13th COTA International Conference of Transportation Professionals (CICTP 2013)

Estimates of pedestrian crossing delay based on multiple linear regression and application
LI Dan\textsuperscript{a}, Xiaofa Shi \textsuperscript{a}*

\textsuperscript{a} Key Laboratory of Road and Traffic Engineering of the State Ministry of Education, Tongji University, Shanghai, 201804, China

Abstract

Within the range of city traffic, a large number of intersections are out of control of traffic light, so the pedestrian delay and car delay, which are caused by the conflict between crossing pedestrian and right-turn cars, are very common. This article, from the pedestrian perspective, gives an analysis for the influence to pedestrian crossing delay made by right-turn cars and models the relationship between independent variables and dependent variable by fitting a linear equation to the observed data with multiple linear regression model. The independent variables includes right-turn car volume, speed, the number of pedestrians on the opposite side and the number of pedestrians on the same side. The dependent variable is pedestrian crossing delay. Then, parameters calibration is made through SPSS by stepwise regression. After the analysis of impact factors of pedestrian crossing delay, this paper gives a restrictive condition for setting exclusive right-turn phase, as when the car volume exceeds a threshold, aiming to provide certain theory basis for the management of city traffic and improve the intersection traffic capacity and service level.

Keywords: multiple linear regression; pedestrian crossing delay; right-turn car flow; the number of pedestrians;

1. Introduction

Delay is loss of travel time due to traffic congestion and traffic control(s/min). It is an indicator for evaluating the level of service of a section of road or an intersection. The major component of urban traffic delays are caused by the intersections. So many scholars have done a great deal of research on intersection delay and provided a large number of theoretical and technical supports for urban traffic management and control. Within the range of city traffic, a large number of intersections are out of the control of traffic light, so the pedestrian delay and car delay, which is caused by the conflict between pedestrian crossing and right-turn cars, is very common. Especially near the shopping malls, schools, hospitals and other sites where there is a large flow of

* Xiaofa SHI. Tel.: +86-13816331510; 
E-mail address: xfshi999@163.com;
people and more serious conflicts. This situation not only brings security risks, but also reduces the efficiency of intersections.

In the case involving a right-turn traffic and pedestrian conflict, most of the literature mainly focuses on the delay of the vehicle and ignores the pedestrian delay. Under the conditions without right-turn signal control, this paper starting from the pedestrian perspective, establishes a multiple linear regression on the relationship between the pedestrian crossing delay and right-turn car volume, speed, the number of pedestrians on the opposite side and the number of pedestrians on the same side, and then analyses the factors of pedestrian crossing delay. At last this paper presents some quantized suggestions for right-turn traffic control.

2. Literature review

The phenomenon of pedestrian and right-turn vehicle conflicting is very normal around the country. There are a rich of previous studies in the area of urban traffic. Li Qingfeng and Wang Zhao’an(2006)estimated pedestrian crossing delay at the intersections through; Gao Liping and Chen Shaokuan(2009) analyzed pedestrian crossing delay in different conditions through simulation and statistical data; Yang (2005) took some intersections in xi’an for instance, established Monte Carlo model to calculate pedestrian delay, this model divided intersection signal phase into a series of 1s-length sub-phase, simulated pedestrian in signal intersection and calculated each sub-phase pedestrian delay and total delay; Yang Xiaoguang(2005,2007)applied crossing theory to infer the turning vehicle capacity model and pedestrian delay model under the condition of pedestrian crossing the street, compared the intersection’s capacity and pedestrian delay before and after setting exclusive right-turn phase and established widely-used pedestrian crossing delay model in a part of road with no signal control to quantificationally compare and analyze pedestrian delay for once-crossing and twice-crossing in roads with different width and volume.

The above scholars from different angles analyzed delays to pedestrian, but did not get to the bottom of it. In this paper, there is a strong practical significance to take the right turn traffic characteristics into account.

