様　式　Ｃ－１９、Ｆ－１９、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

平成　２８　年　　６　月　２２　日現在

機関番号：　１２６０１
研究種目：　挑戦的萌芽研究
研究期間：　2014～2015
課題番号：　２６５４０１６６

研究課題名（和文）Large-Scale Location Mining for Multi-dimensional Risk Factors Discovery: A study for Road Accident Prevention

研究課題名（英文）Large-Scale Location Mining for Multi-dimensional Risk Factors Discovery: A study for Road Accident Prevention

研究代表者

ホラノン　ティーラユット（Horanont, Teerayut）

東京大学・空間情報科学研究センター・客員研究員

研究者番号：００５９８２８７

交付決定額（研究期間全体）：（直接経費）　　2,900,000円

研究成果の概要（和文）：衝突事故は様々な要素の組み合わせによって引き起こされる。例えばある要因は衝突事故を誘発しそれが追突の原因となることもある。ある要因が衝突事故の衝撃を増長し、事故による後遺症を悪化させることもある。幾つかの要因は事故を誘発する直接的な要因になることもあり、またそれが中長期的に事故を起こす構造的な要因となる事もある。交通事故を引き起こすリスク要因を特定することは、交通事故リスクを低減するために重要である。本研究の主な成果は、交通事故の発生に影響をおよぼす要因を特定したことである。それらの要因を明らかにしたことで道路交通や衝突防止システムの安全性向上に貢献することが期待される。

研究成果の概要（英文）：A road traffic crash results from a combination of factors related to the components of the system comprising roads, the environment, vehicles and road users, and the way they interact. Some factors contribute to the occurrence of a collision and are therefore part of crash causation. Other factors aggravate the effects of the collision and thus contribute to trauma severity. Some factors may not appear to be directly related to road traffic injuries. Some causes are immediate, but they may be underpinned by medium-term and long-term structural causes. Identifying the risk factors that contribute to road traffic accidents is important in identifying interventions that can reduce the risks associated with those factors. The main finding of this study is the nfluential factors in traffic accidents.
Those independent variables are found to be influential and thus the
results can help increases the safety of vehicle collision avoidance system.

研究分野：　交通計画

キーワード：　Road traffic accident　Traffic volume　Risk factors　Spatial correlation　Modelling

１．研究開始当初の背景

A road traffic crash results from a combination of factors related to the components of the system comprising roads, the environment, vehicles and road users, and the way they interact. Some factors contribute to the occurrence of a collision and are therefore part of crash causation. Other factors aggravate the effects of the collision and thus contribute to trauma severity. Some factors may not appear to be directly related to road traffic injuries. Some causes are immediate, but they may be underpinned by medium-term and long-term structural causes. Identifying the risk factors that contribute to road traffic accidents is important in identifying interventions that can reduce the risks associated with those factors, and thus increases the safety of vehicle collision avoidance system.

２．研究の目的

This proposal is devoted to construct a novel methodological framework to identify risk factors for road traffic accidents. Long-term human mobility data, weather condition and Geographic Information System (GIS) are combined to provide analysis. The results give new mean to understand the full complex causation information in car accident cases.

３．研究の方法

The proposed methodologies provide an exhaustive 'state-of-the-art' to retrieve road accident risk factors. Road and Environment features such as speed, total trip distance, place, day of week and so on are extracted from massive spatio-temporal dataset. These features will be learned to estimate the multidimensional risks of road accidents and therefore risk of crash, which may have occurred on the road segment during a defined time period.

The accident data were divided into two sets; one is those accidents occurred in weekdays and the other one is those occurred in weekends and holidays. The traffic accidents data, land use data and road characteristics were imported to QGIS along with the 1km by 1km mesh data in order to calculate the values explanatory variables mesh by mesh. The basic statistics of objective and explanatory variables are summarized in Table 1. The candidate explanatory variables will be selected using an R function for model selection in the model calibration process.

3.1 Zero-inflated Negative Binomial Regression

A zero-inflated model has a mixed distribution function of a binary distribution expressing the probability of excessive zero when the count of accidents is zero and a count distribution such as the Poisson or negative binomial distribution which represents the probability of accidents count greater than zero. Mathematical expressions are shown below. A zero-inflated negative binomial (ZINB) regression model for a count random variable Y_i are shown in Eq. (1).

$$P_{NB}(Y_i = 0) = \omega_i + (1-\omega_i)\left(1+\frac{\lambda_i}{\tau}\right)^{-\tau}$$

$$P_{NB}(Y_i = y_i) = (1-\omega_i)\left(\frac{\Gamma(y_i+\tau)}{y_i!\,\Gamma(\tau)}\right)\left(1+\frac{\lambda_i}{\tau}\right)^{-\tau}\left(1+\frac{\tau}{\lambda_i}\right)^{-y_i},\ y_i = 1,2,\ldots \quad (1)$$

where ω_i is a zero-inflation factor and i is the mesh number.

3.2 Model Calibration

The correlations coefficients among variables have been conducted and reviewed using R programming. The independent variables with high correlation coefficients with the dependent variables- number of traffic accidents on workday and number of traffic accidents on weekend and holiday - have been chosen as candidate explanatory variable. However, independent variables with high correlation coefficients with each other are avoided to prevent the multi-colinearity. Thus, only one will be chosen among them, generally the one with higher correlation coefficient with the dependent variable. The correlation coefficients among variables are presented, for example, in Figure 1 and 2.

