Identify high-risk locations using large-scale real-world connected vehicle data

Dr. Kun Xie

Lecturer in Transportation Engineering
University of Canterbury
Contact: kun.xie@canterbury.ac.nz

Source: Laird Technologies
About University of Canterbury

University of Canterbury is located in Christchurch, New Zealand.

- Southern Alps
- New Regent Street
- Lyttelton Harbour
- Botanic Garden
About University of Canterbury

Overview
- Founded by scholars of Oxford and Cambridge universities in 1873
- Second oldest university in New Zealand
- 16,906 students in 2018

Notable Alumni
- **Ernest Rutherford** – “the father of atom”
- Āpirana Ngata - the foremost Māori politician to have ever served in Parliament
- Helen Connon - the first woman student in the British Empire to receive an Honours degree
About University of Canterbury

- Department of Civil and Natural Resources Engineering
  - Civil Engineering is ranked 7th in Academic Ranking of World Universities (ARWU)
  - About 250 undergraduates every year

- Connected Traffic Systems Lab
  - 5 faculty members, covering research areas of traffic safety, network modeling, signal control, connected and autonomous vehicles
Road traffic crashes result in approximately **1.24 million** deaths and **20 to 50 million** injuries globally each year (World Health Organization, 2013)

Road traffic injuries are the **leading cause** of death among young people, aged 15–29 years.
Plans/Coalitions for Road Safety

Source: https://visionzeronetwork.org/resources/vision-zero-cities/
In New Zealand, the total social cost of motor vehicle injury crashes in 2015 is estimated to be $3.8 billion.
Road Safety in New Zealand

Figure 4: Road deaths per 100,000 population in seven countries

Source: Ministry of Transport
Road Safety in New Zealand

- **Safer Journeys** is the government's strategy to guide improvements in road safety over the period 2010 to 2020.

- **Safe System approach** aims for a more forgiving road system that takes human fallibility and vulnerability into account.
Figure 5: Areas of high concern (percentage reduction/increase since 2010) per 100,000

- **Red**: Deaths or serious injuries in head on/run off road crashes
- **Blue**: Deaths or serious injuries in alcohol/drug crashes
- **Green**: ACC entitlement claims from motorcyclists
- **Orange**: 15-24 yr old drivers killed or seriously injured
- **Dark Brown**: Deaths or serious injuries in open road (80-100 km/h) crashes

Source: NZ Transport Agency
Road Safety in New Zealand

- NZ Open Crash Dataset (published in Sep. 2018)

Source: https://opendata-nzta.opendata.arcgis.com/
In the last few decades, road safety management heavily relied on historical crash data.
Reactive Safety Solutions: An Example

- Identify Hotspots of Pedestrian Crashes


Hotspots Identified by Potential for Safety Improvement
Measure Safety Performance without Crash Data?
**Surrogate Safety Measures (SSMs)** are used to quantify risks

- SSM describes the **“near-crash” scenarios** in which a vehicle would collide with another vehicle if they did not change their current intentions.
- Commonly used SSMs: Time to collision (TTC), Deceleration rate to avoid crash (DRAC)

---

**Proactive Safety Solutions**

- **Hotspot Identification**
- **Countermeasure Development**
- **Before-after Evaluation**

- Detect high-risk locations before the occurrence of crashes
- Enable active actions to prevent crashes
- Evaluate the safety treatments once they are implemented
Proactive Safety Solutions: Video-based safety assessment

Connected Vehicle Technology

- Connected vehicles (CVs) can be seen as moving sensors in the road network.
- Rich information generated by CVs then can be used to detect and detect high-risk locations.
Connected Vehicle Data

- Connected vehicle data from Michigan Safety Pilot Model Deployment (SPMD)
  - Data Acquisition System (DAS)
    - **DataFrontTargets File (4.34 G)**
      - Distance
      - Speed difference
    - **DataWSU File (11.2 G)**
      - Speed of the following vehicle
      - GPS locations
    - **DataLane File (3.78 G)**
  - Basic Safety Message (BSM) (68.2 G unzip)
  - Roadside Equipment (29.6 G unzip)

