A STATISTICAL ANALYSIS OF VEHICLE PEDESTRIAN ACCIDENTS THAT MAKE USE OF PEDESTRIAN CRASH WARNING INFORMATION

Toyohashi University of Technology, Student Member, Franziska Miksch
Toyohashi University of Technology, JSCE Member, Kojiro Matsuo
Toyohashi University of Technology, JSCE Member, Nao Sugiki

1. Introduction

One important element to examine pedestrian safety at each location is “exposure”: how many pedestrians are exposed to the accident risk at the location[1)]. The most accurate measure for evaluating the pedestrian exposure would be to count the number of pedestrians at every location. It is not possible in a reasonable time due to the high number of locations. On the other hand, recent technologies in-vehicles have allowed in-vehicle *pedestrian crash warning* (PCW) systems. PCW information must be including the information of the presence of pedestrian at the location, which would make it possible to estimate the pedestrian exposure. The goal of this study is to make a statistical model of vehicle-pedestrian accident frequencies at intersections by considering the PCW frequencies as well as vehicle traffic volumes.

2. Analysis

In this study, we focus on vehicle-pedestrian accidents at intersections on community streets. For the analysis, different types of data were used which were needed to create an exact model as possible.

At first the vehicle-pedestrian accidents in Toyohashi from 2008 to 2015 was used, which is an amount of 1465 accidents (5.68% of total accidents). A vehicle-pedestrian accident that happened in a radius of 15m around the center of an intersection was defined as an intersection accident. Then, the number of accidents at each intersection is counted.

Fig.1 PCW locations in a community street area

General vehicle probe data (corrected by Pioneer Corp.) for a year in 2013 is map-matched and used to evaluate the relative amount of each link’s traffic volume. Then, traffic volume for each intersection is calculated as the total of traffic volume of links that connect to the intersection.

For PCW information, data for a year (from Sep. 2017 to Aug. 2018) of 39 normal size vehicles of a transport service company was used. The vehicles are equipped with “Mobileye 570” (provided by Mobileye Corp.) that is an aftermarket *advanced driver assistant system* (ADAS) based on a single front camera and can make several types of warnings including PCW. The PCW is occurred when the vehicle is evaluated to be going to crash in two seconds. The targeted vehicles are also equipped with an internet communication system that can send the probe data (e.g. latitude, longitude, vehicle speed, and direction) to the data server by around 10 seconds as well as when a warning occurred. **Fig.1** shows the exact locations that PCW occurred in a community street area. The probe data was map-matched, and the number of these vehicles passing through each intersection was aggregated. The frequency of PCW in a radius of 30m around the center of each intersection was also calculated. Then, PCW incidence rate for each intersection was calculated as the frequency of PCW divided by the number of the vehicles passing through. Additionally, some road and area environment variables such as the number of legs and land use were used.

For the statistical analysis, the Negative Binomial Regression Model was applied. The probability can be calculated as follows:

$$P_{NB}(Y_i = y|\lambda_i, \phi) = \binom{y+\phi+1}{\phi-1}\left(\frac{\phi}{\phi+\lambda_i}\right)^{\phi}\left(\frac{\lambda_i}{\phi+\lambda_i}\right)^{y}, \qquad \lambda_i = \exp(\boldsymbol{x}_i^t\beta),$$

where, Y_i is the number of vehicle-pedestrian accidents at an intersection i during the study period; $\boldsymbol{x}_i^t$ is the predictor variables vector for the intersection i; β is the coefficient parameters vector; and ϕ is the dispersion parameter. Additionally, the expectation and the variance are:

$$E_{NB}(Y_i|\lambda_i, \phi) = \lambda_i, \qquad V_{NB}(Y_i|\lambda_i, \phi) = \lambda_i + \frac{\lambda_i^2}{\phi}$$

These parameters were estimated by MLE.

The estimated final model can be seen in **Table 1.** The small dispersion parameter indicates that the variance of the accident count is higher than the expectation, suggesting that this Negative Binomial model is better than the Poisson model. The first variable that is reviewed is the log nature of general probe traffic volume which describes the trip counts at every intersection. The coefficient value show that the higher the general traffic volume is the more accidents tend to happen at an intersection. This relationship is easy to understand and the low value for the standard error even strengthens this outcome. The second variable contains the amount of legs that every intersection has. The lowest value for it is three. The positive value of the coefficient shows that at a crossing with more legs also more accidents happen. Like the first variable the positive relationship is quite intuitive because the amount of conflict points grows between a three-legged and a four-legged intersection. Additionally, the land use of the area where the intersections are located, is analyzed. Three different types of land uses were found to be statistically significant. The highest correlation showed areas with high rise buildings. That means areas with residential buildings with four or more floors as also buildings with commercial use. The lowest corresponding parameter was calculated in areas with a low density of population and low-rise buildings which means less than three floors. In between the value for low rise buildings with a high density in population was estimated. These results show that in areas with high rise buildings the most accidents are happening. Areas with low rise buildings, but higher population density are more endangered for pedestrian accidents than areas with lower population density. The last variable, which is called PCW incidence rate sets the count of the pedestrian crash warnings in relation to the number of equipped vehicles' trips. In this case also a positive relation can be seen what means that the higher the pedestrian flow is at an intersection the more accidents will happen. Noticeable is the high value for β which shows an especially high relation between the pedestrian flow and the crash warnings. This results also indicates that the expectation of the frequency of pedestrian-vehicle crash increases by $\exp(3.31 * 0.01) - 1.0 = 0.34$ (or 3.4%) as the PCW incidence rate increases by 1%.

Table 1 Estimated results of the negative binomial model

Predictor variable	Coefficient	p-value
Intercept	-8.95	<< 0.001
Log nature of general probe traffic volume	0.515	<< 0.001
Number of legs	0.598	<< 0.001
Land use high rise building	1.55	<< 0.001
Land use low rise and low density residential	0.500	0.018
Land use low rise and high density residential	1.08	<< 0.001
PCW incidence rate	3.31	<< 0.001
Dispersion parameter	0.446	
Sample size	5462	
McFadden's ρ^2	0.15	
AIC	1804	

3. Conclusion

All in all, the model helped to understand more deeply the relationship between pedestrian crash warnings and vehicle-pedestrian accidents. The estimated results of the model showed that all analyzed variables are statistically significant. Especially the different results in the areas with different kind of land use show that the surroundings of the intersection have an impact on the occurrence of vehicle-pedestrian accidents. In the future the model can be expanded by introducing a fitting model for the PCW and derive more relations by examining the deviance. Additionally, only data from a target area can be analyzed which gives the possibility to examine some intersections in detail.

REFERENCES

1. Lee J. Abdel-Aty, M. and Shah I.: Evaluation of surrogate measures for pedestrian trips at intersections and crash modeling, Accident Analysis and Prevention, May, 2018 (In Press).