3	1,474	5	0.48
4	340	7	0.21

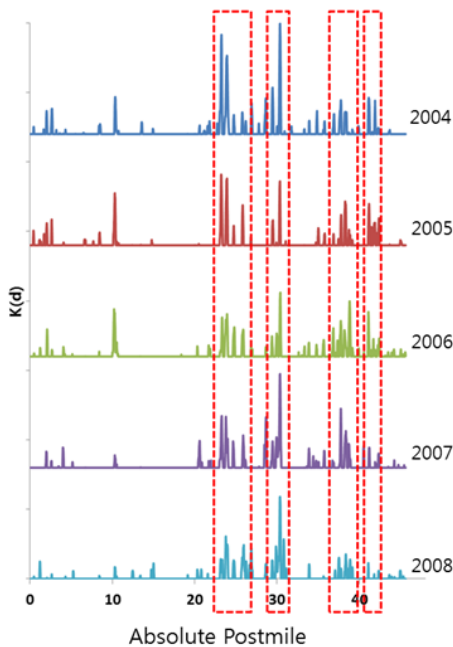


Figure 5. Hotspots identification

여 분석 시나리오마다 총 4개의 그룹으로 분류되었으며, 각 그룹별 평균사고건수를 산정하였다. 이 값이 규모조정계수 $B(d)$ 로 적용된다.

4. 재현성 검토 및 최종 사고취약구간 선정

$M(d)$ 에 규모조정계수를 적용하여 나온 값을 $K(d)$ 라 하면, $K(d)$ 가 특정 Postmile에서 연도별로 예측 값을 초과한 사고건수를 보였는지 연도별 재현성을 검토하는 단계이다. 연속된 5개년도(2004-2008)의 예측 값을 초과한 사고건수가 누락되지 않고 기록된다면, 해당 구간은 연도별 재현성이 검토되었다고 할 수 있다.

재현성까지 검토가 완료된 구간을 CRP에 의한 최종 사고취약구간을 선정한다. Figure 5는 기존의 CRP에 의해 선정된 I-880 northbound의 2004년부터 2008년까지의 주요 사고취약구간이다.

IV. 결과 분석

1. I-880 Northbound의 사고취약구간

세 가지 군집 i) AADT+차로 수, ii) 차로 수, iii)

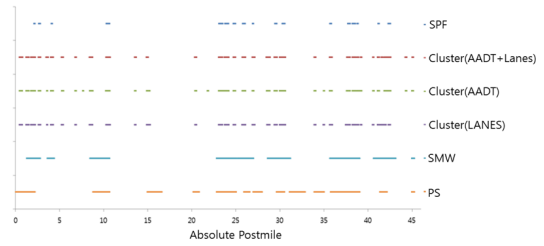


Figure 6. Comparison between absolute postmile of hotspots on the I-880 northbound freeway

AADT를 이용하는 CRP에 의해 선정된 사고취약구간이 기존의 CRP에 의해 선정된 사고취약구간을 모두 포함한다면, 이는 실제 사고취약구간임에도 불구하고 사고 잦은 지점으로 검지되지 않는 구간(false negatives)이 발생하지 않음을 의미한다. 또한 세 가지 군집을 이용하는 CRP에 의해 선정된 사고취약구간의 길이가 SMW와 PS에 의해 선정된 사고취약구간의 길이보다 상대적으로 짧다면, 이는 군집을 이용한 CRP의 false positive rate이 작다는 것을 의미한다.

Figure 6은 본 연구에서 분석된 다양한 방법론에 의해 선정된 실제 사고취약구간의 이점과 길이를 비교한 것이다. 분석된 모든 방법론에서 false negatives는 발견되지 않았다. false negatives는 예산의 배분문제가 아니라 운전자의 안전과 직접적인 연관이 있으므로 새로운 방법이 false negatives를 발생시킨다면 적용이 불가능하다는 의미가 된다. 계층적 군집분석을 활용한 CRP에 의해 도출된 사고취약구간이 기존의 CRP에 의한 사고취약구간을 모두 포함하므로 본 논문에서 제시하고 있는 새로운 CRP 방법론은 false negatives를 발생시키지 않음을 알 수 있다.

분석 대상구간의 총 길이는 45.72mile로, Table 6을 참고하여 선정된 사고취약구간의 길이가 짧은 순서대로 나열할 수 있다. 그 순서는 기존의 CRP, 차로 수를 변수로 한 군집법을 이용한 CRP, AADT를 변수로 한 군집법을 이용한 CRP, AADT와 차로 수를 변수로 한 군집법을 이용한 CRP, SMW, PS 순이다. 즉, 해당 결과는 false positive rate이 작은 순서도 이와 같다는 의미이며, 이를 통해 계층적 군집분석을 이용한 CRP가 안전성능함수를 이용하는 기존의 CRP보다는 성능이 다소 떨어지지만, 다른 사고취약구간 선정 방법론인 SMW와 PS보다 성능이 우수하다고 할 수 있다.

