Waze Annotated Bibliography

Dan Flynn

February 2018

Table of Contents

Annotated bibliography of tools and approaches used in existing analyses of road safety data, relevant for the Safety Data Initiative Waze/EDT pilot project.

Machine learning approaches in transportation safety

1 Modeling the dynamics of driver's dilemma zone perception using agent based modeling techniques

- In a driving simulation study, used Agent Based Models (ABM) to investigate how drivers impacts of driving in a 'dilemma zone', such as too close to an intersetion to safely stop. Models include the MATsim dyanamic agent-based traffic simulation model.

2 Influence of injury risk thresholds on the performance of an algorithm to predict crashes with serious injuries

- Optimization for first responders, using logistic regression on National Automotive Sampling System / Crashworthiness Data System (NASS/CDS) to determine how variation in injury risk thresholds affects crash predictions. Standard logistic regression approach, where outcomes are binary for each type of crash (separate models, not ordinal).

3 A two-stage mining framework to explore key risk conditions on one-vehicle crash severity

- This research combines data mining and a logistic regressino approach to identify crash severity in one-vechicle crashes. Genetic mining rule (GMR) model developed, to identify 'rules' which correspond to variables most associated with risk of a crash. The variables were then used in a hierarchical logistic regression (mixed logit model) to identify road conditions associated with serious crashes.
- Similar to proposed SDI Waze project approach, where random forests used to identify combinations of variables highly associated with EDT-level crashes, and then logistic regression used to assign probability of a crash to Waze events and test statistical significance. Use a training/validation approach for the rule-mining, 70% of data for training, 30% for validation.

4 Estimating likelihood of future crashes for crash-prone drivers

5 Investigating injury severity risk factors in automobile crashes with predictive analytics and sensitivity analysis methods

6 Empirical Bayes approach for estimating urban deer-vehicle crashes using police and maintenance records

7 Development of a Prediction Model for Crash Occurrence by Analyzing Traffic Crash and Citation Data

8 Application of classification algorithms for analysis of road safety risk factor dependencies

9 Data science application in intelligent transportation systems: An integrative approach for border delay prediction and traffic accident analysis

10 Estimating the safety performance of urban road transportation networks

11 Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory

12 The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives

13 Performance Measures for Prioritizing Highway Safety Improvements Based on Predicted Crash Frequency and Severity

14 Factors influencing specificity and sensitivity of injury severity prediction (ISP) algorithm for AACN

15 Proactive Assessment of Accident Risk to Improve Safety on a System of Freeways

16 Prioritizing Highway Safety Manual's crash prediction variables using boosted regression trees

17 An Exploratory Computational Piecewise Approach to Characterizing and Analyzing Traffic Accident Data

18 Risk Factors Analysis for Drivers with Multiple Crashes

- Identifying high-risk drivers using demographic charascteristic, historical violations, and specific violation types with negative binomial regression. Crash estimation model identifies the set of predictors most strongly associated with high-risk drivers. Standard regression approach, models evaluated by AIC.

19 Crash Prediction Method for Freeway Facilities with High Occupancy Vehicle (HOV) and High Occupancy Toll (HOT) Lanes

20 Developing Crash Models with Supporting Vector Machine for Urban Transportation Planning

21 Predicting Traffic Flow Regimes From Simulated Connected Vehicle Messages Using Data Analytics and Machine Learning

22 Exploration of Advances in Statistical Methodologies for Crash Count and Severity Prediction Models

23 Analyzing Traffic Crash Severity in Work Zones under Different Light Conditions

24 Predicting motor vehicle collisions using Bayesian neural network models: An empirical analysis

25 Modeling crash and fatality counts along mainlines and frontage roads across Texas: The roles of design, the built environment, and weather

Crowdsourced data analysis approaches

26 Learning from the crowd: Road infrastructure monitoring system

21 Predicting Traffic Flow Regimes From Simulated Connected Vehicle Messages Using Data Analytics and Machine Learning

Spatial regression for road safety

References

1. Abbas, M. M. & Machiani, S. G. Modeling the dynamics of driver's dilemma zone perception using agent based modeling techniques. *International journal of transportation* **4,** 1–14 (2016).

2. Bahouth, G., Digges, K. & Schulman, C. Influence of injury risk thresholds on the performance of an algorithm to predict crashes with serious injuries. in *Annals of advances in automotive medicine/annual scientific conference* **56,** 223 (Association for the Advancement of Automotive Medicine, 2012).

3. Chiou, Y.-C., Lan, L. W. & Chen, W.-P. A two-stage mining framework to explore key risk conditions on one-vehicle crash severity. *Accident Analysis & Prevention* **50,** 405–415 (2013).

4. Das, S., Sun, X., Wang, F. & Leboeuf, C. Estimating likelihood of future crashes for crash-prone drivers. *Journal of Traffic and Transportation Engineering (English Edition)* **2,** 145–157 (2015).

5. Delen, D., Tomak, L., Topuz, K. & Eryarsoy, E. Investigating injury severity risk factors in automobile crashes with predictive analytics and sensitivity analysis methods. *Journal of Transport & Health* **4,** 118–131 (2017).

