



K-평균법을 이용한 고속도로 사고분석구간 분할기법 개발

Selecting Technique of Accident Sections using K-mean Method

이 기 영* 장 명 순*

Lee, Ki Young Chang, Myung Soon

Abstract

A selection of the analysis section for traffic accidents is used to analyze definitely the cause of accidents sorting similar accidents by a group and to raise the effect of improvement projects deciding the priority of accidents. In the existing method, an uniformly dividing method based on road mileages has been used, which has no consideration for similarities among accidents. Consequently, in recent, a slider-length method considering accident types rather than road mileages is widely used. In this study, using K-mean method, a non-hierarchical grouping technique used in the Cluster Analysis as a applicatory method for the slider length method, a method classifies accidents that occurred the most nearby mileages into one group is proposed. To verify the proposed method, a comparison between the K-mean method and the dividing method at regular intervals on the data of a total of 25.6 km lengths along Kyung-bu freeway in Pusan direction was made so that the K-mean method was proved to an effective method considering the similarities and adjacencies of accidents.

Keywords : cluster analysis, K-mean method, number of accident, black spot

요 지

교통사고 분석구간 선정은 유사한 사고유형을 하나의 동일집단으로 처리함으로써 분석의 효율성을 높이고, 효과적인 개선 사업을 위해 그 순위를 결정하는데 필요한 작업이다. 기존에는 도로이정에 기초하여 균등하게 분할하는 방법을 주로 사용하여 왔는데, 사고간의 유사성을 전혀 고려하지 못하는 단점이 있다. 따라서 최근 도로이정보다는 사고간의 유사성을 고려하여 구간을 선정하는 방식인 Slider length 적용기법이 사용되고 있다. 본 연구에서는 Slider length 적용기법의 한 방법론으로써, 군집분석에 사용되는 비계층적 분류기법인 K-평균법을 사용하여 가장 인근거리에 발생된 사고들이 최대한 하나의 집단으로 분류될 수 있는 기법을 제시하고자 한다. 또한 이의 검증을 위해 경부고속도로 부산방향으로 총연장 25.6km구간에 대하여 균일간격에 의한 분할방식과 K-평균법을 이용한 분할방식에 대해 상호 비교를 통해 그 효율성을 검토하였으며, K-평균법이 사고의 유사성이나 인접성을 감안하여 효율적으로 분석구간 선정에 적용될 수 있음을 검증하였다.

핵심용어 : 군집분석, K-평균법, 사고수, 사고 잦은 구간

* 정희원 · 한국도로공사 도로교통기술원 책임연구원

** 한양대학교 교통시스템공학과 정교수



1. 서론

교통사고 분석구간을 구분하는 목적은 상호 인근에 위치한 유사한 형태의 사고들을 하나의 집단으로 처리함으로써 사고 발생의 원인을 명확히 해석하고, 사고 잦은 구간에 대한 순위를 결정하여 효율적인 교통사고 개선사업을 시행하고자 하는데 있다.

여기서 “교통사고 잦은 구간”은 다른 유사한 조건을 가진 장소와 비교해서 사고율이 평균치보다 높은 곳을 말하는데, 단일로, 교차로, 도로접속지점, 교량, 터널 등 교통사고가 빈번히 발생하는 특정구간으로 상대적, 절대적 선정기준에 의해 선정된 구간을 말한다.

통상 도로개선사업을 위해 사고 잦은 구간을 선정하는데, 그 척도가 되는 주요 지표로는 사고건수, 사상자수, 사고의 피해정도를 고려한 대물피해사고환산법(EPDO:Equivalent Property Damage only) 등이 적용되고 있다.

이러한 사고 잦은 구간을 선정하는 작업도 중요하나, 그 전에 어떻게 유사사고를 최대한 하나의 구간에 소속되도록 구간을 잘 분할할 수 있는가에 대한 검토가 더 중요하다.

기존에는 주로 도로이정에 따라 균등한 간격으로 전 구간을 구분하는 방식이 주로 사용되어 왔으나, 최근 도로이정 기준이 아닌 사고지점을 중심으로 한 사고간의 유사성을 고려하여 분할하는 slider length 적용기법이 사용되고 있다.

