Uncovering Deep Structure of Determinants in Large Truck Fatal Crashes

Subasish Das, Ph.D. Associate Transportation Researcher, Texas A&M Transportation Institute (s-das@tti.tamu.edu)
Mouyid Islam, Ph.D. Research Faculty, CUTR, University of South Florida (mouyid@usf.edu)
Anandi Dutta, Ph.D., Senior Lecturer, Department of CSE, Ohio State University (dutta.34@osu.edu)
Tahmida Hossain Shimu, Traffic Designer, HDR (Tahmida.Shimu@hdrinc.com)

Research conducted by

Texas A&M Transportation Institute
THE TEXAS A&M UNIVERSITY SYSTEM
CUTR CENTER FOR URBAN TRANSPORTATION RESEARCH
THE OHIO STATE UNIVERSITY COLLEGE OF ENGINEERING

Abstract

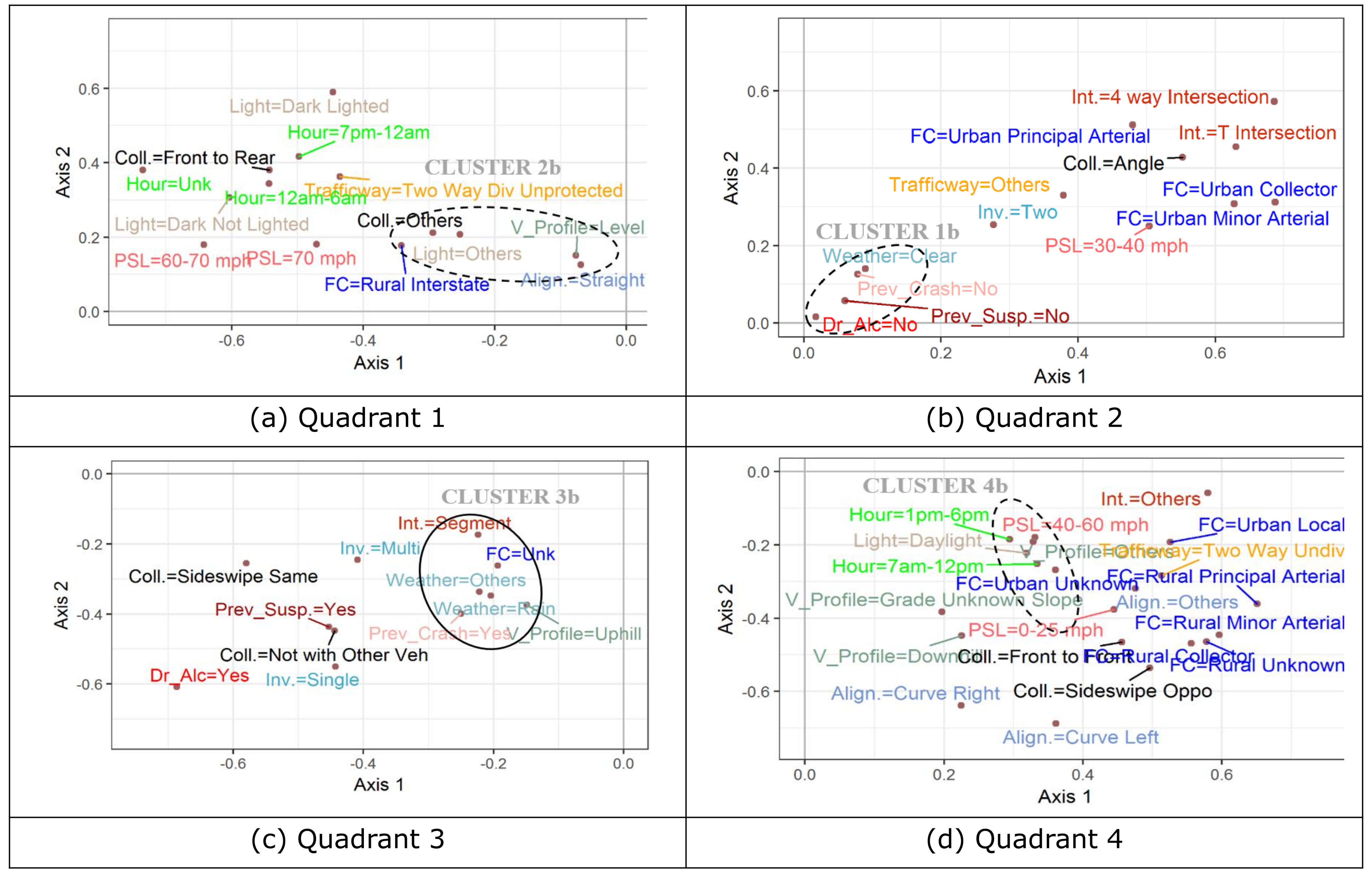
- The number of fatalities and severe injuries in large truck-related crashes has significantly increased since 2009.
- Used six years (2010–2015) of large truck fatal crash data from the Fatality Analysis Reporting System (FARS) were used.
- Applied taxicab correspondence analysis (TCA), an artificial intelligence method known for dimension reduction, to large truck fatal crash data in order to investigate the complex interaction between multiple factors under a two-dimensional map.