3. Data sources

This article data is from the survey data at Cao’an Road and Moyu Road in Shanghai. The data only concludes the west-import sidewalk data; pedestrian green time (30s), west imported Lane (4, one right lane, one left lane and two through lanes), export lanes (3), non-motorized lane (1 for each side) and don’t consider traffic control, crossing facilities, road condition, weather and so on. Pedestrian crossing delay is the D-value of the actual travel time and the ideal travel time, measured according to the following formula:

\[
Y = t_r - t_i \\
t_r = t_a - t_f \\
t_i = l_i/v_i \\
d — pedestrian crossing delay \\
tr — the actual travel time \\
ti — the ideal travel time \\
ta — the arrival time at the crosswalk \\
tf — the finish time of crossing \\
l — the length of crosswalk \\
v_i — the pedestrian’s ideal speed
\]

According to “Traffic Engineering Manual”: \( v_i = 1.2 \text{m/s} \). The result of ideal travel time in this paper is 23s.
4. Modeling

4.1 Variable definition

The dependent variable is pedestrian crossing delay—delays resulted from the conflict between the right-turn traffic and crossing pedestrian.

The probable independent variables are as follows:
- The number of pedestrians on the opposite side (X1)—the number of people who want to across the road on the opposite side of the respondent;
- Right-turn volume (X2)—cars passing the crosswalk during the pedestrian green time (G);
- The number of pedestrians on the same side (X3)—the number of people who want to across the road on the same side of the respondent;
- Right-turn traffic speed (X4)—the speed (m/s).

4.2 Y-X scatter plot analysis

For the above four kinds of influencing factors, this paper uses scatter plots to determine whether its impact on the dependent variable is significant to decide whether it is an independent variable.

We can see from the above scatter plot analysis, right-turn car volume, the number of pedestrians on the opposite side and the number of pedestrians on the same side has a very obvious linear relationship with the pedestrians crossing delay with R² respectively 0.8182, 0.6729, 0.7525, 0.8411, while R² for right-turn speed is 0.059, so the effect for pedestrians crossing delay is not significant. Therefore, right-turn speed cannot be used as the independent variable in the multiple linear regression.
4.3 correlation test

The result of correlation analysis using SPSS is shown in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Pedestrian delay</th>
<th>right-turn volume</th>
<th>right-turn speed</th>
<th>the number of pedestrians on the same side</th>
<th>the number of pedestrians on the opposite side</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedestrian delay</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>right-turn volume</td>
<td>0.905</td>
<td>1</td>
<td>-0.250</td>
<td>0.867</td>
<td>0.917</td>
</tr>
<tr>
<td>right-turn speed</td>
<td>-0.250</td>
<td>-0.090</td>
<td>1</td>
<td>-0.228</td>
<td>-0.298</td>
</tr>
<tr>
<td>the number of pedestrians on the same side</td>
<td>0.867</td>
<td>0.814</td>
<td>-0.228</td>
<td>1</td>
<td>0.852</td>
</tr>
<tr>
<td>the number of pedestrians on the opposite side</td>
<td>0.917</td>
<td>0.903</td>
<td>-0.298</td>
<td>0.852</td>
<td>1</td>
</tr>
</tbody>
</table>

Seen from the Pearson correlation test, right-turn volume, the number of pedestrians on the opposite side and the number of pedestrians on the same side has a very obvious linear relationship with the pedestrians crossing delay while right-turn speed has little, which corresponds with the scatter plot analysis. So we exclude right-turn speed as independent variable.

4.4 Stepwise regression

In this paper, a stepwise regression method is used to model the linear correlation between Y (pedestrian delays) and all the independent variables. The method is including the independent variables in the regression equation in turn. All influencing factors were combined to establish the one linear or multiple linear regression equation and we get some equations. Compare the adjusted R2s and take the one which has the largest adjusted R2. Because we put the independent variables into the equation one by one, it can avoid including into less important variables and dropping the important ones.

