Influence of Crash Report Forms on Red Light Running Crash Data

Khaled Shaaban
Qatar University/Department of Civil Engineering, Doha, Qatar
Email: kshaaban@qu.edu.qa

Dina Elnashar
Qatar Mobility Innovations Center, Doha, Qatar
Email: kshaaban@qu.edu.qa

Essam Radwan
University of Central Florida/Department of Civil, Environmental, and Construction Engineering
Email: kshaaban@qu.edu.qa

***Abstract*—Red light running is a significant problem in urban areas in Florida and the United States and one of the major causes for crashes at signalized intersections. In order to provide potential countermeasures to the red light running problem, these types of crashes need to be studied accurately. Some researchers use crash databases to obtain red light running crash information through filtering crashes reported as “disregarded traffic signal” without going through the actual crash reports. The purpose of this paper was to analyze the actual long form crash reports for a sample of 1,600 crashes and compare them with the data obtained from the database to investigate if errors often occur in how information on red light running crashes are. The analysis of the random sample showed that the percentage of crashes related to red light running reported in the database was 3.1% compared to 5.8% from the actual crash reports. The difference was mainly due to the way of filling out the crash report by the law enforcement personnel. Although, it is always hard to draw solid conclusions about the attributing cause of a collision, it is necessary to standardize the format and coding process for the crash reports and to train the law enforcement personnel to help accurately identify the real cause of the crash. Finally, the findings presented in this research imply that the red light running crashes are being underreported, and that long form crash reports should be considered as part of any future analysis.**

***Index Terms*—Red light running, crashes, traffic signal, crash report**

I. Introduction

Researchers use crash databases for studies related to safety including red light running (RLR) crashes. For example, Retting et al. [1] used two national databases to quantify the occurrence of red-light running crashes, as well as to summarize the characteristics of red-light runners. The databases included the Fatal Analysis Reporting System (FARS), which collects virtually all U.S. police-reported crashes involving a fatality, and the General Estimates System (GES), which is based on a nationally representative probability sample of crashes with a varying degree of injury and property damage.

Mohamedshah et al. [2] developed regression models that indicated that the variables affecting RLR crashes include the traffic volume on both the entering and crossing streets, the type of traffic control provided at intersections, and the width of the crossing street of the intersection. The study suggested using the results to better target the specific intersections for traffic law enforcement measures. For example, instead of installing a RLR camera on all the intersections or even just intersections with past RLR crash histories, this study suggests guidelines for such targeting. Therefore, traffic and safety engineers and enforcement agencies should consider for treatment those intersections which have higher entering traffic volume on the mainline and cross-street, particularly those intersections with high cross-street (minor road) volumes; intersections where the volume on a minor road is relatively high, coupled with a wide mainline street; and locations with fully-actuated traffic signals.

Yan et al. [3] used the 1999-2001 Florida traffic crash data to investigate the crash propensity of three aspects of risk factors related to traffic environment, driver characteristics, and vehicle types. The results showed that seven environmental factors (number of lanes, crash time, weather, highway character, day of week, urban/rural, and speed limit), four factors related to driver characteristics (driver age, alcohol/drug use, physical defect, and driver residence), and type of vehicle are significantly associated with the risk of red-light running crashes. Furthermore, the study confirmed significant interaction effects between risk factors including crash time and Highway character, number of lanes and urban/rural, weather condition and driver age, driver age

Manuscript received January 10, 2013; revised February 15, 2013; accepted February 28, 2013

doi: 10.12720/jtle.1.2.197-201

and gender, alcohol/drug use and gender; and type of vehicle and gender.

In orders to better understand the underlying crash mechanisms, Wang and Abdel-Aty [4] analyzed left-turn crashes occurring at 197 four-legged signalized intersections over 6 years. Initially, 1,575 left-turn crashes were determined from the Crash Analysis Reporting (CAR) system maintained by the Florida Department of Transportation (FDOT). In addition to the initial 1,575 left-turn crashes, additional 1,523 crashes, which were originally recorded as other crash types, were determined to be left-turn crashes since at least one of the involved vehicles was turning left by inspecting their crash reports through the state crash report image retrieval system. From reviewing the crash reports, it was also found that, for around 30% of left-turn crashes, left-turning vehicles' traveling directions were recorded as the destination direction, but not the initiating direction before turning.

