Relating Freeway Traffic Accidents to Inductive Loop Detector Data Using Logistic Regression

Junhyeong Park, Graduate Student Researcher
Cheol Oh, Ph.D. Assistant Professor
Department of Transportation Systems Engineering
Hanyang University at Ansan
Contents

1. Introduction
2. Data Preparation
3. Methodology and Model Development
4. Application and Technical Issues
5. Conclusion
1. Introduction

Background

- Behavior/Aggressiveness
- Driving Skill/Experience
- Psycho-physiological loads

Vehicle Characteristics
- Malfunctioning components

Driver Characteristics

Traffic Dynamics
- Flow
- Occupancy
- Speed

Environment
- Weather Conditions
- Geometric Design elements

Accident
1. Introduction

Background

Freeway Traffic Surveillance System

Real-time traffic data collection

Detection of disruptive traffic condition

Analysis of traffic variables in abnormal traffic condition

Traffic accident indicator

Accident occurs

Implication starts

Normal traffic condition

Disruptive traffic condition

Time

T

T-x
1. Introduction

Data

- Seohae Freeway: Seoul – Mokpo (339.51km)
- Seohaeen freeway loop data from detectors and traffic accident data for 3 years – 2004, 2005 and 2006
  - Occupancy, Speed, Volume

Traffic accident data

- Traffic accident data from Seohaeen freeway for three years (2004~2006)
  - Date, Time, Location, Postmile, Direction, Type

Loop data

- Loop data from Seohaeen freeway for three years (2004~2006)
  - Extraction of traffic data from accident-involved loop data

Derivation of explanatory variables for traffic conditions
# 2. Data Preparation

Conceptual illustration for loop detector data processing

<table>
<thead>
<tr>
<th>Space</th>
<th>Time</th>
<th>Upstream Loop 2</th>
<th>Upstream Loop 1</th>
<th>Downstream Loop 2</th>
<th>Downstream Loop 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(t-15) ~ (t-30)</td>
<td>up2_t2_O, up2_t2_DO</td>
<td>up1_t2_O, up1_t2_DO</td>
<td>dn1_t2_O, dn1_t2_DO</td>
<td>dn2_t2_O, dn2_t2_DO</td>
</tr>
<tr>
<td></td>
<td></td>
<td>up2_t2_V, up2_t2_DV</td>
<td>up1_t2_V, up1_t2_DV</td>
<td>dn1_t2_V, dn1_t2_DV</td>
<td>dn2_t2_V, dn2_t2_DV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>up2_t2_S, up2_t2_DS</td>
<td>up1_t2_S, up1_t2_DS</td>
<td>dn1_t2_S, dn1_t2_DS</td>
<td>dn2_t2_S, dn2_t2_DS</td>
</tr>
<tr>
<td></td>
<td>(t) ~ (t-15)</td>
<td>up2_t1_O, up2_t1_DO</td>
<td>up1_t1_O, up1_t1_DO</td>
<td>dn1_t1_O, dn1_t1_DO</td>
<td>dn2_t1_O, dn2_t1_DO</td>
</tr>
<tr>
<td></td>
<td></td>
<td>up2_t1_V, up2_t1_DV</td>
<td>up1_t1_V, up1_t1_DV</td>
<td>dn1_t1_V, dn1_t1_DV</td>
<td>dn2_t1_V, dn2_t1_DV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>up2_t1_S, up2_t1_DS</td>
<td>up1_t1_S, up1_t1_DS</td>
<td>dn1_t1_S, dn1_t1_DS</td>
<td>dn2_t1_S, dn2_t1_DS</td>
</tr>
<tr>
<td></td>
<td>Accident occurrence time (t)</td>
<td>up2_t_O, up2_t_DO</td>
<td>up1_t_O, up1_t_DO</td>
<td>dn1_t_O, dn1_t_DO</td>
<td>dn2_t_O, dn2_t_DO</td>
</tr>
<tr>
<td></td>
<td></td>
<td>up2_t_V, up2_t_DV</td>
<td>up1_t_V, up1_t_DV</td>
<td>dn1_t_V, dn1_t_DV</td>
<td>dn2_t_V, dn2_t_DV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>up2_t_S, up2_t_DS</td>
<td>up1_t_S, up1_t_DS</td>
<td>dn1_t_S, dn1_t_DS</td>
<td>dn2_t_S, dn2_t_DS</td>
</tr>
</tbody>
</table>

- O: Occupancy
- V: Volume
- S: Speed
- DO: Differ. Occupancy
- DV: Differ. Volume
- DS: Differ. Speed
2. Data Preparation

**Loop data processing procedure**

**STEP 1**
Archiving loop detector data for analysis

**STEP 2**
Missing data imputation

**STEP 3**
Allocation of individual accident ID

**STEP 4**
Derivation of explanatory Variables for traffic conditions

**STEP 5**
Dataset establishment for normal traffic conditions

Spatio-temperal traffic variable analysis
3. Methodology and Model Development

Binary Logistic Regression

- BLR can be applied to binary classification problem

1: Hazardous traffic condition
2: Accident-free condition

Binary Classification

\[ \text{Pr} (ACC_i = 1 | X_i) = \frac{\exp[f(X_i, \beta) + \beta_i]}{1 + \exp[f(X_i, \beta) + \beta_i]} \]

where,
- \( \text{Pr}(ACC_i) \): Probability of accident likelihood
- \( X_i \): accident involved traffic variables
- \( f(X_i, \beta) \): Function of \( X_i \) and parameters
3. Methodology and Model Development