본 연구에서는 사고구간 분할시, 군집분석에 사용되는 비계층적 분류기법인 K-평균법을 사용하여 인근거리에 발생한 사고들이 최대한 하나의 집단으로 분류할 수 있는 기법을 제시하고자 한다. 이러한 사고구간 분할방법은 기존 균일간격에 의한 분할방식이 가지는 사고간의 거리적 유사성을 고려하지 못하는 단점을 보완할 수 있으며, slider length 적용기법의 장점을 살리는 하나의 방법일 것이다.

K-평균법은 전체 사고를 K개의 군집으로 분류하는데 있어 각 군집에 속한 데이터의 평균값에 소속

변수들이 가장 밀집하도록 군집을 이루게 함으로써, 이점(거리)에 따라 가장 인근거리에 있는 사고들을 하나의 집단으로 처리되도록 하는 기법이다.

다만, K-평균법을 활용하기 위해서는 분석전에 미리 군집수(K)를 결정해야 하는 이론적 제약이 존재한다. 따라서 전체구간에 사고건수가 매우 적을 때 적절한 군집수를 설정하지 못할 경우, 상호 거리가 먼 사고들이 하나의 집합으로 처리되는 모순이 발생할 수도 있다. 따라서 이를 해결하기 위해 가상의 사고를 균등한 간격으로 삽입함으로써 이론적 제한조건을 완화시키는 방법을 적용하였다.

이 모델의 검증을 위해 경부고속도로 부산방향으로 서울~신갈구간의 총연장 25.6km에 대해 200m를 단위구간으로 하여, 균일간격에 의한 분할방식과 K-평균법을 이용한 방식에 대한 상호 비교를 통해 그 효율성을 검증하였다. 이때 두 분석기법을 동일조건으로 하기 위해 균일간격에 의한 분할기법에서의 분석구간 140개에 맞추어 K-평균법도 140개의 군집수를 적용하였으며, 여기에 200m단위로 가상 사고건수(140건)를 삽입하여 적용하였다.

2. 기존방법 및 연구사례

2.1 기존방법

교통사고 분석구간을 구분하는 것은 크게 두 가지 목적이 있다. 하나는 전체 도로구간을 조사하는데 필요한 인력, 비용 및 시간적 한계를 고려하여 이를 순차적으로 관리하는데 있으며, 또 하나는 유사한 유형을 가진 사고들을 집합화하여 사고분석의 효율성을 기하고자 하는데 있다.

먼저, 전체구간을 어떠한 방식으로 분할하는가에 대한 접근방법론은 두 가지로 나눌 수 있다.

첫째, 도로이정에 따라 일정간격으로 균일하게 나누어 분석하는 방법이다. 이러한 방법은 간편하고 처리속도가 빠른 장점이 있는 반면, 사고간의 유사성을

고려하지 못하는 단점이 있다. 그러나 분석자의 잘못된 주관적 판단을 배제시키거나, 사고가 전체구간에 균일하게 분포되어 있을 경우에는 그 활용도가 높은 방법이다.

둘째, 사고의 유사성을 고려하여 모든 사고지점을 분석구간의 시점, 중점, 종점으로 구분하여 상호 유사성을 고려하여 구간을 결정하는 방법이 있다. slider length 적용기법이 대표적인 예라 할 것이다. 이런 방식은 분석의 정확도는 높으나 많은 분석시간이 소요되며, 무엇보다도 구간 길이를 어떻게 결정할 것인가, 또한 어떠한 방법으로 이동시키면서 사고 구간을 구별하는가에 대한 방법에 따라 그 효율성이 달라지게 된다.

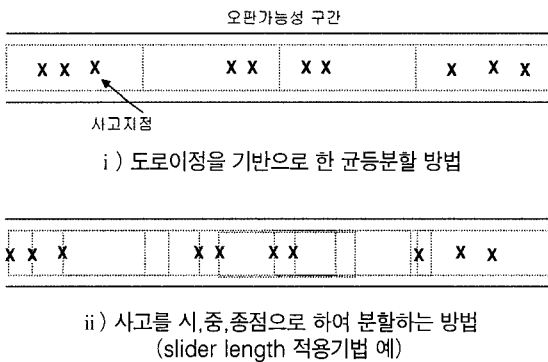


그림 1. 사고구간 분할기법의 종류

적용기법에 앞서 무엇보다도 중요한 점은 분석구간 길이를 어느 정도에서 결정해야 하는가에 대한 부분이다. 구간길이는 소속되어 있는 사고들이 충분한 유사성을 가지고 있어야 하며, 실제 사고분석단계나 대응방안 수립시에 적용이 가능하도록 적정 범위내에서 결정해야 한다. 따라서 너무 길거나 짧아도 그 효율성이 떨어지게 된다.