Wang and Abdel-Aty [5] studied the crashes from 2000 to 2005 for 197 four-legged signalized intersections from Orange and Hillsborough counties in the Central Florida area. They studied 13,218 crashes obtained from the CAR system maintained by FDOT, which is the most complete crash database for the state road intersections in Florida. The purpose of studying the crashes was to identify the angle collisions attributed to red light running. Of 2431 initial angle collisions, 836 were right-angle crashes. Another 97 crashes, which were classified as non-angle crashes, were considered as right-angle crashes by inspecting their original crash reports through the state crash report document image retrieval system. The proportion of right-angle crashes among angle crashes was 34.4%, and 10.4% of right-angle crashes were from other crash types, which shows the importance the crash reports to obtain accurate information about the studied crashes.

Wang et al. [6] also studied the data for 177 four-legged signalized intersections were collected from Florida to investigate how intersection attributes affect this safety influence area and how the varied safety influence areas for intersection approaches improve safety analysis using data from the CAR system. In the state of Florida, crashes within 50 feet of an intersection are classified as "at intersection"; when crashes are within 250 feet of an intersection, the site locations are recorded as "influenced by intersection." The crash location distance is supposed to be measured from the center of the intersection. In practice, when police officers determine the distance for a crash, they are frequently measuring it from the stop bar rather than the center. Among the total 13,218 crashes for the selected intersections, there were 2276 crashes coded as "not at intersection" in the site location. A total of 1940 crashes were changed to "influenced by intersection" after reviewing the crash reports. In general, the crash databases may not provide accurate information about the crashes. Consequently, researchers have sometimes turned to the actual hard copy crash reports filed by victims and/or law enforcement personnel who investigated the report. Using this approach, the purpose of this paper was to analyze the actual hard copy crash reports and compare them with the data obtained from a crash database to investigate if errors often occur in how information on a red light running crash report is coded in the computerized crash database.

II. Literature Review

Researchers use crash databases for studies related to safety including red light running crashes. For example, Retting et al. [7] used two national databases to quantify the occurrence of red-light running crashes, as well as to summarize the characteristics of red-light runners. The databases included the Fatal Analysis Reporting System (FARS), which collects virtually all U.S. police-reported crashes involving a fatality, and the General Estimates System (GES), which is based on a nationally representative probability sample of crashes with a varying degree of injury and property damage.

III. Crash Data Sources

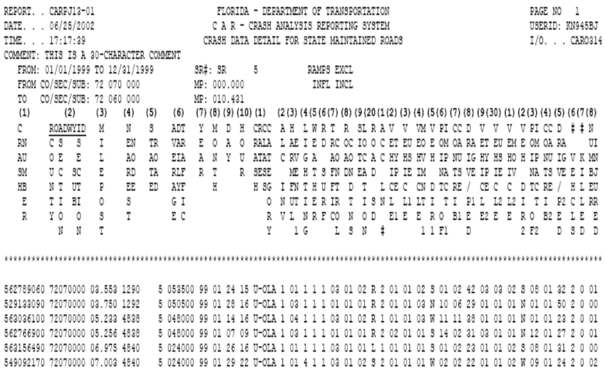
The data for this research was collected from the two databases; the CAR System and the Safety Office's Electronic Document Management System (EDMS).

A. CAR Database

The information contained in the CAR database has been compiled from information collected for the purpose of identifying, evaluating, or planning safety enhancements. This system and its products identify information used for the purpose of developing highway safety construction improvement projects. The database was originally generated by merging crash data from the Department of Highway Safety and Motor Vehicles (DHSMV) with roadway information from FDOT. The database is updated yearly. All reported crashes with a fatality, an injury, and high property damage occurred on state roads are included in this database [8].