6. Gkritza, K., Souleyrette, R. R., Baird, M. J. & Danielson, B. J. Empirical bayes approach for estimating urban deer-vehicle crashes using police and maintenance records. *Journal of transportation engineering* **140,** 04013002 (2013).

7. Gonzalez-Velez, E. & Gonzalez-Bonilla, A. *Development of a prediction model for crash occurrence by analyzing traffic crash and citation data.* (Transportation Informatics Tier I University Transportation Center, University at Buffalo, 2017).

8. Kwon, O. H., Rhee, W. & Yoon, Y. Application of classification algorithms for analysis of road safety risk factor dependencies. *Accident Analysis & Prevention* **75,** 1–15 (2015).

9. Lin, L. Data science application in intelligent transportation systems: An integrative approach for border delay prediction and traffic accident analysis. (State University of New York at Buffalo, 2015).

10. Lord, D. & Persaud, B. N. Estimating the safety performance of urban road transportation networks. *Accident analysis & prevention* **36,** 609–620 (2004).

11. Lord, D., Washington, S. P. & Ivan, J. N. Poisson, poisson-gamma and zero-inflated regression models of motor vehicle crashes: Balancing statistical fit and theory. *Accident Analysis & Prevention* **37,** 35–46 (2005).

12. Lord, D. & Mannering, F. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice* **44,** 291–305 (2010).

13. Morgan, N. S. Performance measures for prioritizing highway safety improvements based on predicted crash frequency and severity. (Auburn, 2013).

14. Pal, C. *et al.* Factors influencing specificity and sensitivity of injury severity prediction (isp) algorithm for aacn. *International journal of automotive engineering* **7,** 15–22 (2016).

15. Pande, A., Nuworsoo, C. & Shew, C. *Proactive assessment of accident risk to improve safety on a system*

of freeways. (Mineta Transportation Institute; MTI Report 11-15, 2012).

16. Saha, D., Alluri, P. & Gan, A. Prioritizing highway safety manual's crash prediction variables using boosted regression trees. *Accident Analysis & Prevention* **79,** 133–144 (2015).

17. Saleem, F., Asa, E. & Membah, J. An exploratory computational piecewise approach to characterizing and analyzing traffic accident data. *International Journal of Scientific and Technical Research in Engineering* (2016).

18. Shawky, M. & Al-Ghafli, A. Risk factors analysis for drivers with multiple crashes. *International Journal of Engineering and Applied Sciences* **3,** 42–48 (2016).

19. Srinivasan, S. *et al. Crash prediction method for freeway facilities with high occupancy vehicle (hov) and high occupancy toll (hot) lanes.* (Florida Department of Transportation, 2015).

20. Sun, X., Das, S. & Broussard, N. Developing crash models with supporting vector machine for urban transportation planning. in *17th international conference road safety on five continents (rs5c 2016), rio de janeiro, brazil, 17-19 may 2016.* (Statens väg-och transportforskningsinstitut, 2016).

21. Vasudevan, M., Curtis, C., Lowman, A. & O'Hara, J. *Predicting traffic flow regimes from simulated connected vehicle messages using data analytics and machine learning.* (ITS Joint Program Office, Department of Transportation, 2016).

22. Wang, K. Exploration of advances in statistical methodologies for crash count and severity prediction models. (University of Connecticut, 2016).

23. Wei, X., Shu, X., Huang, B., Taylor, E. L. & Chen, H. Analyzing traffic crash severity in work zones under different light conditions. *Journal of Advanced Transportation* **2017,** (2017).

24. Xie, Y., Lord, D. & Zhang, Y. Predicting motor vehicle collisions using bayesian neural network models: An empirical analysis. *Accident Analysis & Prevention* **39,** 922–933 (2007).

25. Xu, J., Kockelman, K. M. & Wang, Y. MODELING crash and fatality counts along mainlanes and 1 frontage roads across texas: 2 the roles of design, the built environment, and weather 3. in *Presented at the 93rd annual meeting of the transportation research* **22,** 24 (2014).

26. Masino, J., Thumm, J., Frey, M. & Gauterin, F. Learning from the crowd: Road infrastructure monitoring system. *Journal of Traffic and Transportation Engineering (English Edition)* **4,** 451–463 (2017).

27. Gill, G., Sakrani, T., Cheng, W. & Zhou, J. Comparison of adjacency and distance-based approaches for spatial analysis of multimodal traffic crash data. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* **42,** (2017).

28. Rhee, K.-A., Kim, J.-K., Lee, Y.-I. & Ulfarsson, G. F. Spatial regression analysis of traffic crashes in seoul. *Accident Analysis & Prevention* **91,** 190–199 (2016).

29. Schultz, B., Grant G. *Use of roadway attributes in hot spot identification and analysis.* (Utah Department of Transportation; Brigham Young University-Provo, 2015).

30. Zeng, Q. & Huang, H. Bayesian spatial joint modeling of traffic crashes on an urban road network. *Accident Analysis & Prevention* **67,** 105–112 (2014).