본 연구에서 제시한 K-평균법은 구간수의 개념이 아닌 군집수로 표현된다. 또한 K-평균법에 의한 분할은 사고간의 거리적 인접성을 고려한 분할방법이므로 각 군집의 길이는 다양하게 나타날 것이다.

2.2 기존 연구사례

기존연구는 적정 사고분석구간 길이를 산정하는가에 대한 부분보다는 어떠한 구간을 사고 잦은 구간으로 선정하는가에 대한 부분에 초점이 맞추어져 있어 연구사례는 매우 제한적이다.

김정현, 이수범, 박병정(2002)은 적정 교통사고 잦은 구간 길이를 산정하기 위해 최소정지시거에 의한 접근방법을 제시하였다. 즉 설계속도가 80km/h 인 도로구간에서의 최소정지거리인 140m를 준용하여 적정 구간 길이로 제시하였다. 또한 이를 바탕으로 사고 잦은 구간 선정을 위해 모든 사고를 분석구간의 시점, 중점, 종점으로 적용하여 분석하는 방법 (slider length 적용기법)을 제안하였다. 특히 사고지점의 시점과 종점을 기준으로 하였을 경우는 동일한 결과를 도출한 반면, 중점을 기준으로 하였을 경우 사고의 잠재적 위험성을 가진 지점을 더 많이 찾게 되는 결과를 나타냈다. 여기서는 분석구간을 140m의 단일거리로 분할하였으므로, 만약 인근에 위치한 5개의 사고가 200m에 걸쳐 발생할 경우 이를 동일구간으로 처리하지 못하는 단점이 존재한다.

김태완(1995)은 표1에서와 같이 도로의 기하구조를 대상으로 3가지 방법의 구간분할법을 제시하였다. 이중 분할법 C로 구간을 분할하여 사고분석을 한 경우가 사고추정에 대한 높은 결정력을 보여주고 있음을 밝혀냈다. 즉 동일한 기하구조를 갖는 구간으

표 1. 경부고속도로 기하구조관련 구간분할 방법

분할법	기 준	구 간 수
A	경부고속도로 전구간을 100m 단위로 균등 분할	4,179구간
B	분할법 A로 분할 후 기하구조 및 교통량을 감안하여 인접구간을 묶음	0.1~1.8km 연장의 3,137개 구간
C	중심이 되는 평면선형·종단선형을 기준으로 비교적 긴 구간을 한개의 관측 구간으로 선정	0.2~2.8km 연장의 760개 구간

자료 : 김태완(1995)



로 사고분석구간을 선정하는 것이 사고분석에 가장 유리한 조건임을 보여주고 있다.

Shaw-Pin Miaou(1994)는 등구간이 지나치게 짧거나 등구간으로 분할하지 않았을 경우의 문제점을 제시하였다. 즉 구간이 1 mile 이하일 경우는 사고 분석시 바람직한 결과를 얻기 힘들다고 주장하였으며, 또한 등구간으로 분할하지 않을 경우 너무 긴 구간이 발생하면 동일구간에서 기하구조의 연속성을 유지하지 못하는 단점도 지적하였다.

Bhagman Persaud(2000)등은 도로의 직선부와 곡선부에 대한 구간분할을 통해 곡선부의 사고율에 대한 사고빈도가 높음을 증명하였는데, 이 같은 결과는 도로기하구조의 특성에 따라 사고유형에 차이가 있음을 나타냄으로써 구간분할시 도로기하구조 특성을 고려하여 분할하는 것이 타당함을 제시하고 있다.