The database basically contains all the information recorded in the long form crash report. For each crash, there are more than 300 variables used to describe the site and time of the crash, the geometric conditions, the traffic control, and drivers and pedestrian's characteristics. The downloaded detailed data for the state roads in Florida from the CAR system contained more than 80 columns supplying information from the crash record for example, crash report number, crash date, time of the accident, day of the week, distance, mile post, county, section number, location mile post, location node, location distance, average daily traffic, crash-level alcohol involved code, total amount of damage, injury level, number of people injured in the crash, number of fatalities, first harmful event in the crash, second harmful event, contributing cause, and other important information about the crash.

For each variable, several code values were assigned to represent different categories of the variable. For the variable "Contributing Cause", the code 01 is used to denote "no improper driving/action", 02 denotes "careless driving", 03 denotes "failed to yield right-of-way", 04 denotes "Improper backing", 05 denotes "improper lane change”, 06 denotes “improper turn”, 07 denotes “alcohol – under influence, 08 denotes “drugs – under influence”, 09 denotes “alcohol and drugs – under influence”, 10 denotes “followed too closely”, 11 denotes “disregarded traffic signal”, 12 denotes “exceeded safe speed limit”, 13 denotes “disregarded stop sign”, 14 denotes “failed to maintain equipment/vehicle”, 15 denotes “improper passing”, 16 denotes “drove left of center”, 17 denotes “exceeded stated speed limit”, 18 denotes “obstructing traffic”, 19 denotes “improper load”, 20 denotes “disregarded other traffic control”, 21 denotes “driving wrong side/way, 22 denotes “fleeing police”, 23 denotes “vehicle notified”, 24 denotes “driver distraction”, and 77 denotes “all other”. A snapshot of the data downloaded from the CAR system is shown in Fig. 1 [9].



```
REPORT. . CARPJ13-01                FLORIDA - DEPARTMENT OF TRANSPORTATION                  PAGE NO   1
DATE. . . 06/25/2002                C A R - CRASH ANALYSIS REPORTING SYSTEM                 USERID: KN945BJ
TIME. . . 17:17:39                  CRASH DATA DETAIL FOR STATE MAINTAINED ROADS            I/O. . . CAR0314
COMMENT: THIS IS A 30-CHARACTER COMMENT
  FROM: 01/01/1999 TO 12/31/1999        SR#: SR      5          RAMPS EXCL
  FROM CO/SEC/SUB: 72 070 000           MP: 000.000              INFL INCL
  TO   CO/SEC/SUB: 72 060 000           MP: 010.431

562789060 72070000 03.553 1290    5 053500 99 01 24 15 U-OLA 1 01 1 1 1 03 01 02 R 2 01 01 02 S 01 02 42 03 03 02 S 08 01 32 2 0 01
529133090 72070000 03.750 1292    5 050500 99 01 28 16 U-OLA 1 03 1 1 1 01 01 01 R 2 01 01 03 N 10 06 29 01 01 01 N 01 01 50 2 0 00
563036100 72070000 05.233 4838    5 048000 99 01 14 16 U-OLA 1 04 1 1 1 03 01 02 R 1 01 01 03 W 11 11 38 01 01 01 N 01 01 23 2 0 01
562766900 72070000 05.256 4838    5 048000 99 01 07 09 U-OLA 1 03 1 1 1 01 01 01 R 2 02 01 01 S 14 02 31 03 01 01 N 12 01 27 2 0 01
563156490 72070000 06.975 4840    5 024000 99 01 26 16 U-OLA 1 01 1 1 1 03 01 01 L 1 01 01 01 S 01 02 23 01 01 02 S 08 01 31 2 0 00
549092170 72070000 07.003 4840    5 024000 99 01 29 22 U-OLA 1 01 4 1 1 03 01 02 S 2 01 01 01 W 02 02 22 01 01 02 W 09 01 24 2 0 02
```

Figure 1. Sample of a Report from the CAR System

B. EDMS Database

In general, traffic crashes can be reported by the use of two forms commonly referred as the “long form” and the “short form”. The long form report includes the narrative, the diagram, and any necessary update/continuation. The long form is used when the following criteria are met: death or personal injury; leaving the scene involving damage to attend vehicles or property; and driving under the influence. The short form report is used to report other types of traffic crashes. FDOT use the Safety Office’s EDMS to scan and import the long form crash reports received from the DHSMV into the Department Management System and to ensure those records are legible, properly indexed/archived, and retrievable [10].

