Traffic Injury Prevention

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gcpi20

Mining groups of factors influencing bus/minibus crash severities on poor pavement condition roads considering different lighting status

Reuben Tamakloe, Enoch F. Sam, Madiha Bencekri, Subasish Das & Dongjoo Park

TRAFFIC INJURY PREVENTION
https://doi.org/10.1080/15389588.2022.2066658

Mining groups of factors influencing bus/minibus crash severities on poor pavement condition roads considering different lighting status

Reuben Tamakloe[a], Enoch F. Sam[b], Madiha Bencekri[a], Subasish Das[c], and Dongjoo Park[a]

[a]Department of Transportation Engineering, University of Seoul, Seoul, South Korea; [b]Department of Geography Education, University of Education, Winneba, Ghana; [c]Texas A&M Transportation Institute, Texas A&M University, College Station, Texas, USA

ABSTRACT

Objective: This study employs a data mining approach to discover hidden groups of crash-risk factors leading to each bus/minibus crash severity level on pothole-ridden/poor roads categorized under different lighting conditions namely daylight, night with streetlights turned on, and night with streetlights turned off/no streetlights.

Methods: The bus/minibus data employed contained 2,832 crashes observed on poor roads between 2011 and 2015, with variables such as the weather, driver, vehicle, roadway, and temporal characteristics. The data was grouped into three based on lighting condition, and the association rule data mining approach was applied.

Results: Overall, most rules pointing to fatal crashes included the hit-pedestrian variable, and these crashes were more frequent on straight/flat roads at night. While median presence was highly associated with severe bus/minibus crashes on dark-and-unlighted roads, median absence was correlated with severe crashes on dark-but-lighted roads. On-street parking was identified as a leading contributor to property-damage-only crashes in daylight conditions.

Conclusions: The study proposed relevant countermeasures to provide practical guidance to safety engineers regarding the mitigation of bus/minibus crashes in Ghana.

ARTICLE HISTORY
Received 5 November 2021
Accepted 12 April 2022

KEYWORDS
Developing country; data mining; severity; crash; bus; lighting condition

Introduction

Although public transport in developed countries is perceived to have a high level of safety per distance traveled compared to private vehicles (Kaplan and Prato 2012), the case of developing countries (DC's) is different. Due to the low vehicle ownership rates in DC's, there is a high dependence on privately operated, deregulated commercial, public transport vehicles (Sam, Daniels, et al. 2018). These vehicles are primarily old and largely less maintained – increasing the risk of occupants sustaining severe injuries should a crash occur. The worsening safety condition is further exacerbated by the large proportions of poor roadway infrastructure, including pavement and lighting conditions, characteristic of DC's.

Studies from DC's have identified that inadequate street lighting on roads is a proximate cause of crashes (Guobadia 2021). Nevertheless, as demonstrated in the detailed literature review on bus safety (see Text A1 in Appendix A of the supplementary material), it is evident that no study comprehensively explores the key patterns in crash data to identify groups of factors influencing bus/minibus crash severity considering lighting condition at the time of the crash. Buses/minibuses convey many people per mile traveled compared to other vehicle types, and crashes involving them are associated with enormous economic losses. It is imperative to conduct evidence-based studies to identify the risk factors associated with bus/minibus crashes.

The leading factors influencing injury severity in bus-involved crashes were identified as overspeeding, careless driving, loss of control, improper safety gap, driver intoxication, roadway geometry, and violation of the right of way (Barua and Tay 2010; Sam, Daniels, et al. 2018; Tamakloe et al. 2020). Regarding road infrastructure, poor pavement conditions and lighting systems are significant factors that critically impact crash likelihood and severity (Mukherjee and Mitra 2020). In DC's, where the quality of traffic infrastructure is predominantly low, the presence of unlighted and poor roads (pothole-ridden roadways) has had significant ramifications on severity outcomes (Sam, Daniels, et al. 2018). An investigation into the effect of lighting conditions on crashes also revealed that injury severity increases the most on dark roads (Wanvik 2009), especially during adverse weather. This makes it imperative to consider analyzing crashes on poor roads given each lighting condition to understand the risk factors associated with such crashes.

Existing literature reveals that bus safety researchers largely employed parametric methods (logit/probit models) for analysis (Sam, Daniels, et al. 2018; Tamakloe et al. 2020). Even though these have been successfully employed, certain assumptions regarding their use can negatively affect their

CONTACT Dongjoo Park djpark@uos.ac.kr Department of Transportation Engineering, University of Seoul, Seoul, South Korea.