이기영, 이용택, 장명순(2003)은 경부고속도로 하행 전체구간을 200m단위의 등구간으로 분할하고, 각 구간별로 발생한 사고와 사고피해도를 고려하여 양적, 질적피해에 대한 각각의 피지함수를 지정하여 이의 조합을 통해 사고취약구간을 선정하는 기법을 제시하였다.

2.3 K-평균법

군집분석은 실험의 결과나 표본 등에서 얻어진 자료들을 어떤 성질에 따라 분류하여 동질적인 몇 개의 군집(cluster)으로 나누어 분류하고자 할 경우에 사용한다. 즉 이 기법은 n 개의 개체를 대상으로 p 개의 변수들을 측정하였을 때, 이 변수들을 이용하여 n 개의 개체들 사이의 유사성 또는 비유사성의 정도를 측정하여 개체들을 가까운 순서대로 연결해 나가는 분석방법이다.

본 연구에서는 군집분석의 하나인 K -평균법을 사용하여 사고구간을 분할하는 기법으로 활용하였다. MacQueen(1967)은 가장 가까운 중심점을 갖는 군집에 각 항목을 할당하는 알고리즘을 설명하면서 K -평균법이란 용어를 처음으로 사용하였다.

K -평균법은 계층적 군집분석과는 달리 개체가 어느 한 군집에만 속하도록 하는 상호배반적 군집방법이다. 이 방법은 군집의 수를 미리 정하고, 각 개체가 어느 군집에 속하는지를 분석하는 방법으로서 대규모의 자료에 대한 군집분석에 유용하게 이용되는 방법이다. 이를 활용한 군집절차는 다음과 같다.

- (1) 전체 개체를 K 군집으로 초기화한다. 이러한 군집의 초기 분류는 분석자가 지정할 수도 있고, 프로그램에서 제공할 수 있다.
- (2) 각 군집의 각각의 변수에 대해서 중심점들을 계산한다.
- (3) 주어진 케이스에 대해서 중심점과의 거리를 계산한다. 만약 현재 속하고 있는 군집의 중심점과 가까우면 그대로 놔두고, 그렇지 않으면 다른 군집으로 배정한다.
- (4) 각 케이스에 대하여 (3)번 절차를 시행한다.
- (5) 어느 케이스도 다른 군집에 재배정되지 않을 때까지 (2), (3), (4) 과정을 반복한다.

여기에서 통상 K -평균법에서 자료의 상호 연관성을 판단하는 지표는 유클리드 거리를 적용한다. 즉 p 차원 상의 두 관찰값 $X = (x_1, x_2, \dots, x_p)$ 와 $Y = (y_1, y_2, \dots, y_p)$ 간의 유클리드 거리는,

$$d(X, Y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_p - y_p)^2} \\ = \sqrt{(X - Y)'(X - Y)} \quad \langle \text{식1} \rangle$$

와 같이 정의된다.

3. 분석기법 선정 및 분석

3.1 분석기법 선정

사고구간을 선정할 때는 크게 3가지 측면을 고려해야 한다. 첫째, 유사한 사고유형이 하나가 되도록 묶을 것, 둘째, 기하구조적인 특징이 같을 것, 셋째, 구간이 너무 지나치게 길거나 짧아 현장조사 및 사고

대책을 세우는데 어려움이 없도록 해야 한다. 이러한 기본 원칙하에 구간을 분할하고 이를 이용해 사고 잦은 구간을 선정하여 개선사업을 수행해야 한다.

본 연구에서는 균일하게 도로구간을 분할하는 방식보다는 사고간의 유사성(상호 거리)을 고려한 접근 방식을 제안하고자 한다. 이러한 분석방법은 사고를 기반으로 하나 동일한 길이로 구간을 분할되는 기존 slider length 적용기법의 단점을 보완하는 방법이다. 이를 위해 여러 군집분석 중 비계층적 방법의 하나인 K-평균법을 이용하여 사고구간을 분할하고자 한다.

K-평균법은 전체 데이터를 K개의 군집으로 분류할 때 각 군집의 속한 데이터의 평균값에 소속 변수들이 가장 밀집하도록 개체를 분할하는 것이다. 즉 군집을 결정할 때 사고가 발생한 위치에 따라 가장 인근거리에 있는 사고들이 최대한 하나의 군집으로 묶여지게 된다.

