Automated Proactive Road Safety Analysis
Transportation Research At McGill Seminar

Nicolas Saunier
nicolas.saunier@polymtl.ca

ÉCOLE
POLYTECHNIQUE
MONTREAL

November 25th 2010
1 Motivation

2 Probabilistic Framework for Automated Road Safety Analysis

3 Experimental Results using Video Data

4 Investigating Collision Factors Using Microscopic Data

5 Conclusion
A World Health Issue

Over 1.2 million people die each year on the world’s roads, and between 20 and 50 million suffer non-fatal injuries. In most regions of the world this epidemic of road traffic injuries is still increasing. (Global status report on road safety, World Health Organization, 2009)
# A World Health Issue

N. Saunier

<table>
<thead>
<tr>
<th>RANK</th>
<th>LEADING CAUSE</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ischaemic heart disease</td>
<td>12.2</td>
</tr>
<tr>
<td>2</td>
<td>Cerebrovascular disease</td>
<td>9.7</td>
</tr>
<tr>
<td>3</td>
<td>Lower respiratory infections</td>
<td>7.0</td>
</tr>
<tr>
<td>4</td>
<td>Chronic obstructive pulmonary disease</td>
<td>5.1</td>
</tr>
<tr>
<td>5</td>
<td>Diarrhoeal diseases</td>
<td>3.6</td>
</tr>
<tr>
<td>6</td>
<td>HIV/AIDS</td>
<td>3.5</td>
</tr>
<tr>
<td>7</td>
<td>Tuberculosis</td>
<td>2.5</td>
</tr>
<tr>
<td>8</td>
<td>Trachea, bronchus, lung cancers</td>
<td>2.3</td>
</tr>
<tr>
<td>9</td>
<td>Road traffic injuries</td>
<td>2.2</td>
</tr>
<tr>
<td>10</td>
<td>Prematurity and low birth weight</td>
<td>2.0</td>
</tr>
<tr>
<td>11</td>
<td>Neonatal infections and other</td>
<td>1.9</td>
</tr>
<tr>
<td>12</td>
<td>Diabetes mellitus</td>
<td>1.9</td>
</tr>
<tr>
<td>13</td>
<td>Malaria</td>
<td>1.7</td>
</tr>
<tr>
<td>14</td>
<td>Hypertensive heart disease</td>
<td>1.7</td>
</tr>
<tr>
<td>15</td>
<td>Birth asphyxia and birth trauma</td>
<td>1.5</td>
</tr>
<tr>
<td>16</td>
<td>Self-inflicted injuries</td>
<td>1.4</td>
</tr>
<tr>
<td>17</td>
<td>Stomach cancer</td>
<td>1.4</td>
</tr>
<tr>
<td>18</td>
<td>Cirrhosis of the liver</td>
<td>1.3</td>
</tr>
<tr>
<td>19</td>
<td>Nephritis and nephrosis</td>
<td>1.3</td>
</tr>
<tr>
<td>20</td>
<td>Colon and rectum cancers</td>
<td>1.1</td>
</tr>
</tbody>
</table>