Each crash report consists of four pages, and the following section documents entries needed for these four pages. Page 1 includes the date of crash, time of crash, crash report number, county code, city code; distance from the intersecting roadway in miles, street names, roadway characteristics, drivers information, vehicles information, directions of travel, point of impact, alcohol involvement, gender, race, and type of injury.

Page 2 includes the second and third vehicle information, the first contributing cause, vehicles’ defect if any, vehicles’ movement, type of collision, lighting condition, road surface condition, weather, and traffic control. A sample of Pages 1 and 2 is shown in Fig. 2.

First, page 3 includes a narrative explaining how the crash occurred exactly. This part is very important since it can help determine the exact contributing cause of the crash. The information in this section depends on the amount of information entered by the officer. Second, it includes also the information related to the passengers in every vehicle including the passenger name, current address, city and state, zip code, date of birth, race, sex, location, injury type, and if he/she was ejected or not. Third, it lists the violators’ details including the vehicle, name of violator, Florida statute number, charge, and citation number. Fourth, it includes the information related to any witnesses for the crash including the witness name, current address, city, state, and zip code. Finally, it lists the information of the investigator including badge number, department, and date of the report.

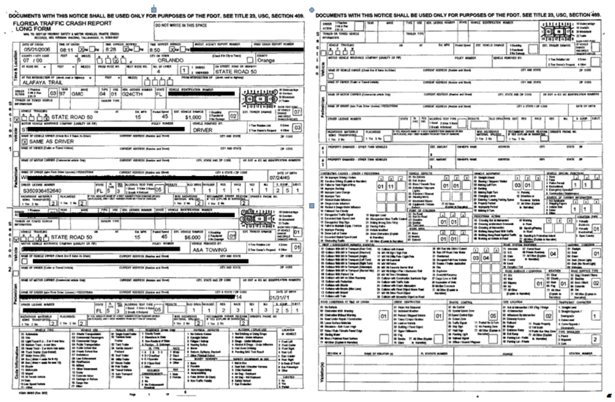


Figure 2. Long Form Crash Report Sample (Pages 1 and 2)

Page 4 provides a diagram for the intersection including a north arrow to show the directions, lanes, median, striping, vehicle 1, vehicle 2, vehicle 3 if applicable, pedestrian if involved in the crash, and bicycle if involved in the crash. The diagram also describes the location of every vehicle including the direction, movement, and point of contact. This page provides detailed information that can assist in determining the real cause of a crash especially if the law enforcement officer made an error in recording the correct code previous page. This information on this page will be used to compare the data from the long form crash reports with the data from the CAR system. A sample of Pages 1 and 2 is shown in Figure 3.

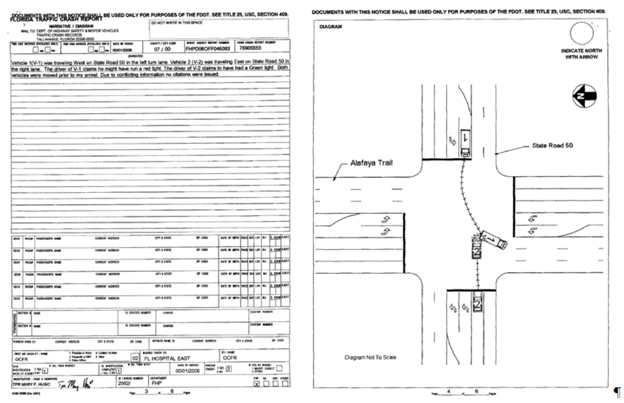


Figure 3. Long Form Crash Report Sample (Pages 3 and 4)

IV. DATA COLLECTION

The purpose of this paper was to analyze the actual long form crash reports for a sample of 1,600 crashes and compare them with the data obtained from the database to investigate if errors often occur in how information on the red light running crash report is coded in the crash reports. The random sample was selected using the SAS software.