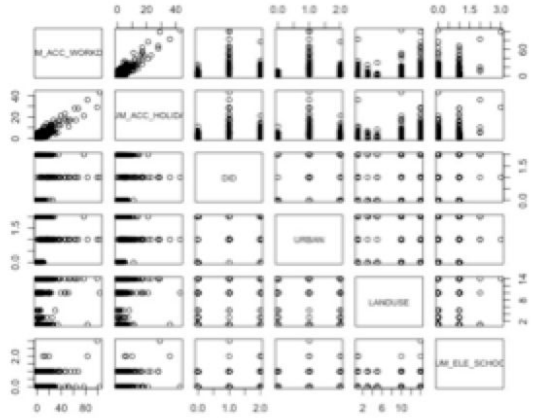

Fig. 1. Correlation coefficients of land use variables

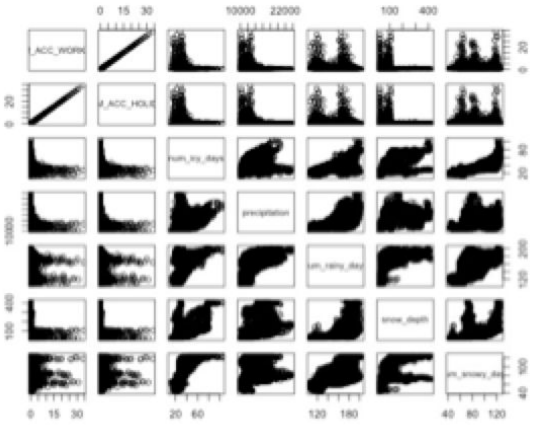

Fig. 2. Correlation coefficients of climate variables

The candidate variables are then included in the model one by one while monitoring the AICc value. The models with lowest AICc value and most significant variables are chosen, prefecture by prefecture.

4．研究成果

4.1 Results and discussions

The results of model estimates for Hokkaido and Tohoku region, Kanto region, and Chubu region are calculated for traffic accident on weekend and holiday as well as the traffic accident on workday. Sets of independent variables are varied from prefecture to prefecture due to the differences in prefectural characteristic which affect traffic accidents; however, independent variables found to be significant for at least one prefecture are number of shops, number of elementary schools, number of sport facilities, number of intersections, number of people commuting to work by train, automobile, bicycle, and on foot, number of rainy days, and number of snowy days for traffic accident on workday count component, and urban area, and total length of roads for zero-inflated component.

4.2. Conclusions and future work

The main finding of this study is the influential factors in traffic accidents. Those independent variables are found to be influential; that is, number of shops, number of elementary schools, number of sport facilities, number of intersections, number of people commuting to work by train, automobile, bicycle, and on foot, number of rainy days, number of snowy days, urban area, and total length of roads are statistically significant in the traffic accidents models on workday; and number of shops, number of department stores, number of intersections, number of residents, number of rainy days, number of snowy days, urban area, and total length of roads are significant in the traffic accidents models on weekend and holiday.

After the identifying factors among land use, climate, road and demographic variables, it is necessary to develop models that consider temporal changes. Land use factors, which rarely change in short period of time, and demographic factors, which change also in a longer time period should be considered. In addition, it is necessary to include factors which change daily or in a shorter time period using real time cli-mate data and road related data such as traffic volume and speed to alert users when they enter high risk zones.

By developing this approach, it could improve the road safety and reduce traffic accidents; it is anticipated that heavy or fatal accidents will decrease and that awareness of road users will be raised at all time and all places.

5．主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計　0　件）

〔学会発表〕（計　2　件）

1. Paweenuch Songpatanasilp; Harutoshi Yamada; Teerayut Horanont; and Ryosuke Shibasaki

(2015). Traffic accidents risk analysis based on road and land use factors using GLMs and zero-inflated models. In Proceedings of 14th International Conference on Computers in Urban Planning and Urban Management (CUPUM 2015), 7-10 July 2015, Massachusetts, USA, pp. 320-1 - 320-26.

2. Paweenuch Songpatanasilp; Teerayut Horanont; Harutoshi Yamada; and Ryosuke Shibasaki (2015). Modeling traffic accidents occurrences based on land use and road factors using geographically weighted regression models. In Proceedings of the 10th International Conference on Knowledge Information and Creativity Support Systems (KICSS 2015) [CD-ROM], 12-14 November 2015, Phuket, Thailand, pp. 481-492.

〔図書〕（計 0 件）

〔産業財産権〕
○出願状況（計 0 件）

○取得状況（計 0 件）

〔その他〕
ホームページ等
http://shiba.iis.u-tokyo.ac.jp/

６．研究組織
(1)研究代表者
ティーラユット ホラノン（TEERAYUT HORANONT）
東京大学・空間情報科学研究センター・客員研究員
研究者番号：00598287

(2)研究分担者
関本 義秀（SEKIMOTO YOSHIHIDE）
東京大学・生産技術研究所・准教授
研究者番号：60356087

(3)連携研究者
( )

研究者番号：