- Scope
  - Time: April, 2013
  - Sample rate: 10 Hz
  - A total of 62,589,725 messages
  - Locations: 75 selected highways
Connected Vehicle Data

Procedure of data processing:

1. **DataFrontTargets**
   - Key Variables: Device, Trip, Time, Range, RangeRate, CIPV
   - Select observations with CIPV = 1

2. **DataWsu**
   - Key Variables: Device, Trip, Time, GpsValidWsu, GpsTimeWsu, LatitudeWsu, LongitudeWsu, AltitudeWsu, GpsSpeedWsu, ValidCanWsu, SpeedWsu, AxWsu
   - Remove Duplicates
   - Select observations with GpsValidWsu = 1
   - Select observations with ValidCanWsu = 1

3. **Merged Dataset**
   - Key Variables: Device, Trip, Time, Range, RangeRate, CIPV, GpsValidWsu, GpsTimeWsu, LatitudeWsu, LongitudeWsu, AltitudeWsu, GpsSpeedWsu, ValidCanWsu, SpeedWsu, AxWsu
   - Remove outliers based on GpsSpeedWsu, SpeedWsu, AxWsu, AltitudeWsu
   - Remove all the observations with RangeRate < 0 and RangeRate - GpsSpeedWsu > 1
   - Select observations based on study horizon
   - Calculate SSMS using Key Variables

Merge DataFrontTargets and DataWsu together based on Device, Trip, and Time.
**Surrogate Safety Measures (SSM)**

**Risky Scenario 1:** Speed of the following vehicle is higher than that of the leading vehicle

**Risky Scenario 2:** Speed of the following vehicle is lower than or equal to that of the leading vehicle but the spacing between them is small.

\[
TTC = \begin{cases} 
\frac{l_0 - l_v}{v - v_1}, & v_2 > v_1 \\
\infty, & \text{otherwise}
\end{cases}
\]

\[
DRAC = \begin{cases} 
\frac{(v_2 - v_1)^2}{2(l_0 - l_v)}, & v_2 > v_1 \\
0, & \text{otherwise}
\end{cases}
\]

- \(v_1\): Initial speed of the leading vehicle
- \(v_2\): Initial speed of the following vehicle (remains constant in the car following scenario)
- \(l_v\): The length of the vehicle
- \(l_0\): Initial relative distance between the leading vehicle and the following vehicle

A new surrogate safety measure (SSM) is needed.
**Time to Collision with Disturbance (TTCD)**

- **TTCD**: the time interval between the given of the disturbance and the collision of two vehicles
- **Two possible collision outcomes** depending on the deceleration rate:
  - Outcome 1 – the leading vehicle is still decelerating when collision occurs;
  - Outcome 2 – the leading vehicle is fully stopped when collision occurs.

Assume the disturbance will result in a constant deceleration of the leading vehicle.
**Time to Collision with Disturbance (TTCD)**

**Critical Scenario** - the following vehicle collides with the leading vehicle exactly at the time when the leading vehicle stops

\[
l_0 + l_1 = l_2 + l_v
\]

\[
l_1 = \frac{v_1^2}{2d^*}
\]

\[
l_2 = v_2 t^* = \frac{v_1 v_2}{d^*}
\]

\[
d^* = \frac{2v_1 v_2 - v_1^2}{2(l_0 - l_v)}
\]

**Nomenclature**

- \(d\) The deceleration rate of the leading vehicle
- \(d^*\) The deceleration rate of the leading vehicle that causes the collision to occur exactly when the leading vehicle stops
- \(v_1\) Initial speed of the leading vehicle
- \(v_2\) Initial speed of the following vehicle (remains constant in the car following scenario)
- \(t_0\) The time when a disturbance is given
- \(t^*\) The time interval between the given of the disturbance and the full stop of the leading vehicle
- \(l_0\) Initial relative distance between the leading vehicle and the following vehicle
- \(l_1\) Distance travelled by the leading vehicle after being given a disturbance
- \(l_2\) Distance travelled by the following vehicle before colliding with the leading vehicle after a disturbance has been given to the leading vehicle
- \(N\) Number of Monte Carlo simulations
- \(TTCD\) The time interval between the given of the disturbance and the collision of the two vehicles
Time to Collision with Disturbance (TTCD)

