INTEGRATION OF SAFETY SCORING AND CLUSTER ANALYSIS TO STUDY THE RELATIONSHIP BETWEEN ROAD GEOMETRY AND ROAD TRAFFIC CRASHES

Shibaura Institute of Technology Ph.D. Student Student Member Nitesh Acharya
Shibaura Institute of Technology Associate Professor Regular Member Michael Henry

1. INTRODUCTION

Each year approximately 1.3 million people die in road traffic crashes and around 93% of those crashes and death are in low- and middle-income countries [1]. Although there are several factors contributing for a road crash to occur, poor road geometry is one of the major causes. Generally, most of the roads in developing countries are not designed and engineered properly and as a result, road crashes occur. It is a huge challenge to identify deficiencies in road geometry and in addition, rectifying problems related to road geometry requires huge investment.

Many studies have been conducted to explore relationship between road geometry and road crashes and its outcomes suggest on proper identification of road crash locations as a preliminary task. After the crash spots have been identified and assessed in detail, various techniques can be used to group incident spots for collective management and selection of countermeasures. In this paper, two analytical tools of Safety scoring and Cluster analysis, is integrated to identify the similarities between road segments based upon safety scores of their road geometric features and the relationship between the cluster characteristics with road crashes is explored. This proactive approach will help road agencies and decisionmakers to plan their interventions and future investments.

2. METHODOLOGY

2.1. Field inspection and data collection

A 74 kms stretch of rural road of Nepal was considered for the study and the whole stretch was divided into 74 segments of 1 kms each. Five geometric features which are understood to contribute to road crashes were considered. These features were Carriageway width (V1), Shoulder width (V2), Road gradient (V3), Sight distance (V4) and Presence of access (V5). In May 2013, visual inspection was carried out by driving from start to end of each road segments and scores were assigned for each geometric feature. The road network had varying carriageway width of 3.80 to 7.00 meters, most of the sections did not have shoulder and the longitudinal gradient varied from 2% to 14%. Since the road passed through hilly terrain, there were numerous sharp curves and few property accesses. To address these variations in road features and to assess its impact accordingly, the ordinal scoring scale of 1, 2 and 3 was adopted, where "1" reflects low risk, "2" denotes medium risk and "3" represents high risk. Since, assigning the scores based on visual assessment is very subjective, a set of criteria as a guideline was developed using published domestic and international literatures [2]. The crash data record for three years was obtained from archive of Nepal police.

2.2. Data Analysis

The data set was a 74 × 5 matrix which comprised of safety scores for five road geometry features of each 74 segments of roads. To group road segments demonstrating similar road geometric features, Data clustering technique was used. The clustering was done using the open-source statistical analysis software R.

Since number of clusters in the dataset was not known, Hierarchical clustering method with Agglomerative approach was used. This approach is good at identifying small clusters (treats each observation as a cluster) and clusters are formed by grouping together observations or clusters with similar characteristics [3]. Euclidean distance was measured between each observation and distance matrix was formed. Using this matrix and Mcquitty type linkage, a dendrogram was plotted. Considering the shape of the diagram, 9 unique clusters were identified, and are shown in Figure 1. These clusters have been formed by combining road segments having highly similar road geometric features.

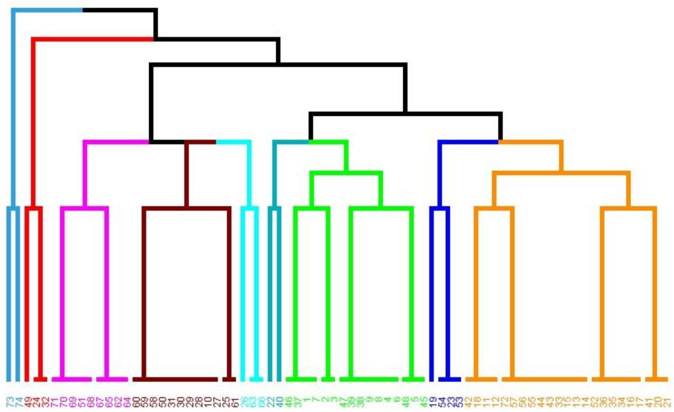

Figure 1: Dendrogram of road sections by Safety scores

3. RESULTS AND DISCUSSION

Table 1 shows the mean and standard deviation values of the safety scores for each geometric feature for all 9 clusters. From the results, it can be concluded that Shoulder width (V2) is one of the most critical and

Keywords: road geometry, road traffic crash, safety scores, cluster analysis
Address: Toyosu 3-7-5, Koto-ku, Tokyo, Japan 135-8548; TEL +81-03-5859-8363; email: na21504@shibaura-it.ac.jp

© Japan Society of Civil Engineers

dominant road features in all clusters with score of 3.0, except in Cluster 1 and 2.

The next objective of the study was to explore the relationship between the formed clusters and road crashes recorded in that clusters. The average number of road crashes occurring per kilometer of road stretch is shown in Figure 2 below.