Supplemental data for this article is available online at http://dx.doi.org/10.1080/15389588.2022.2066658

results. Pande and Abdel-Aty (2009) mentioned that classifying certain variables as dependent and independent is arbitrary, and that the correlations between the individual crash-causing variables may affect the results. Besides, parametric methods cannot show associations among variables in a crash dataset.

In contrast, Association Rules Mining (ARM) algorithm, a non-parametric technique, is deemed superior regarding prediction accuracy and requires no prespecified assumptions. According to Hong et al. (2020a), ARM can uncover latent relationships/groups of factors in datasets and handle many variables effectively. ARM has been demonstrated to help discover insightful trends and patterns in even small crash datasets ranging from 126 to 449 observations (Montella 2011; Das et al. 2020; Hong et al. 2020a; Tamakloe et al. 2022). For this reason, researchers have favored ARM for data mining investigations. Samerei et al. (2021) recently employed ARM to explore data pertaining to bus crashes in Australia and identified that groups of factors likely to lead to fatal crashes on weekends are pedestrians, overspeeding, and darkness. Text A1 and Table A1 in Appendix A of the supplemental material provide a detailed summary of the bus safety literature.

Despite the numerous studies conducted regarding bus safety, researchers have paid no attention to identifying the groups of factors influencing the severity of bus crashes on poor roads in DC's. Poor pavement conditions affect the safety and maneuverability of vehicles, particularly these buses/minibuses, which form the backbone of mobility and convey many people per distance traveled (Sam, Daniels, et al. 2018). Due to the vital role buses/minibuses play in these countries and the issue of poor road infrastructure in DC's, it is crucial to conduct a more comprehensive study to improve its safety. Against this backdrop, this study aims at employing the robust data mining technique to explore insightful groups of factors influencing bus/minibus crash severity on poor roads in Ghana, considering different lighting conditions present at the time of the crash. The study is expected to identify groups of factors influencing bus/minibus severity to develop more targeted policies to improve transport safety in DC's.

Methods

Data description

A detailed description of the study area and some issues associated with public transport safety in Ghana are presented in Text B1 and Figure B1 of Appendix B in the supplementary material. In total, 2,832 police-reported bus/minibus crashes were acquired from the Building and Road Research Institute (BRRI) for the study (2011–2015). These crashes occurred on both rural and urban roads in the country. Table B1 in Appendix B of the supplementary material summarizes the dataset segregated according to the different combinations of poor pavement and lighting conditions: PD = poor pavement + daylight; PNLO = poor pavement + night with streetlights turned on; PNLOFF = poor pavement + night with streetlights turned off/no streetlights. It is noteworthy in this study that poor roads connote roads that have potholes on them or are untarred. These roads were grouped together as they are uncomfortable to ply and pose significant safety challenges to drivers.

For this study, factors relating to drivers' characteristics and errors, temporal characteristics, weather conditions, roadway characteristics, vehicle characteristics, crash characteristics, and crash location characteristics were explored. Note that crash severity outcomes in Ghana is classified into four–fatal (at least one life is lost within thirty days of occurrence), hospitalized/serious/severe injury (at least one casualty is detained at the hospital for over a day), injured-not-hospitalized/minor injury (the most severely injured casualty requires not more than just first-aid), and property-damage-only/PDO (only roadway infrastructure, nearby properties or the vehicles are damaged) (Sam, Daniels, et al. 2018). See Text B2 in the supplementary material for the study framework description.

Association rule mining

This article employed the ARM methodology based on the Apriori algorithm proposed by Agrawal et al. (1993). Given an x-itemset, I, of x unique items i such that $I = \{i_1, i_2, \ldots, i_x\}$ and the bus/minibus crash database $D = \left\{d_1, d_2, \ldots, d_y\right\}$ where each crash record consists of a subset of items contained in I. Rules generated are of the form $A \Longrightarrow B$ (A = antecedent, and B = consequent), such that $A, B \subseteq I$ and $A \cap B = \emptyset$. The rule is explained as follows: if A exists as a crash-risk factor, then B is also likely to occur. The extraction of interesting rules depends on three main measures: support, confidence, and lift. The support, for an itemset shows how popular that itemset is, and is computed as:

$$Support(A \Longrightarrow B) = \frac{|A \cup B|}{|D|} \tag{1}$$

where $|A \cup B|$ shows the number of times itemsets A and B co-occur, and $|D|$ is the number of items in the database. The measure, confidence, shows how likely crash-risk factors or items in itemset A occurs when those in itemset B also occur, and is estimated as:

$$Confidence(A \Longrightarrow B) = \frac{Support(A \cup B)}{Support\ \ (A)} \tag{2}$$