V. DATA ANALYSIS

The percentage of red light running crashes reported in the CAR system for the random sample of 1,600 crashes was 3.1%. These crashes had "disregarded traffic signal" as the contributing cause in the database. Based on the detailed analysis of the long form crash reports, it was found that the actual percentage of red light running was 5.8%. Figure 4 depicts the percentage of the different contributing causes listed for the random sample.

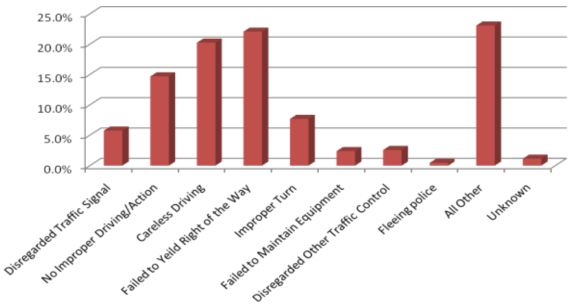


Figure 4. Contributing Causes for the Random Sample from the Long Form Crash Reports

Some of the frequent problems found in filling the red light running crash reports included:

- Not filling out the contributing cause at all.
- The no fault driver was coded first although the second driver had Disregarded Traffic Signal as a contributing cause.
- Coding left turning vehicles running red light as "improper turn".
- When both the left turning vehicle and the through vehicle claim to have a green arrow and a solid green respectively, the contributing cause is coded as "all other".
- Coding failing to stop and entering the intersection as "careless driving".
- Coding failing to stop and entering the intersection as "failed to yield right-of-way".
- Mistakenly coding the contributing cause as "disregarding other traffic control".

The previous results show the importance of standardizing the format and coding process of the crash reports by the enforcement personnel to help identify the real pattern of crashes at the studied location.

VI. CONCLUSIONS AND RECOMMENDATIONS

In this paper, the actual hard copy crash reports files were analyzed and compared with the data obtained from the CAR database for a random sample of 1,600 crash reports. The purpose of the analysis was to investigate if errors often occur in how information on a red light running crashes are coded. The analysis showed that results from the CAR system and from the crash report forms are different in reporting the actual number of red light running related crashes. The findings presented in this research imply that the red light running crashes are being underreported if the long form crash reports are not used. The percentage of red light running crashes reported in the CAR system was 3.1%. These crashes had "disregarded traffic signal" as the first contributing cause. The actual percentage based on the detailed analysis of the long form crash reports of red light running was found to be 5.8%. This study shows the importance of standardizing the format and coding process for the long form crashes by the police officers to help identify the real pattern of crashes at any studied location.

Representing red light running crashes only through "disregard traffic signal" noted reports may underestimate the extent of red light running effects at a given intersection. Researchers may have to turn to the actual hard copy crash reports filed by the law enforcement personnel who investigated the crash to obtain more accurate results.

REFERENCES

[1] R. A. Retting, A. F. Williams, and M. A. Greene. "Red-Light Running and Sensible Countermeasures," T*ransportation Research Record: Journal of the Transportation Research Board*, no. 1640, Transportation Research Board of the National Academies, Washington, D.C., pp. 23-26, 1998.

[2] Y. M. Mohamedshah, L. W. Chen, and F. M. Council. "Association of Selected intersection factors with red light running crashes," in *Proceedings of the 70th Annual ITE Conference*, Institute of Transportation Engineers, Washington, D.C., 2000.

[3] X. Yan, E. Radwan, and E. Birriel. "Analysis of red-light running crashes based on quasi-induced exposure and multiple logistic regression method," presented at the 84st Annual Meeting of the Transportation Research Board, Washington, D.C., January 9-13, 2005.