Table 1: Mean and Standard Deviation of Safety scores for each road geometry features

		V1	V2	V3	V4	V5
Cluster 1	Mean	2.00	2.00	1.00	3.00	2.50
n= 2	S.D.	0.00	0.00	0.00	0.00	0.71
Cluster 2	Mean	1.67	2.00	2.00	2.00	1.00
n= 3	S.D.	0.58	0.00	0.00	0.00	0.00
Cluster 3	Mean	2.44	3.00	1.00	3.00	1.00
n= 9	S.D.	0.53	0.00	0.00	0.00	0.00
Cluster 4	Mean	2.00	3.00	1.00	2.00	1.17
n= 12	S.D.	0.00	0.00	0.00	0.00	0.39
Cluster 5	Mean	3.00	3.00	1.00	2.67	2.00
n= 3	S.D.	0.00	0.00	0.00	0.58	0.00
Cluster 6	Mean	3.00	3.00	2.50	2.00	1.00
n= 2	S.D.	0.00	0.00	0.71	0.00	0.00
Cluster 7	Mean	2.00	3.00	2.38	2.00	1.31
n= 16	S.D.	0.00	0.00	0.50	0.00	0.48
Cluster 8	Mean	2.25	3.00	2.00	3.00	2.00
n= 4	S.D.	0.50	0.00	0.00	0.00	0.00
Cluster 9	Mean	2.30	3.00	2.39	3.00	1.00
n= 23	S.D.	0.47	0.00	0.50	0.00	0.00

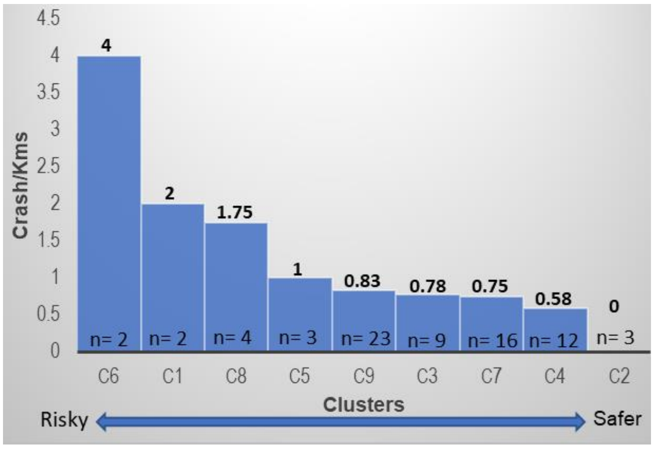

Figure 2: Crash rate in each cluster

In the histogram above, different clusters have different crash rates i.e., there exists clusters with very high average number of crashes and cluster having no crashes at all. The road segments in Cluster 6 have recorded 4 crashes per kilometer and are very risky and vulnerable to future crashes. In contrast, road sections in Cluster 2 are relatively safer without any crash records.

Different conclusions can be drawn after analyzing together the results of mean Safety scores and crash rates. For example, Cluster 2 has no road crashes at all and the mean Safety scores for each road features are 2 or less. This exhibits relationship between safety condition of road stretches in the clusters and the safety performance (crash rates) in those clusters. Similarly, Cluster 6 has very high crash rate of 4 crashes/kilometers and the mean safety scores for the features V1, V2 and V3 is also high. This implies that these geometric features (carriageway width, shoulder width and gradient) could be contributing factors for the crashes occurring in those segments. Furthermore, both the cluster C5 and C6 have same score of 3.00 for features V1 and V2, and both V4 and V5 scores for Cluster 5 are higher than of Cluster 6; however, Cluster 6 has a higher crash rate. Therefore, this might suggest that the effect of V3 (gradient) can be significant in the roads with narrow carriageway and shoulder.

Improving these geometric features may reduce road crashes in those road stretches. However, it is very challenging and requires huge investment, hence the impact of these features can be reduced using countermeasures like regulating speed, traffic calming measures, installation of safety barriers, etc.

4. CONCLUSIONS

Using the above techniques, clustering of road segments having similar geometric features and safety performance was done. In this way, 74 different types of road segments were reduced into 9 types, by forming the clusters. These results will help the road agency to plan their interventions accordingly as the segments within a cluster can be treated with same countermeasures. The results can also be used for long-term major safety improvement programs like widening of roadway and major rehabilitation works. Furthermore, the approach of using Safety scores for evaluating the roadway condition is proactive and integrating this approach with Cluster analysis can be an easy, quick, and helpful approach to identify unsafe road sections.

ACKNOWLEDGEMENT

This research was supported by a scholarship for road asset management from the Japan International Cooperation Agency.

REFERENCES

[1] World Health Organization (2018), “Global Status Report on Road Safety 2018
[2] Department of Roads, Government of Nepal (2013), “Nepal Road Standards 2070”
[3] A. Kassambara (2017), “Multivariate Analysis I-Practical Guide to Cluster Analysis in R, Unsupervised Machine Learning”

© Japan Society of Civil Engineers