Table 6. Length of hotspots on the I-880 northbound freeway

Method	Rescaling factor	Lengths of hotspots (miles)	Rate (%)
Existing CRP	SPFs	2.42	5.29
CRP with clustering	AADT + # of lanes	6.3	13.78
	AADT	6.15	13.45
	# of lanes	5.54	12.12
SMW	-	15.95	34.89
PS	-	17.61	38.52

Table 7. Length of hotspots on the I-880 southbound freeway

Method	Rescaling factor	Lengths of hotspots (miles)	Rate (%)
Existing CRP	SPFs	2.15	4.70
CRP with clustering	AADT + # of lanes	7.23	15.81
	AADT	7.41	16.21
	# of lanes	8.03	17.56
SMW	-	26.37	57.68
PS	-	23.12	50.59

2. I-880 Southbound의 사고취약구간

Figure 7은 I-880 Northbound에 적용한 과정과 마찬가지로, Southbound의 전체 이정에 대하여 각 사고취약구간 선정 방법론을 적용하여 실제 사고취약구간의 이정과 길이를 표현한 것이다.

Figure 7을 살펴보면 앞서 기술된 I-880 Northbound의 결과와 마찬가지로 세 가지 군집을 이용한 CRP에 의한 사고취약구간이 기존의 CRP에 의한 사고취약구간을 모두 포함하는 것을 확인할 수 있다. 이것은 I-880 Southbound 적용사례에서도 본 연구의 CRP 방법론이 false negatives를 발생시키지 않는다는 것을 의미한다. 선정된 사고취약구간의 길이가 짧은 순서는 기존의 CRP, AADT와 차로 수를 변수로 한 군집법을 이용한 CRP, AADT를 변수로 한 군집법을 이용한 CRP, 차로 수를 변수로 한 군집법을 이용한 CRP, PS, SMW 순이다. Northbound와 Southbound의 경우를 비교해보면, 계층적 군집분석을 이용한 CRP의 변수에 따라 순서의 차이가 있으며, SMW와 PS의 순서에도 차이가 있다.

이상의 분석 결과에서 알 수 있듯이 본 연구에서 제시한 계층적 군집분석 기반의 CRP는 기존의 CRP 기법과 비교할 때, false negatives는 발생시키지 않는 반면,

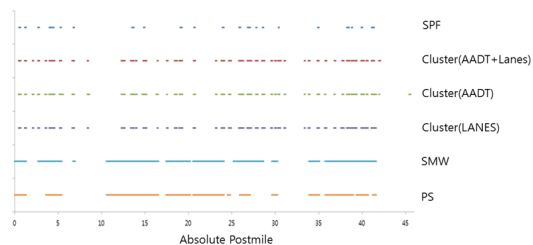


Figure 7. Comparison between absolute postmile of hotspots on the I-880 southbound freeway

양방향 모두 false positive rate이 다소 증가하는 것으로 분석되었다. 그럼에도 불구하고 데이터의 간편성과 분석의 용이성 측면에서 더 넓게 적용이 가능할 수 있으리라 판단되며, 기존에 널리 사용되고 있는 SMW와 PS 보다는 성능이 우수하다는 것이 양방향에서 모두 확인되었다.

V. 결론 및 향후 연구과제

CRP 방법론을 통해 선정된 사고취약구간이 SMW나 PS를 통해 선정된 사고취약구간에 비해 false positive rate이 적으며, 이는 CRP가 다른 사고취약구간 방법론에 비해 효율성이 우수하다는 것을 의미한다. 이러한 결과는 다양한 연구를 통해 입증된 바 있다.

본 연구에서는 안전성능함수 구축을 필요로 하는 기존의 CRP 기법과 달리 다양한 통계자료 구축이 미비한 국가나 지역에서도 활용이 가능한 방법론을 모색하는 데 중점을 두었다. 연구의 목적을 달성하기 위해 안전성능함수 대신 간단한 고속도로 정보만을 가지고 계층적 군집분석을 시행하여 얻은 그룹별 평균사고건수를 사용하였다. 그룹별 평균사고건수를 규모조정계수로 이용한 CRP의 성능을 확인하기 위해, 안전성능함수가 구축되어 있는 캘리포니아주 Alameda의 I-880 고속도로 사고 자료를 이용하여 다양한 방법론이 사고취약구간을 선정하는 성능을 false positive rate과 false negative rate 측면에서 비교하였다.

계층적 군집분석(변수: AADT, 차로 수, AADT+차로 수)을 통해 얻은 그룹별 평균사고건수를 규모조정계수로 이용한 CRP를 통해 선정된 사고취약구간이 기존의 CRP로 선정된 사고취약구간을 모두 포함하며(false negatives가 발생하지 않음), SMW나 PS로 선정된 사

고취약구간보다 false positive rate이 작다는 것을 확인할 수 있었다. 이는 군집법을 이용한 CRP의 성능이 기존의 CRP보다는 다소 떨어지지만, 다른 사고취약구간 선정 방법론보다는 우수하다는 의미를 지닌다. 이를 통해 안전성능함수가 구축되어 있지 않은 국가나 지역에서 계층적 군집분석을 이용하여, 다른 방법론보다 성능이 우수한 CRP를 사고취약구간 선정에 적용할 수 있을 것이다. 또한 고속도로 DB의 구축능력을 보유하고 있는 국가나 지역이라 할지라도 누적된 자료가 부족한 신규 고속도로의 안전계획을 수립하는데 활용할 수 있을 것이다.