**TOTAL 2004**

<table>
<thead>
<tr>
<th>RANK</th>
<th>LEADING CAUSE</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ischaemic heart disease</td>
<td>12.2</td>
</tr>
<tr>
<td>2</td>
<td>Cerebrovascular disease</td>
<td>9.7</td>
</tr>
<tr>
<td>3</td>
<td>Chronic obstructive pulmonary disease</td>
<td>7.0</td>
</tr>
<tr>
<td>4</td>
<td>Lower respiratory infections</td>
<td>5.1</td>
</tr>
<tr>
<td>5</td>
<td>Road traffic injuries</td>
<td>3.6</td>
</tr>
<tr>
<td>6</td>
<td>Trachea, bronchus, lung cancers</td>
<td>3.5</td>
</tr>
<tr>
<td>7</td>
<td>Diabetes mellitus</td>
<td>2.5</td>
</tr>
<tr>
<td>8</td>
<td>Hypertensive heart disease</td>
<td>2.3</td>
</tr>
<tr>
<td>9</td>
<td>Stomach cancer</td>
<td>2.2</td>
</tr>
<tr>
<td>10</td>
<td>HIV/AIDS</td>
<td>2.0</td>
</tr>
<tr>
<td>11</td>
<td>Nephritis and nephrosis</td>
<td>1.9</td>
</tr>
<tr>
<td>12</td>
<td>Self-inflicted injuries</td>
<td>1.9</td>
</tr>
<tr>
<td>13</td>
<td>Liver cancer</td>
<td>1.7</td>
</tr>
<tr>
<td>14</td>
<td>Colon and rectum cancer</td>
<td>1.7</td>
</tr>
<tr>
<td>15</td>
<td>Oesophagus cancer</td>
<td>1.5</td>
</tr>
<tr>
<td>16</td>
<td>Violence</td>
<td>1.4</td>
</tr>
<tr>
<td>17</td>
<td>Alzheimer and other dementias</td>
<td>1.4</td>
</tr>
<tr>
<td>18</td>
<td>Cirrhosis of the liver</td>
<td>1.3</td>
</tr>
<tr>
<td>19</td>
<td>Breast cancer</td>
<td>1.3</td>
</tr>
<tr>
<td>20</td>
<td>Tuberculosis</td>
<td>1.1</td>
</tr>
</tbody>
</table>

**TOTAL 2030**

Road Safety Analysis

N. Saunier

Motivation

Probabilistic Framework

Experimental Results

Investigating Collision Factors

Conclusion

- Limits of the traditional approach based on historical collision data:
  - problems of availability and quality
  - insufficient data to understand the processes that lead to collisions
  - reactive approach
  - pedestrians: issues are made more acute by the rarity of collisions and the lack of data (exposure)

- Need for proactive approaches and surrogate safety measures that do not depend on the occurrence of collisions
Surrogate Safety Measures

Research on surrogate safety measures that
- bring complementary information
- are related to traffic events that are more frequent than collisions and can be observed in the field
- are correlated to collisions, logically and statistically
Traffic Conflicts

A traffic conflict is “an observational situation in which two or more road users approach each other in space and time to such an extent that a collision is imminent if their movements remain unchanged” [Amundsen and Hydén, 1977]

- Traffic Conflict Techniques
- Limits caused by the data collection process (human observers in the field)
  - cost
  - intra- and inter-observer variability
- Mixed validation results
The Safety/Severity Hierarchy

Accidents
- Serious Conflicts
- Slight Conflicts
- Potential Conflicts
- Undisturbed passages

Various severity measures

Various severity measures

The Safety/Severity Hierarchy

F
I
PD

Undisturbed passages
Potential Conflicts
Slight Conflicts
Serious Conflicts
Accidents

The Safety/Severity Hierarchy

F
I
PD

Various severity measures
Motivation

Need for automated tools to address the shortcomings of reactive diagnosis methods and traffic conflict techniques
The Collision Course

A traffic conflict is “an observational situation in which two or more road users approach each other in space and time to such an extent that a collision is imminent if their movements remain unchanged” [Amundsen and Hydén, 1977]

The extrapolation hypotheses must be specified.
Rethinking the Collision Course

For two interacting road users, many chains of events may lead to a collision.

It is possible to estimate the probability of collision if one can predict the road users’ future positions.

- Learn road users’ motion patterns (including frequencies), represented by actual trajectories called prototypes.
- Match observed trajectories to prototypes and extrapolate.

[Saunier et al., 2007, Saunier and Sayed, 2008]
A Simple Example

$t_1$, $t_2$, $t_3$, $t_4$, $t_5$
Collision Points

- Using of a finite set of extrapolation hypotheses, the collision points $CP_n$ are enumerated.
- The probability of collision $P$ is computed by summing the probabilities of reaching each potential collision point.

$$P(\text{Collision}(U_i, U_j)) = \sum_{n} P(\text{Collision}(CP_n))$$

- The expected Time To Collision is also computed (if there is at least one collision point, i.e. $P(\text{Collision}(U_i, U_j)) > 0$)

$$TTC(U_i, U_j, t_0) = \frac{\sum_{n} P(\text{Collision}(CP_n)) t_n}{P(\text{Collision}(U_i, U_j))}$$