**Collision Outcome 1:** \( d \leq d^* = \frac{2v_1v_2 - v_1^2}{2(l_0 - l_v)} \)

\[
l_0 + l_1 = l_2 + l_v
\]

\[
l_1 = v_1 \times TTCD - \frac{1}{2}d \times TTCD^2
\]

\[
l_2 = v_2 \times TTCD
\]

\[
TTCD = \frac{(v_1 - v_2) + \sqrt{(v_1 - v_2)^2 + 2d(l_0 - l_v)}}{d}
\]

**Nomenclature**

- \( d \): The deceleration rate of the leading vehicle
- \( d^* \): The deceleration rate of the leading vehicle that causes the collision to occur exactly when the leading vehicle stops
- \( v_1 \): Initial speed of the leading vehicle
- \( v_2 \): Initial speed of the following vehicle (remains constant in the car following scenario)
- \( l_0 \): The time when a disturbance is given
- \( l_v \): The time interval between the given of the disturbance and the full stop of the leading vehicle
- \( l_v \): The length of the vehicle
- \( l_0 \): Initial relative distance between the leading vehicle and the following vehicle
- \( l_1 \): Distance travelled by the leading vehicle after being given a disturbance
- \( l_2 \): Distance travelled by the following vehicle before colliding with the leading vehicle after a disturbance has been given to the leading vehicle
- \( N \): Number of Monte Carlo simulations
- \( TTCD \): The time interval between the given of the disturbance and the collision of the two vehicles
Collision Outcome 2: \( d > d^* = \frac{2v_1v_2 - v_1^2}{2(l_0 - l_v)} \)

\[
\begin{align*}
  l_0 + l_1 &= l_2 + l_v \\
  l_1 &= \frac{v_1^2}{2d} \\
  l_2 &= v_2 \times TTCD \\
  TTCD &= \frac{2d(l_0 - l_v) + v_1^2}{2dv_2}
\end{align*}
\]
Time to Collision with Disturbance (TTCD)

To sum up, TTCD can be expressed as:

\[
TTCD = \begin{cases} 
\frac{(v_1 - v_2) + \sqrt{(v_1 - v_2)^2 + 2d(l_0 - l_v)}}{2d}, & d \leq \frac{2v_1v_2 - v_2^2}{2(l_0 - l_v)} \\
\frac{2d(l_0 - l_v) + v_1^2}{2dv_2}, & d > \frac{2v_1v_2 - v_2^2}{2(l_0 - l_v)}
\end{cases}
\]

If TTCD is less than a predefined threshold TTCD*, a conflict is detected. Define Conflict Risk with Disturbance (CRD) as the probability of being involved with conflicts under hypothetical disturbance d:

\[
CRD = \Pr(TTCD < TTCD^*)
\]

We assume that d follows a shifted gamma distribution (17.315, 0.128, 0.657) (calibrated by Kuang and Qu (2015) using NGSIM data). Monte Carlo method is used to compute CRD.
Correlation between Risk Identified by SSMs and Rear-end Crash Data

- Risk of a car-following scenario identified by SSMs
  - TTC - conflict occurrence
  - DRAC - conflict occurrence
  - TTCD - CRD (the probability of conflict occurrence)
- Aggregate trip-based risk into location-based risk
- Potential confounding effect
Correlation between Risk Identified by SSMs and Rear-end Crash Data

- Control for the traffic exposure effect
  - \( \text{Risk rate} = \frac{\text{Risk identified by each SSM}}{\text{Number of CV GPS points}} \)
  - \( \text{Crash rate} = \frac{\text{Rear–end crash count}}{\text{Traffic volume}} \)

- Investigate the correlation between risk rate and crash rate
Identify the Optimal Threshold for Each SSM

- Risks identified by TTC, DRAC and TTCD are subject to the selection of thresholds
- To obtain the optimal SSM thresholds, every possible threshold value incremented by 0.1 within a reasonable range was tested.