먼저 각 사고간의 유사성을 상호간의 거리(이정)로 정한다면, 모든 사고들이 군집을 이루기 위한 K-평균법의 목적함수는 아래와 같다. 즉 각 군집의 중점을 기준으로 그 군집에 소속된 모든 사고들과의 거리의 합이 최소화가 되도록 군집을 이루게 하는 것이다.

$$\text{Min.} \sum_{i=1}^n \sqrt{(d_m^c - d_i^c)^2} \quad \langle \text{식2} \rangle$$

여기서, d_m^c : 군집 c 의 중점(이정)($c=1, \dots, k$)
 d_i^c : 군집 c 에 속한 사고 i 의 값(이정)

K-평균법을 적용하기 위해서는 기본적인 이론적인 특성을 감안하여, 사고발생 현황 및 특성에 맞추어 미리 군집수(K)를 결정해 주어야 한다.

한 예로써, 도로 전체구간 길이가 10km이고 사고가 10건이 발생하였을때, 이를 200m단위로 균등하게 분류할 경우 50개의 구간으로 나누어지게 된다. 이 같은 조건에서 K-평균법을 적용할 경우, 군집수(K)를 사고건수인 10개 이상으로 정할 수는 없다. 또한 군집수를 10개 미만으로 하더라도 각 사고간 평균거리가 1km에 이르러 같은 군집에 소속된 사고

간의 거리적 유사성을 찾기는 불가능하다. 따라서 군집수를 적절히 조정하는 작업이 매우 중요하다.

따라서 K-평균법의 적용을 위해 단위구간 200m마다 가상사고를 군집수(K)만큼 삽입하였다. 이러한 방식은 실제 사고가 나지 않는 구간은 가상의 사고만이 포함된 군집으로 처리하여 그 영향을 배제시키고, 실제 사고가 포함된 각 군집의 구간길이를 200m 전후에서 결정되게 하는 효과를 가지게 된다. 아래의 그림 2는 사고가 집중 발생하는 구간을 하나의 군집으로 처리할 수 있음을 보여주고 있다.

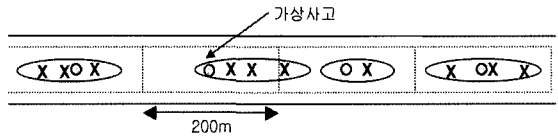


그림 2. K-평균법을 적용한 구간분할기법

위의 예에서 K-평균법을 적용할 경우, 10km의 전체구간에 200m 간격으로 가상의 사고건수 50개가 배정되고, 따라서 총 사고건수 60개는 50개의 군집으로 분리되게 된다.

이러한 가상의 사고를 삽입하는 방식은 K-평균법의 한계를 극복하고, 인근에 위치한 사고간의 군집을 최적화하기 위한 방법으로써 사고에 기반을 둔 slider length 적용기법의 한 예라고 볼수 있다.

본 연구에서 제시한 K-평균법에 의한 분석방법은 다음과 같은 절차에 의해서 수행된다.

- (1) 분석구간의 군집수(K)를 미리 결정한다. 통상 도로교통안전관리공단과 한국도로공사는 사고 잦은 지점 관리를 위해 도로구간을 200m단위로 분할하고 있으며, 따라서 이를 기준으로 군집수를 결정한다.

$$\text{군집수}(K) = \frac{\text{전체구간길이(m)}}{200}$$

- (2) 임의의 가상 사고를 단위구간길이 200m마다 1건씩 발생시켜 총 K개의 가상 사고를 도로에 배정한다.



$$CN_t = CN_r + CN_j$$

CN_t : 전체 사고건수

CN_r : 실제발생한 사고건수

CN_j : 가상 사고건수

(3) K -평균법을 사용하여 모든 사고를 K 개의 군집으로 분할한다. 즉 각 군집에 포함된 사고들과 그 군집의 중점과의 거리를 계산하고, 만약 현재 군집에 소속된 사고가 현재 소속된 군집의 중점(이정)보다 타 군집의 중점(이정)보다 더 가까우면 그대로 군집에 남게 되고, 그렇지 않으면 타 군집으로 배정한다.