[4] X. Wang and M. Abdel-Aty, "Modeling left-turn crash occurrence at signalized intersections by conflicting patterns," *Accident Analysis and Prevention Journal*, no. 40, pp. 76-88, 2008.

[5] X. Wang and M. Abdel-Aty. "Investigation of Right-Angle Crash Occurrence at Signalized Intersections," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2019, Transportation Research Board of the National Academies, Washington, D.C., pp. 156-168, 2007.

[6] X. Wang, M. Abdel-Aty, A, Nevarez, and J. Santos. "Investigation of safety influence area for four-legged signalized intersections: Nationwide survey and empirical inquiry," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2083, Transportation Research Board of the National Academies, Washington, D.C., pp. 86-95, 2008.

[7] R. A. Retting, R. G. Ulmer, and A. F. Williams, "Prevalence and characteristics of red light running crashes in the United States", *Accident Analysis and Prevention Journal*, no. 31, pp. 687-694, 1999.

[8] L. Lu, J. J. Lu, and P. S. Lin. "Development of an interface between FDOT'S crash analysis reporting system and the safety analyst," *Publication BDK84 977-01. FDOT*, U.S. Department of Transportation, 2009.

[9] Crash Data Reports User Manual for the Crash Analysis Reporting System (CAR), Florida Department of Transportation, Safety Office, Crash Records Division October 2007 Version.

[10] Electronic Document Management System (EDMS) Requirements, Florida Department of Transportation, August 2007 Version.

Khaled Shaaban is an Assistant Professor of Transportation Engineering in the Department of Civil Engineering at Qatar University. He received his Ph.D. degree in transportation engineering from the University of Central Florida, Orlando, Florida in 2005. His research interests include traffic operations, traffic simulation, and traffic safety.

He is a registered Professional Engineer in several states in the United States and a Professional Traffic Operations Engineer with over twenty years of experience in the transportation engineering industry. He has been a principal investigator or co-principal investigator for many research projects. The outcomes of these projects include numerous published papers and presentations on a variety of transportation engineering topics. Dr. Shaaban is an active member on many committees and international professional organizations.

Dina Elnashar holds an Master degree in transportation engineering from the University of Central Florida, Orlando, Florida. Her field of expertise covers traffic safety and traffic simulation. She has more than 11 years of industrial experience. She joined Qatar Mobility Innovation Center (QMIC), Doha, Qatar in 2010 to work on applied research projects related to utilizing real-time data in intelligent transportation, logistics management, and road safety applications. Previously, she worked as Research Assistant at the University of Central Florida, Orlando, Florida.

Essam Radwan is the Interim Chair of The Department of Civil, Environmental, and Construction Engineering at the University of Central Florida (UCF). He also serves as the Executive Director of the Center for Advanced Transportation Systems Simulation (CATSS). He received a BS degree in Civil Engineering with honors from Cairo University, Egypt and MS and Ph.D. Degrees from Purdue University. In 1978, he started his teaching career at Virginia Tech where he was tenured and promoted to Associate Professor then he moved to Arizona State University in 1984 and was promoted to full professor in 1990. Through a national search he was recruited to UCF to chair the Department of Civil and Environmental Engineering and he remained in that position till 2003. In 1998 he was appointed to be the founding Director of the federally established transportation center (CATSS).

His research interest is traffic characteristics and traffic signal control of freeway and street intersections. Included in this area are theoretical and applied models, discrete event simulation models, and large-scale databases developed to describe traffic behavior. He was instrumental in bridging the gap between human factors researchers, computer engineers, and transportation engineers. The "Human Centered Transportation Simulation Program" that he created under CATSS auspices has been viewed as a unique and one of the leading programs in the US. He directed and co-directed close to 55 research projects totaling well over $11 million in external funded research projects. Through research and graduate advising he published more than two hundred and fifty (250) technical papers and reports and his work has been extensively cited by his peers in the US and around the world.