Although the support-confidence framework prunes infrequent itemsets from the dataset, many more uninteresting rules remain. To ensure that more interesting rules are generated, a third measure known as the lift is usually used as the performance measure of the rule (Tamakloe et al. 2022). The lift of a rule $A \Longrightarrow B$ describes how likely a crash-risk factor B occurs given that A has occurred, while controlling for the popularity of crash-risk factor B. It can also be seen as a measure of the statistical interdependence of a rule. Lift is computed as follows:

$$Lift(A \Longrightarrow B) = \frac{Support(A \cup B)}{Support\ \ (A) \cdot Support(B)} \tag{3}$$

When $lift(A \Longrightarrow B) > 1$, the rule is said to be valuable/interesting (positive interdependence between the antecedent and consequent); otherwise, there is either no association or correlation between both itemsets (when $lift(A \Longrightarrow B) = 1$) or the occurrence of itemset A is entirely based on the occurrence of itemset B (when $lift(A \Longrightarrow B) < 1$ there is negative interdependence between the antecedent and consequent). The higher the lift of the rule, the more interesting/valuable it is.

Appropriate support, confidence, and lift thresholds must be selected to obtain insightful rules. As no specific rules govern the selection of support, confidence, and lift thresholds, this study employed the following thresholds after thorough trial-and-error experiments: support =1%; confidence =30%; lift =1.50.

Results

Figure C1 in Appendix C of the supplementary material displays the top thirty frequent items in the full and segregated datasets, and Figure 1 presents a wordcloud of the most frequent items in the full dataset. Essentially, the words or risk factors having the highest frequency of occurrence are displayed in larger fonts. Thus, the font sizes decrease with decreasing frequency of word occurrence. The five most frequent items are (in no exact order) as follows: {Driver's gender: male}, {Roadway surface condition: dry}, {DUI: negative}, {Roadworks in progress: no}, and {Hit-and-run: no}. It was observed that drivers aged between 30 and 49 were more associated with crashes on poor pavement condition roads than younger drivers (age <30) irrespective of the lighting condition. The most common severity level is hospitalized/severe injury. Besides, the most common driver error is inattentiveness.

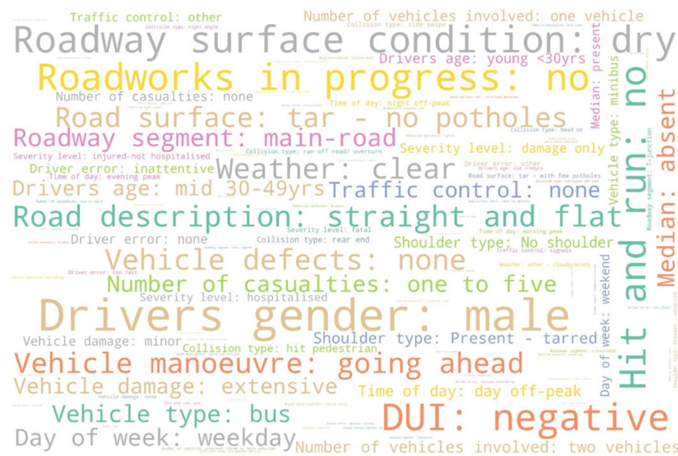

Figure 1. Wordcloud showing the most frequent crash-risk factors in the bus/minibus-involved crash dataset.

Rules for PD bus/minibus-involved crashes

Rules generated for each crash severity outcome are displayed in Table C1 (found in Appendix C of the supplementary material). The first rule in the fatal severity category (rule 1.1) highlights strong associations among items such as {Roadway segment: main road} and {Traffic control: none} with fatal bus/minibus crashes on poor pavement condition roads during the daylight condition. The second rule (rule 1.2) also revealed that other crash types involving going ahead maneuvers at the main road segment are likely fatal if they occur on poor pavements during daylight conditions (PD). Besides, it was observed in rules 1.9 and 1.10 that fatal bus/minibus crashes on PD roads are likely to occur on mainline segments without median separations. Buses/minibuses using median-unseparated mainline segments are exposed to much interaction with other vehicle streams. Sam, Brijs, et al. (2018) also identified that the absence of road medians increased the chance of observing increased bus crash severities.

Rule 2.1 suggests that non-defective vehicles involved in hit-pedestrian collisions at tarred shoulder present sections were likely to result in hospitalized/severe injury crashes. Rule 2.2 also showed that inattentive driving at untarred shoulder sections leading to hit-pedestrian crashes results in severe injuries. In general, severe/hospitalized hit-pedestrian crashes on weekdays during the daytime are likely to occur due to inattentive driving (rule 2.7: {Collision type: hit-pedestrian, Day of the week: weekday, Driver error: inattentive} => {Severity level: hospitalized}). However, severe/hospitalized hit-pedestrian crashes involving mid-aged drivers are not likely attributable to driver error (rule 2.4). In addition, {Shoulder type: present - tarred} is also commonly associated with severe crashes occurring on PD roads. This is likely as drivers mostly perceive shoulders as an extension of narrow roads – increasing their speeds to compensate for the broader roads (Sam, Daniels, et al. 2018). Besides, these commercial drivers are likely to focus more on finding passengers instead of driving safely during the daytime to make more profits. Thus, pedestrians and street hawkers who use these portions of the roadway are likely to be hit by these vehicles.