(4) (3)의 반복에 따라 어느 사고도 타 군집으로 이동되지 않으면 분석을 완료한다. 만약 특정군집 N_j 에 가상 사고만이 포함될 경우, 그 군집은 사고가 발생되지 않은 집단을 의미한다.

K -평균법에 의한 사고구간 분할이 합리적으로 되었는가에 대한 평가를 위해 본 연구에서는 균일간격에 의한 분할방법과 상호 비교하고자 한다. 그러나 현재로서는 적합하게 사고구간이 구분되었는가를 평가하는 공식적인 기준은 없다. 따라서 각 기법에 의해 선정된 구간(군집)에 소속된 사고의 양을 기준으로 사고 잦은 구간의 순위를 정하고, 이를 통해 비교하고자 한다.

즉, 두 기법은 전체구간을 200m단위에 기초하여 구간을 나누어서 분석하므로 구간수(또는 군집수)는 K 로 동일하다. 따라서 동일한 구간수와 군집수를 가

지게 되므로 그 상태에서 사고 잦은 구간을 효율적으로 결정하는지에 대한 부분을 검토하고자 한다.

3.2 자료수집 및 분석

분석대상지역으로 경부고속도로 부산방향으로 서울기점에서 신갈JCT에 이르는 총 28km에 이르는 구간을 선정하였다. 이 구간은 2002년부터 2004년까지 3년동안 총 88건의 교통사고가 발생하였으며, km당 3.14건이 발생하였다.

본 연구에서는 균일간격에 의한 분할 기법과 K -평균에 의한 분할 기법을 비교하기 위해, 표2에서와 같이 동등한 조건으로 선정하고자 한다.

먼저 구간분할 기준거리를 200m로 정하면 총 140개의 구간으로 분할된다. 즉 기존의 균일간격에 의해 분할하는 방법(A), 200m마다 가상 사고건수 140개를 배치한 K -평균법을 이용한 방법(B)에 대해 분석하여 비교하고자 한다. 또한, K -평균법 중 군집수를 달리 적용할 때의 결과를 비교하고자 C방법과 같이 전체 이정을 300m로 나누었을 때의 값을 군집수로 적용하여 분석해 보고자 한다.

3.3 분석 결과

K -평균법은 통계 패키지인 SPSS를 이용하여 분석하였으며, 아래의 표 2는 두 기법에 의해 분할된 사고분석구간에 대한 결과를 나타내고 있다.

표 2. 경부선 주요구간의 사고분석구간 선정결과 (사고건수 기준)

방법	분할 방법	분석구간	총이정 (km)	총연장 (km)	구간수 (군집수)	사고건수 (건)	사고포함 구간수 (군집수)
A	균일간격에 의한 분할기법	경부고속도로	397.5~425.5 (부산방향)	28	140	88	54
B	K -평균법을 이용한 분할기법 (군집수 140개 : 200m기준)	경부고속도로	397.5~425.5 (부산방향)	28	140	88	55
C	K -평균법을 이용한 분할기법 (군집수 94개 : 300m기준)	경부고속도로	397.5~425.5 (부산방향)	28	94	88	55

표 3. 경부고속도로 부산방향(397.5~425.5km)의 사고 잦은 구간 선정 결과

순위	균일간격에 의한 분할기법 (A방법)			K-평균법을 이용한 구간분할기법 (B방법)			K-평균법을 이용한 구간분할기법 (C방법)		
	구간 번호	구간이정	사고 건수	군집 번호	군집내 사고간의 이정	사고 건수	군집 번호	군집내 사고간의 이정	사고 건수
1	43	406.1~406.3	7	75	406	7	69	406~406.2	8
2	58	408.9~409.1	4	119	409.6~409.8	5	67	409.8~410	6
3	63	409.9~410.1	4	105	409	4	91	409~409.2	6
4	3	397.9~398.1	3	122	410	4	6	398.1~398.3	4
5	28	402.9~403.1	3	7	398.1~398.2	3	93	409.5~409.7	4
6	61	409.5~409.7	3	83	406.2~406.4	3	36	403~403.1	3
7	88	414.9~415.1	3	107	415~415.1	3	32	402.7~402.8	3
8	4	398.1~398.3	2	40	402.7	2	38	415~415.1	3
9	10	399.3~399.5	2	45	403	2	88	418.3~418.5	3
10	29	403.1~403.3	2	48	403.1~403.2	2	15	399.4~399.5	2
계			33			35			42

참고 : 1. 위의 사고기법 중 K-평균법에 제시한 각 구간별 사고건수는 가상 사고를 배제한 실제 사고건수만을 표시한 것임.
 2. B방법중 군집내 사고이정이 단일이정으로 표시된 것은, 그 주변에 사고가 발생하지 않거나 한 지점에 집중발생된 경우임.