Do a stepwise regression analysis using SPSS software fitting the obtained Y and (X1, X2, X3). The introducing order for independent variables to the line equation is successively the number of pedestrian on the opposite side, right-turn traffic flow, the number of pedestrian on the same side, and the results are as follows. See Table 2.

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>adjusted R Square</th>
<th>Std. error of the estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.917a</td>
<td>0.841</td>
<td>0.838</td>
<td>1.66685</td>
</tr>
<tr>
<td>2</td>
<td>0.934b</td>
<td>0.873</td>
<td>0.868</td>
<td>1.50428</td>
</tr>
<tr>
<td>3</td>
<td>0.943c</td>
<td>0.890</td>
<td>0.884</td>
<td>1.41086</td>
</tr>
</tbody>
</table>

a. predictive variables: (constant), the number of pedestrian on the opposite side.
b. predictive variables: (constant), the number of pedestrian on the opposite side, right-turn volume.
c. predictive variables: (constant), the number of pedestrian on the opposite side, right-turn volume, the number of pedestrian on the same side.
Seen from Table 2, we got three models. With the gradual increase of the independent variable, all models’ adjusted R2 are increasing. So we got the best model: Model 3, Adjusted R2=0.884, which means (X1, X2, X3) can explain an 88.4% change in Y. Therefore, the choice of multiple linear regression equation is Model3.

Table 3 Anova\(^a\)

<table>
<thead>
<tr>
<th>Model</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>regression</td>
<td>887.134</td>
<td>3</td>
<td>295.711</td>
<td>148.560</td>
<td>0.000*</td>
</tr>
<tr>
<td>residual</td>
<td>109.478</td>
<td>55</td>
<td>1.991</td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>996.613</td>
<td>58</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) predictive variables: (constant), the number of pedestrian on the opposite side.

b. predictive variables: (constant), the number of pedestrian on the opposite side, right-turn volume.

c. predictive variables: (constant), the number of pedestrian on the opposite side, right-turn volume, the number of pedestrian on the same side.

Seen from Table 3, conduct F-test, the value of significant level sig=0.000<0.05, so the coefficients in the regression equation are not zero. The assumption of zero coefficients is rejected, and equation is fitting well.

Table 4 Coefficients\(^a\)

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized coefficients</th>
<th>Standardized coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>the number of pedestrian on the opposite side</td>
<td>0.949</td>
<td>0.340</td>
<td>2.794</td>
<td>0.007</td>
</tr>
<tr>
<td>Right-turn volume</td>
<td>0.339</td>
<td>0.104</td>
<td>0.382</td>
<td>3.255</td>
</tr>
<tr>
<td>the number of pedestrian on the same side</td>
<td>0.383</td>
<td>0.116</td>
<td>0.351</td>
<td>3.308</td>
</tr>
<tr>
<td>3</td>
<td>0.249</td>
<td>0.084</td>
<td>0.256</td>
<td>2.943</td>
</tr>
</tbody>
</table>

\(^a\) dependent variable: pedestrian crossing delay

Establish a linear regression model based on each variable coefficient shown in Table 4:

\[
Y = 0.949 + 0.339X_1 + 0.383X_2 + 0.249X_3
\]

The values of significant level are 0.007, 0.002, 0.002, 0.005, 0.005, all beneath 0.05, meaning all Xs affect Y significantly.

4.5 Model analysis and discussion

Seen by the model, the coefficients on each variable for pedestrian delays are positive, indicating that they are positively correlated. With the right-turn volume, number of pedestrian increasing, pedestrian delay will increase, and in theory, will reach positive infinity. Right-turn volume affects pedestrian delay mostly. When right-turn
volume raises 1(pcu/T), the delay increases 0.383s. So right-turn traffic affects pedestrian delay more than pedestrian themselves and it’s the only chief factor in addition to signal control factors. Therefore, it is a sally port to reduce pedestrian crossing delay to control right-turn traffic.