<table>
<thead>
<tr>
<th>SSM</th>
<th>Optimal Threshold</th>
<th>Pearson’s correlation coefficient</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTC</td>
<td>2.3 s</td>
<td>0.41</td>
<td>0.0002</td>
</tr>
<tr>
<td>DRAC</td>
<td>3.0 m/s²</td>
<td>0.39</td>
<td>0.0005</td>
</tr>
<tr>
<td>TTCD</td>
<td>1.7 s</td>
<td>0.45</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
Risk Identified for One Trip

Identified Risks from 3762th to 3768th Time Intervals

<table>
<thead>
<tr>
<th>Time Interval</th>
<th>Relative Distance (m)</th>
<th>Relative Speed (m/s)</th>
<th>Following Vehicle Speed (m/s)</th>
<th>Conflict Presence Identified by TTC</th>
<th>Conflict Presence Identified by DRAC</th>
<th>CRD Identified by TTCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>3762</td>
<td>2.77</td>
<td>0.21</td>
<td>13.05</td>
<td>0</td>
<td>0</td>
<td>0.14</td>
</tr>
<tr>
<td>3763</td>
<td>2.79</td>
<td>0.21</td>
<td>13.09</td>
<td>0</td>
<td>0</td>
<td>0.07</td>
</tr>
<tr>
<td>3764</td>
<td>2.81</td>
<td>0.21</td>
<td>13.07</td>
<td>0</td>
<td>0</td>
<td>0.14</td>
</tr>
<tr>
<td>3765</td>
<td>2.83</td>
<td>0.21</td>
<td>13.11</td>
<td>0</td>
<td>0</td>
<td>0.11</td>
</tr>
<tr>
<td>3766</td>
<td>2.85</td>
<td>0.21</td>
<td>13.17</td>
<td>0</td>
<td>0</td>
<td>0.17</td>
</tr>
<tr>
<td>3767</td>
<td>2.88</td>
<td>0.19</td>
<td>13.21</td>
<td>0</td>
<td>0</td>
<td>0.10</td>
</tr>
<tr>
<td>3768</td>
<td>2.90</td>
<td>0.17</td>
<td>13.31</td>
<td>0</td>
<td>0</td>
<td>0.08</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Detect High-risk Locations Using CV Data

- TTCD have greater potential to infer crash risks than AADT data.
Detect High-risk Locations Using CV Data

- It shows the potential of using CV data to detect risk and thus supports more proactive safety management.

Crashes in April, 2013

Risks identified by TTCD each day in April, 2013
Related Study: Smartphone Data for Risk Detection

- Four dangerous driving behaviors
  - Fast acceleration, hard breaking, phone use while driving, and speeding

Related Study: Prevent Secondary Crashes Using Connected Vehicles

An example of primary and secondary crashes.

Use of connected vehicles for reducing secondary crash risk

Related Study: Prevent Secondary Crashes Using Connected Vehicles

- Simulation Setting
  - Transmission range: 1,000 meters
  - Transmission frequency: 10 time per second
  - No communication latency and information loss

Spatiotemporal Distribution of Conflicts under Different Market Penetration Rates (MPRs)
Real-world connected vehicle pilot test data collected in Ann Arbor, Michigan was used to generate surrogate safety measures (SSMs) for risk identification.

By imposing a hypothetical disturbance, TTCD is able to detect rear-end collision risks in various car following scenarios, even when the leading vehicle has a higher speed.

Results showed that risk data captured by TTCD could achieve the highest level of correlation with historical rear-end crash data compared with other traditional SSMs.

Connected vehicle data has the potential to advance proactive road safety management.


Thank You!

Dr. Kun Xie
Lecturer in Transportation Engineering
University of Canterbury
Contact: kun.xie@canterbury.ac.nz