Methodology

- Collected crash data file, vehicle data file, and person data file from FARS.
- Conducted preliminary exploratory analysis at the beginning to examine the significant factors that may contribute to crash occurrence.
- The final dataset contains 14 variables.
- Applied TCA to identify the key association patterns.

Why TCA?

- Summarizes the key aspects of a complex data set by projecting the multivariate data on two-dimensional or three-dimensional maps.
- One important advantage of TCA over CA is that it remains closest to the original data as it directly acts on the correspondence matrix P without calculating a dissimilarity (or similarity) measure between the rows or columns.



(a) Quadrant 1 (b) Quadrant 2 (c) Quadrant 3 (d) Quadrant 4

MCA plot of attributes

Log Odds Ratio of Four Crash Prevalence Scenarios

Row Labels	Count	Prev_Crash (Y_vs_N)	Prev_Suspen (Y_vs_N)	Alc (Y_vs_N)	SingleVeh (Y_vs_N)
Cluster01	7	2.07	3.82	3.97	3.27
Cluster02	72	2.67	2.72	1.61	1.42
Cluster03	16	2.26	3.97	1.74	2.58
Cluster04	275	1.77	1.82	1.51	1.59
Cluster05	3	0.00	0.00	0.00	0.00
Cluster06	30	0.47	1.34	2.30	2.49
Cluster07	29	0.00	0.00	5.03	3.64
Cluster08	5	0.00	0.00	0.00	0.00
Cluster09	764	1.14	1.09	0.82	1.01
Cluster10	6	0.00	2.03	0.00	0.00
Cluster11	5	0.00	0.00	0.00	0.00
Cluster12	72	0.47	1.69	0.13	1.93
Cluster13	9	0.00	2.72	0.00	0.00
Cluster14	1623	0.61	0.80	0.60	0.66
Cluster15	4	0.00	0.00	0.00	0.00

Key Clusters

- **Cluster 1a** (Urban collector or minor arterial, Intersection= T-intersection or 4-way intersection, Posted speed limit= 30–40 mph)
- **Cluster 2a** (Lighting=Dark lighted or dark not lighted, Roadway= Two-way divided unprotected, Collision=Front to rear)
- **Cluster 3a** (Impaired driver, Single truck or sideswipe same direction, and Single or multiple vehicle involvement)
- **Cluster 3b** (Segment, Weather=Rain or others, Uphill, Previous crash conducted by the truck driver=Yes and unknown functional class)
- **Cluster 4a** (Functional class=Rural principal or minor arterial, collector & urban local, Roadway=Two-way undivided, Collision=Sideswipe opposite, low-posted speed limit, and Alignment=Others)

Conclusions

- Urban intersections are the setting for a disproportionate number of large truck fatal crashes.
- There is a strong association between two-way roadways with an unprotected median and large truck fatal crashes.
- Two distinct clusters (impaired driver's involvement in single-vehicle crashes, and drivers with the past crash record being involved in inclement weather crashes) indicate human error associated patterns in large truck fatal crashes.
- Driving in non-daytime hours is associated with a high number of truck-involved crashes.
- Individual-level TCA analysis identified 69 distinct clouds based on four prevalence driving behaviors. A total of 13 clusters show LOR values greater than zero for all prevalence behavioral groups. These clusters represent 8,339 large truck drivers and occupants (LTDOs). Out of these 8,339 LTDOs, the people that possess these prevalence traits are over-represented.