The scope of application of the model: intersections setting right turn lanes without setting the right-turn traffic signal lights. The model is very effective to calculate pedestrian crossing delays in such kind of intersections. Compared to other non-linear model, its calculations are more quick and easy. And this model is very flexible to add or reduce the independent variables, focusing on different points. However, the model’s dependent variable—pedestrian crossing delay only includes part of the delay rather than the overall delay. If you want to know anything about the overall delay of pedestrians, you need to learn from some of the other mature models.

The model is not perfect. The coefficient of right-turn vehicles is 0.383 and the number of pedestrian on the opposite side 0.339, the difference between both is only 0.044s, which means almost no difference. From a common sense point of view, it is unnormal that a car or a person has the same effect on the pedestrian crossing delay. The reason for this phenomenon may be the small number of samples or the peak hours. The coefficient of the number of pedestrian on the same side is 0.249, and differs from the number of pedestrian on the opposite side by 0.1, which means the number of pedestrian on the same side affect the delay less than the number of pedestrian on the opposite side. However, in actual observations, when the number of pedestrian crossing the street is very small, the effect on pedestrian crossing delay by the pedestrian on the same side is very little. The effect by the pedestrian on the same side increases sharply along with the increase of the number of pedestrian crossing the street. So the effect on pedestrian crossing delay by the pedestrian on the same side is affected by the number of pedestrian on the opposite side. Through quantitative analysis of survey data, due to the colinearity between the number of pedestrians on the same side and the number of pedestrians on the opposite side is in the acceptable range, this paper assumes no colinearity.

4.6 Recommendations for right-turn traffic signal control

According to the model, for a particular intersection, set right-turn traffic signal control from the point of view of pedestrian delay according to the right-turn traffic volume.

Seeking critical right-turn traffic volume, assume:
(1) The length of the crosswalk pedestrian green light for G;
(2) The pedestrian maximum tolerable waiting time is T;
(3) \(X_3 = 1, X_1 = 0;\)
(4) \(Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3\) (for different intersections, the coefficient \(\beta\) can be obtained according to above method)

Also \(Y \leq G\) then get \(\beta_0 + \beta_1X_1 + \beta_2X_2 \leq G\), then get \(X \leq \min\{\frac{G - \beta_0 - \beta_1}{\beta_2}, \frac{T - \beta_0 - \beta_1}{\beta_2}\}\)

When \(X\) gets \(X \leq \min\{\frac{G - \beta_0 - \beta_1}{\beta_2}, \frac{T - \beta_0 - \beta_1}{\beta_2}\}\), the intersection need to set the right turn signal lights control.

5 Conclusions

Around domestic intersections, not only pedestrian crossing delays are deeply affected by the right turn traffic, but also intersection traffic safety is reduced, on the other hand, pedestrian delay and the crowd itself have certain correlation. In order to study the specific relationship between pedestrian crossing delay and right-turn traffic and crowd itself, based on multiple linear regression model, this paper analysed its influencing factors and presented
quantitative recommendations for right-turn car signal control in urban road intersections and provided certain theoretical support for traffic management and control.

References

NIU Zhiqiang, QIAN Dalin and WANG Xiaoxiao (2011), intersection pedestrian signal delay model of qualitative reasoning. Transportation Engineering and Information Technology.
GAO Liping (2009), Pedestrian-motor vehicle delay model based on the crosswalk location. Jilin University (Engineering Science).
XIONG Wen, CHEN Xiaohong, HU Xianbiao (2009), The study of pedestrian crossing facilities time and space threshold on urban trunk road, Urban Transport.
DAI Tongyu, Du Renbing, PEI Yulong (2008), Simulation of pedestrian crossing delays at intersections of city roadway, Computer and Communications.
FENG Shuming, PEI Yulong (2007), The study of pedestrian crossing delays. Harbin Institute of Technology.