[Saunier et al., 2010]
Video Sensors

Video sensors have distinct advantages:

- they are easy to install (or can be already installed)
- they are inexpensive
- they can provide rich traffic description (e.g. road user tracking)
- they can cover large areas
- their recording allows verification at a later stage
Automated Proactive Road Safety Analysis

N. Saunier

Motivation

Probabilistic Framework

Experimental Results

Investigating Collision Factors

Conclusion

Video-based System

Image Sequence

Road User Trajectories

Interactions

Camera Calibration

Applications

Labeled Images for Road User Type

Motion patterns, volume, origin-destination counts, driver behavior

Traffic conflicts, exposure and severity measures, interacting behavior
Automated Proactive Road Safety Analysis

N. Saunier

Motivation
Probabilistic Framework
Experimental Results
Investigating Collision Factors
Conclusion

Feature-based Road User Tracking in Video Data

Good enough for safety analysis and other applications, including the study of pedestrians and pedestrian-vehicle interactions [Saunier and Sayed, 2006]
Motion Pattern Learning

Traffic Conflict Dataset, Vancouver
58 prototype trajectories (2941 trajectories)

Reggio Calabria, Italy
58 prototype trajectories (138009 trajectories)
The Kentucky Dataset

- Video recordings kept for a few seconds before and after the sound-based automatic detection of an interaction of interest
  - 229 traffic conflicts
  - 101 collisions
  - The existence of an interaction or its severity is not always obvious
  - The interactions recorded in this dataset involve only motorized vehicles
  - Limited quality of the video data: resolution, compression, weather and lighting conditions
- Calibration done using the tool developed by Karim Ismail at UBC [Ismail et al., 2010b]
Severity Indicators

Side conflict
Severity Indicators

Side conflict
Severity Indicators

Parallel conflict
Severity Indicators

Side collision
Severity Indicators

Side collision
Severity Indicators

Parallel collision
Distribution of Indicators

Maximum Collision Probability

Minimum TTC
Spatial Distribution of the Collision Points

Traffic Conflicts

0
8
16
24
32
40
48
56
64
72
Spatial Distribution of the Collision Points

Collisions

Spatial Distribution of the Collision Points

Collisions
Study Before and After the Introduction of a Scramble Phase

Data collected in Oakland, CA [Ismail et al., 2010a]
Distribution of Severity Indicators

Histogram of Before-and-After TTC

Frequency of traffic events
TTC$_{min}$ in seconds
Before and After Distribution of the Collision Points

Motivation

Probabilistic Framework

Experimental Results

Investigating Collision Factors

Conclusion
Objectives

- Understand collision processes by studying the similarities of interactions with and without a collision (conflicts)
- There is some evidence that evasive actions undertaken by road users involved in conflicts may be of a different nature than the ones attempted in collisions [Davis et al., 2008]
  - Importance for surrogate safety measures: what interactions without a collision may be used as surrogates for collisions?
- Use of data mining techniques (k-means and hierarchical agglomerative clustering method) to cluster the data

[Saunier et al., 2011]
# Description of Interactions

<table>
<thead>
<tr>
<th>Categorical attributes</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of day</td>
<td>weekday, week end</td>
</tr>
<tr>
<td>Lighting condition</td>
<td>daytime, twilight, nighttime</td>
</tr>
<tr>
<td>Weather condition</td>
<td>normal, rain, snow</td>
</tr>
<tr>
<td>Interaction category</td>
<td>same direction (turning left and right, rear-end, lane change), opposite direction (turning left and right, head-on), side (turning left and right, straight)</td>
</tr>
<tr>
<td>Interaction outcome</td>
<td>conflict, collision</td>
</tr>
</tbody>
</table>
## Description of Interactions