Result of Model Development – Case 4

- **R-square**

<table>
<thead>
<tr>
<th>CASE</th>
<th>-2 Log Likelihood</th>
<th>Cox와 Snell의 R-Square</th>
<th>Nagelkerke R-Square</th>
<th>Number of accident</th>
</tr>
</thead>
<tbody>
<tr>
<td>CASE 4</td>
<td>264.329</td>
<td>0.226</td>
<td>0.301</td>
<td>142 건</td>
</tr>
</tbody>
</table>

- **Correct classification Rate (CCR)**

<table>
<thead>
<tr>
<th>CASE</th>
<th>Observed</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accident Occurrence</td>
<td>Percent Correct (%)</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>89</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>37</td>
</tr>
<tr>
<td>case4</td>
<td>1</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>89</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
</tr>
</tbody>
</table>
### Result of Model Development – Case 4

#### Significant Variables

<table>
<thead>
<tr>
<th>CASE</th>
<th>Variable</th>
<th>Beta</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>case 4</td>
<td>up1_t1_O</td>
<td>0.63607</td>
<td>0.03142</td>
</tr>
<tr>
<td>collision</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>and rear-</td>
<td>up1_t2_V</td>
<td>-0.00701</td>
<td>0.03692</td>
</tr>
<tr>
<td>end collision</td>
<td>up2_t1_O</td>
<td>-0.54779</td>
<td>0.02818</td>
</tr>
<tr>
<td></td>
<td>up2_t2_V</td>
<td>0.00671</td>
<td>0.01903</td>
</tr>
<tr>
<td></td>
<td>up1_t1_DO</td>
<td>-1.23492</td>
<td>0.01089</td>
</tr>
<tr>
<td></td>
<td>up1_t2_DO</td>
<td>1.11977</td>
<td>0.02173</td>
</tr>
<tr>
<td></td>
<td>up2_t1_DS</td>
<td>0.06987</td>
<td>0.05738</td>
</tr>
<tr>
<td></td>
<td>dn2_t1_DV</td>
<td>0.00816</td>
<td>0.00346</td>
</tr>
<tr>
<td></td>
<td>dn2_t1_DS</td>
<td>0.20489</td>
<td>0.00047</td>
</tr>
<tr>
<td></td>
<td>dn2_t2_DS</td>
<td>-0.12514</td>
<td>0.02855</td>
</tr>
<tr>
<td></td>
<td>Constant</td>
<td>-1.13773</td>
<td>0.00013</td>
</tr>
</tbody>
</table>
3. Methodology and Model Development

**Significant traffic variables for accident likelihood**

<table>
<thead>
<tr>
<th>Significant traffic variables in Case 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Up2</strong></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>t−30min (t2)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>t−15min (t1)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>t</td>
</tr>
</tbody>
</table>

Acc Occurrence
3. Methodology and Model Development

**Accident Likelihood**

- Estimate probability of accident likelihood using Binary Logistic Regression

\[ Pr(ACC_i = 1 \mid X_i) = \frac{\exp[f(X_i, \beta)]}{1 + \exp[f(X_i, \beta)]} \]

\[ f(X_i, \beta) = 0.636X_1 - 0.007X_2 - 0.548X_3 + 0.006X_4 - 1.235X_5 + 1.120X_6 + 0.070X_7 + 0.008X_8 + 0.205X_9 - 0.125X_{10} - 1.138X_{11} \]

<table>
<thead>
<tr>
<th>up1_t1_O</th>
<th>up1_t2_V</th>
<th>up2_t1_O</th>
<th>up2_t2_V</th>
<th>up1_t1_DO</th>
<th>up1_t2_DO</th>
<th>up2_t1_DS</th>
<th>dn2_t1_DS</th>
<th>dn2_t1_DO</th>
<th>dn2_t2_DS</th>
<th>상수</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.63607</td>
<td>-0.00701</td>
<td>-0.54779</td>
<td>0.00671</td>
<td>-1.23492</td>
<td>1.11977</td>
<td>0.06987</td>
<td>0.00816</td>
<td>0.20489</td>
<td>-0.12514</td>
<td>-1.13773</td>
</tr>
</tbody>
</table>

**Accident-free traffic Condition**

**Abnormal traffic Condition**
4. Application and Technical Issue

Application

- Warning Information System
  - Variable speed limit
  - Accident likelihood information

Real-time Traffic Monitoring

Real-time safety surrogate measure

Estimate accident likelihood

\[
\Pr(ACC_i = 1 | X_i) = \frac{\exp[f(X_i, \beta)]}{1 + \exp[f(X_i, \beta)]}
\]

Traffic condition is safe?
\[
\Pr(ACC_i) \geq \text{threshold}
\]

No

Provide warning information using VMS
  \[\rightarrow\] Accident likelihood information
  \[\rightarrow\] Variable Speed Limit

Yes
4. Application and Technical Issue

Technical Issue

Resolution of traffic data

Duration and frequency for warning information provision

When accident likelihood index exceeds the threshold, provide warning information

<table>
<thead>
<tr>
<th>threshold</th>
<th>Time(min)</th>
<th>Warning Information Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>915</td>
<td>64%</td>
</tr>
<tr>
<td>0.7</td>
<td>615</td>
<td>43%</td>
</tr>
<tr>
<td>0.8</td>
<td>555</td>
<td>38%</td>
</tr>
</tbody>
</table>
5. Conclusion

Future studies

- Adaptive Variable Speed Limit system
- Study for various traffic accident type and model application
- Field tests for enhancing transferability
5. Conclusion

Conclusion

- Proposed a model to estimate traffic accident likelihood using real-time traffic data
- Application for Warning information and VSL
- Proposed system can lead to safer driving leading to accident prevention

---

**Warning information**

**Variable Speed Limit**

**Safe Driving**

Reducing vehicle speed

**Decrease accident likelihood**