계층적 군집분석을 이용한 CRP를 사고취약구간 선정에 적용함으로써, false positive rate을 줄여 사고취약구간을 효과적으로 선정할 수 있으며, 이는 한정된 예산을 사고감소효과가 높은 구간에 효율적으로 배분하는 데에 기여할 수 있다. 또한, 사고취약구간 선정 결과를 바탕으로 고속도로의 실제 위험구간 탐색에 투여되는 전문가들의 시간 및 노동력을 절감하는 효과를 얻을 수 있다.

새로운 방법론은 구득이 용이한 기본적인 고속도로 정보를 활용할 수 있으며, 기존의 방법론에 비해 상대적으로 분석이 쉽고 간단하여 활용성이 높다는 장점이 있다. 그러나 안전성능함수를 활용하는 기존의 방법론에 비해 사고위험구간을 검지하는 성능이 다소 떨어지는 점은 분명한 한계점이라 할 수 있다.

본 연구에서는 노출변수(exposure variable)의 분산을 최소화함으로써 각 그룹별 유사성이 정의된다고 가정하고 Ward 군집법을 활용하였다. 향후 부족한 사고정보를 가지고 좀 더 효율적으로 그룹별 평균사고건수를 산출할 수 있는 군집방법을 연구하는 것은 흥미로운 주제가 될 것이다. 다양한 환경 속에서 적용이 가능한 변수를 탐색하고, 다양한 군집화 방법을 적용하여 개선된 방법론을 제안할 수 있을 것이다.

또한, 고속도로 DB의 구축수준이 다른 실제 국가 및 지역에 대한 사례분석이 필요하다. 사례분석을 토대로 본 연구에서 제시하는 방법론의 전이성 및 활용성에 대한 평가를 수행하고, 이를 토대로 개선방안을 모색하는 것도 중요한 향후 연구과제 중 하나일 것으로 사료된다.

ACKNOWLEDGEMENT

The authors thank Dr. Koohong Chung of

California Department of Transportation for his help and comments.

REFERENCES

- Alluri P., Ogle J. (2012), Effects of State-specific SPFs, AADT Estimations, and Overdispersion Parameters on Crash Predictions Using SafetyAnalyst, TRB 2012 Annual Meeting, CD-ROM 12-4332.
- American Association of State Highway and Transportation Officials (AASHTO) (2010), Highway Safety Manual: 1st Edition, American Association of State Highway and Transportation Officials.
- Berthouex P. M., Linfield C. B. (2002), Statistics for Environmental Engineers: Second Edition, Lewis Publisher.
- California Department of Transportation (2002), Table C task force: Summary Report of Task Force's Findings and Recommendations, Caltrans.
- Chung K. H., Jang K. T., Madanat S., Washington S. (2010), Proactive Detection of High Collision Concentration Locations on Highways, TR Part A: Policy and Practice, Vol.45, No.9, pp.634-645.
- Chung K. H., Ragland D. R. (2007), A Method for Generating a Continuous Risk Profile for Highway Collisions, TRB 2007 Annual Meeting, CD-ROM 07-2935.
- Chung K., Ragland D. R., Oh S. M. (2009), The Continuous Risk Profile Approach for the Identification of High Collision Concentration Locations on Congested Highways, Transportation and Traffic Theory, pp.463-480.
- Hair J. F. Jr., Black W. C., Babin B. J., Anderson R. E. (2010), Multivariate Data Analysis: 7th Edition, Prentice Hall
- Kim Y. I. (2011), A Study on Crash Severity Based Continuous Risk Profile Approach for Identifying Hazardous Expressway Sections, Master's Thesis, Seoul National University.
- Kwon O. H., Park M. J., Chung K. H., Yeo H.

- S.(2012), Comparing the Performance of Sliding Moving Window, Peak Searching, and Continuous Risk Profile Methods for Identifying High Collision Concentration Locations, TRB 2012 Annual Meeting, CD-ROM 12-2293.
- Oh S. M., Chung K. H., Ragland D. R., Chan C. Y. (2009), Analysis of Wet Weather Related Collision Concentration Locations: Empirical Assessment of Continuous Risk Profile, TRB 2009 Annual Meeting, CD-ROM 09-0366.
- Yu J. S. (2008), An Analysis of the Hazardous Highway Segments Using CRP (Continuous Risk Profile) Method, Master's Thesis, Hanyang University.
- ☞ 주 작성 자 : 이서영
 - ☞ 교 신 저 자 : 이철원
 - ☞ 논문투고일 : 2013. 4. 13
 - ☞ 논문심사일 : 2013. 5. 28 (1차)
 - 2013. 6. 27 (2차)
 - 2013. 7. 17 (3차)
 - ☞ 심사판정일 : 2013. 7. 17
 - ☞ 반론접수기한 : 2013. 12. 31
 - ☞ 3인 익명 심사필
 - ☞ 1인 abstract 교정필