<table>
<thead>
<tr>
<th>Numerical attributes</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Road user type</strong></td>
<td></td>
</tr>
<tr>
<td>passenger car</td>
<td>number of road users</td>
</tr>
<tr>
<td>van, 4x4, SUV</td>
<td>number of road users</td>
</tr>
<tr>
<td>bus</td>
<td>number of road users</td>
</tr>
<tr>
<td>truck (all sizes)</td>
<td>number of road users</td>
</tr>
<tr>
<td>motorcycle</td>
<td>number of road users</td>
</tr>
<tr>
<td>bike</td>
<td>number of road users</td>
</tr>
<tr>
<td>pedestrian</td>
<td>number of road users</td>
</tr>
<tr>
<td><strong>Type of evasive action</strong></td>
<td></td>
</tr>
<tr>
<td>No evasive action</td>
<td>number of evasive actions</td>
</tr>
<tr>
<td>Braking</td>
<td>number of evasive actions</td>
</tr>
<tr>
<td>Swerving</td>
<td>number of evasive actions</td>
</tr>
<tr>
<td>Acceleration</td>
<td>number of evasive actions</td>
</tr>
<tr>
<td><strong>3 attributes from $\Delta v$</strong></td>
<td>km/h</td>
</tr>
<tr>
<td><strong>6 values from $s$</strong></td>
<td>km/h</td>
</tr>
</tbody>
</table>

Description of Interactions: 

- **Road user type**
  - passenger car
  - van, 4x4, SUV
  - bus
  - truck (all sizes)
  - motorcycle
  - bike
  - pedestrian

- **Type of evasive action**
  - No evasive action
  - Braking
  - Swerving
  - Acceleration

- **Numerical attributes**
  - 3 attributes from $\Delta v$ (km/h)
  - 6 values from $s$ (km/h)
Distribution of Speed Attributes

Norm of the velocity difference

- DeltaVmin
- DeltaV
- DeltaVmax

Collision
Conflict

Speed

- Smin1
- Smin2
- S1
- S2
- Smax1
- Smax2

Smin1 Smin2 S1 S2 Smax1 Smax2
0 10 20 30 40 50 60
Speed (km/h)

Norm of the velocity difference

Smin1 Smin2 S1 S2 Smax1 Smax2
0 10 20 30 40 50
DeltaV (km/h)

DeltaVmin DeltaV DeltaVmax
0 10 20 30 40 50
DeltaV (km/h)

DeltaVmin DeltaV DeltaVmax
0 10 20 30 40 50
3 Clusters

Interaction outcome

- **Cluster 1**
  - Collision: 25
  - Conflict: 15

- **Cluster 2**
  - Collision: 100
  - Conflict: 30

- **Cluster 3**
  - Collision: 40
  - Conflict: 80

Number of interactions

Clustering: 3 Clusters

Cluster 1

Cluster 2

Cluster 3
Clustering: Speed Attributes

- **DeltaV (km/h)**
  - Cluster 3
  - Cluster 2
  - Cluster 1

- **Speed (km/h)**
  - Cluster 3
  - Cluster 2
  - Cluster 1
Clusters: Interaction Category

- **Cluster 1**: collisions, highest speeds, categories side straight and same direction turning right
- **Cluster 2**: almost pure conflicts, lowest speeds
- **Cluster 3**: collisions, medium speeds, categories same direction turning left and right and same direction changing lanes
Conclusion

- Tools and framework for **automated** road safety analysis using video sensors
- **Large** amounts of data: data mining and visualization for safety analysis
- Future work:
  - Still more work on data collection techniques (computer vision)
  - Validation of proactive methods for road safety analysis
  - Understanding and modelling of collision processes: collect more data
- Need for more **open** science: data and code sharing

http://nicolas.saunier.confins.net
Collaboration with

- Clark Lim and Tarek Sayed (University of British Columbia)
- Karim Ismail (Carleton University)
- Nadia Mourji, Bruno Agard (École Polytechnique de Montréal)

Questions?


In *Transportation Research Board Annual Meeting Compendium of Papers*, Washington, D.C.
10-2715.

**Saunier, N., Mourji, N., and Agard, B. (2011).*
Investigating collision factors by mining microscopic data of vehicle conflicts and collisions.
In *Transportation Research Board Annual Meeting Compendium of Papers*, Washington, D.C.
11-0117.

**Saunier, N. and Sayed, T. (2006).*
A feature-based tracking algorithm for vehicles in intersections.
In *Third Canadian Conference on Computer and Robot Vision*, Québec. IEEE.

**Saunier, N. and Sayed, T. (2008).*
A Probabilistic Framework for Automated Analysis of Exposure to Road Collisions.