전체적으로 사고가 포함된 구간수를 살펴보면, A 방법과 B방법이 비슷한 결과를 나타냈다. 다만 K-평균법에서는 전체적으로 사고간의 거리가 먼 경우에 하나의 군집으로 묶이지 않도록 가상의 사고를 배치 하였으므로, 이 같은 결과는 교통사고가 전 구간에 걸쳐 고르게 발생하고 있다는 점을 간접적으로 증명해 주고 있다.

또한 본 연구에서는 두 가지 기법을 적용하여 사고 건수를 기준으로, 사고 잦은 구간에 대한 순위를 선정하였는데 그 분석결과는 표 3과 같다.

사고건수를 기준으로 A, B방법에 의한 사고 잦은 구간 선정결과, 전체적으로 10위권내의 구간에 대한 사고건수의 합은 K-평균법을 이용한 방식이 2건의 사고를 더 포함하는 것으로 나타났다. 세부적으로 보면, K-평균법에 의한 2순위 선정구간은 200m내에서 발생한 5건의 사고가 하나의 군집으로 처리되었으나, 균일분할방식에서는 이 구간이 두개로 분리되어 6순위(교통사고 3건)에 그치고 있다. 또한 B방법에 의해 결정된 사고 잦은 구간 10위권내를 보면, 각 군집의 사고들은 최대 200m의 범위내에서 상호 군집

을 이루고 있어, 추후 사고원인 및 개선사업을 위한 분석에 유용하게 적용될 수 있음을 보여주고 있다.

또한 K-평균법 중 300m를 기준으로 하여 군집수를 94개로 낮추어서 분석한 결과(C방법)를 보면, B 방법보다 많은 사고를 포함하고 있으면서도 각 구간의 길이가 200m를 초과하지 않은 상대적으로 우수한 결과를 나타내었다.

이러한 결과들은 분석구간의 길이나 사고의 양 또는 분포에 따라 적정 군집수를 선택하는 것이 매우 중요하다는 것을 보여주고 있다. 즉 K-평균법 적용 시, 군집을 몇 개로 분류할 것인가, 가상의 사고건수를 어떤 간격에 따라 배치하느냐에 대한 심도있는 검토가 필요하다.

K-평균법에 의한 분석 결과는 인근의 사고를 하나의 분석구간으로 묶을 수 있는 효율적인 방식이며, 따라서 사고 잦은 구간 선정시에 효율적으로 활용할 수 있다. 또한 이러한 군집분석방식은 사고를 중심으로 한 slider length 적용기법의 한 방법론으로써 효율적인 교통사고 개선사업을 위해서 활용가능한 방법이다.



4. 결론

기존 사고구간 분석방법은 주로 도로이정에 기초하여 거리에 따라 균등하게 분할하고, 분할구간 길이 또한 통상 200m단위를 기준으로 사용하였다. 그러나 이러한 방식은 인근의 유사한 사고가 하나의 구간으로 처리되지 못하는 단점이 존재한다. 또한 최근 사용되고 있는 slider length 적용기법 또한 적절한 분석구간 길이에 대한 해답을 제시하지 못하고 있는 실정이다.

본 연구에서는 사고분석구간 선정을 위해 사고지점을 기반으로 한 분석방법인 slider length 적용기법을 활용하기 위해, 군집분석의 일종인 비계층적 방법인 K -평균법을 활용하였다. 이러한 K -평균법은 전체 사고를 K 개의 군집으로 분류하는데 있어, 각 군집의 중앙지점에 거기에 소속된 사고들이 가장 가깝게 위치하도록 군집을 이루게 함으로써, 하나의 집단으로 처리되도록 하는 기법이다.

K -평균법은 인근거리에 발생된 사고들이 최대한 하나의 집단으로 분류될 수 있도록 하며, 이는 사고분석의 효율성과 사고 잦은 구간 선정을 통한 경제적 개선사업을 수행하는데 있어서 매우 중요한 작업이다.

이 기법의 활용을 위해 사전에 군집수를 결정해야 하는 사용상의 제한조건을 극복하고자, 적정수의 가상 사고를 도로에 배정하여 전혀 인근하지 않은 사고들이 하나의 집단으로 처리되지 않도록 하였다.

모델의 검증을 위해, 경부고속도로 부산방향으로 서울~신갈구간의 총연장 25.6km에 대해서 200m 단위로 분석구간을 나누어 균일간격에 의한 분할방식과 K -평균법을 이용한 방식에 대한 상호 비교를 통해 그 효율성을 검증하였다. 분석 결과, K -평균법에 의한 사고구간 분할방법이 가까운 거리에서 발생한 사고들을 효율적으로 분할하여 집단화시키는 것으로 나타났다. 이러한 군집분석 방식은 보다 효율적으로 사고 잦은 구간이 선정되도록 하는데 적용가능하다.

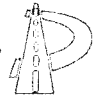
본 논문에서 제시한 K -평균법에 의한 분할방식을 더욱 더 효율적으로 활용하기 위해서는 향후 보완해야 할 몇 가지 과제가 있다.

첫째, 본 연구에서는 균등분할방식과의 동등한 조건을 비교를 위해 군집과 구간의 수를 140개로 균일하게 적용하였다. 따라서 가상 사고건수도 군집수에 맞추어 140개를 적용하였다. 그러나 C방법에서 보듯이, 전체구간의 사고밀도(km당 사고건수 등), 기하구조 특성에 따라 군집수 및 가상 사고는 적절하게 사용하는 것이 바람직하며, 다양한 케이스에 따라 적정 군집수를 결정하는 방법을 개발해야 한다.

둘째, 본 연구에서는 단순히 사고이정에 따라 가까운 거리에 있는 사고들의 집합화를 위한 방법론을 제시하였으나, 사고의 유사성은 거리외에 도로기하구조의 특성, 사고의 유형, 교통 특성등의 다양한 변수를 고려하는 것이 타당하며 이러한 종합적인 분석을 통해 분석구간을 결정하는 방식을 지속적으로 개발해야 할 것이다.

참고문헌

1. 강민욱, 손봉수, 도철웅, 강정규, 고속도로 곡선구간에서의 사고예측모형 개발, 대한토목학회지 제22권 제6-D호, 2002
2. 한국건설기술연구원, 교통사고 다발지점 개선방안에 관한 연구조사 설계 및 도로안전시설 설치편람 작성용역, 1989
3. 김정현, 이수범, 박병정, 교통사고 잦은 지점 및 구간 선정방법 개선에 관한 연구, 교통개발연구원, 2002
4. 김태완, 도로의 기하구조 요소가 교통사고에 미치는 영향 - 경부고속도로를 중심으로, 서울대학교 석사학위논문, 1995
5. 도로교통안전관리공단(2001) 교통사고 잦은 곳 기본 개선계획 및 효과분석
6. 유경수, 이기영, 교통사고 취약구간 개선 종합대책연구, 한국도로공사 연구소, 1997
7. 윤소식, 호주의 Black Spot(사고많은 곳)와 우리나라



- 라 적용방안, 신호등, 2001
8. 한국도로공사, 경부고속도로 사고통계 내부자료, 2002~2004
 9. Bhagman Persaud, Richard A. Retting, and Craig Lyon, Guidelines for Identification of Hazardous Highway Curves, *Transportation Research Record 1717*, Paper No. 00-1685, 2000
 10. Kiyoung Lee, Yong-taeck Lee, Myungsoon Chang, The Evaluation of the Methods Identifying Accident Blackspot on Korean Highway, *World Transport Research (WCTR9 Selected Proceedings from the 9th World Conference on Transport Research)*, Elsevier ltd (CD-ROM), 2003
 11. Shaw-Pin Miaou, The relationship between truck accidents and geometric design of road sections, *Accident Analysis and Prevention*, Vol.25, No6, 1994

〈접수 : 2005. 11. 29〉