Rules generated for injured-but-not-hospitalized/slight injury crashes showed strong associations with items such as {Collision type: run-off-road/overturn}, {Roadworks in progress: no}, {Traffic control: other}, and {Number of vehicles involved: one vehicle}. The most interesting rule among those generated (rule 3.1) shows that run-off/overturn crashes occurring at the main road segments with no traffic signals are likely to have lower severity levels. Moreover, rules 3.2 and 3.3 highlight that these crashes usually occur at median-unseparated roadway segments and those with no roadworks in progress. Besides, from rule 3.9, single-vehicle run-off crashes are likely to cause minor injuries requiring no hospitalization.

Finally, PDO crashes on poor pavement roads during the daytime were mainly linked with roads with no shoulders, straight and flat pavements, and parked vehicles (rule 4.1). Rules 4.3, 4.5, and 4.1 depict that clear weather, main roadway segments without shoulders, and median separations are likely candidates for PDO bus/minibus-involved crashes during the day. Again, these rules are plausible as bus/minibus drivers usually park improperly on the road looking for passengers (not at bus stops). These behaviors are likely to result in PDO crashes such as sideswipes.

Rules for PNLO bus/minibus-involved crashes

Key items found to lead to fatal crashes on PNLO roads include {Collision type: hit-pedestrian} and {Vehicle defects:

none} (Table C2). From rule1.1, hit-pedestrian collisions involving non-defective vehicles during clear weather were primarily fatal when the crash occurred on a roadway with poor pavement during the night with lights on (PLNO). The third rule shows strong associations between variables such as hit-pedestrian crashes, going ahead maneuvers, and straight and flat roads with fatal crashes. Rules 1.6, 1.9–1.10 reveal that single-vehicle-involved hit-pedestrian crashes and non-DUI variable are associated with fatal crashes on PNLO roads.

The first rule under the severe/hospitalized category, {Collision type: hit-pedestrian, Driver's age: mid-30–49 yrs., Vehicle damage: extensive} $\Longrightarrow$ {Severity level: hospitalized}, shows that vehicle-pedestrian crashes caused by drivers aged 30–49 that produces extensive vehicle damage are likely to result in severe injuries. Rules 2.2 and 2.3 show that most hit-pedestrian collisions that are not hit-and-run and occur at shoulder-absent segments are also likely to have severe injury outcomes. Rule 2.8 depicts that those hit-pedestrian collisions during the weekends at median-unseparated roadway sections are also highly associated with severe consequences.

The rules generated with minor injured-but-not-hospitalized/minor injury outcome as a consequence reveal items such as {Driver error: none}, {Shoulder type: no shoulder}, {Driver's age: young <30 yrs.}, {Roadway segment: main road} and {Road description: straight and flat} are persistent antecedents. As displayed in rule 3.1, the support value of 0.061 shows that 6.1% of the bus/minibus crashes on straight and flat roads with no shoulders resulted in minor injuries. Rule 3.5 also highlighted that non-driver error crashes occurring at shoulder-absent and uncontrolled roads highly correlate with minor injury severities.

Finally, rules with PDO crash outcomes were assessed (rules 4.1–4.10). Strong associations between PDO and items such as {Road description: straight and flat}, {Weather: clear}, {Number of vehicles involved: two vehicles}, and {Traffic control: none} were found. The rule {Number of vehicles involved: two vehicles, Road description: straight and flat, Weather: clear} $\Longrightarrow$ {Severity level: PDO} has a lift value of 1.671, indicating a strong correlation between the variables, highlighting the need for ensuring attentiveness among drivers while cruising on straight roads.

Rules for PNLOFF bus/minibus-involved crashes

The first rule in this category outlined in Table C3 has antecedent {Collision type: hit-pedestrian, Driver's age: mid-30–49 yrs., Vehicle damage: minor}, depicting that vehicle-pedestrian crashes involving mid-aged drivers with minor vehicle damage are likely fatal if they occur on poor and dark-unlighted roads. Besides, the second rule shows that 3.1% of the hit-pedestrian crashes that occurred during the weekdays on roadway segments with tarred pavements are also highly likely to be fatal. Rules 1.3–1.4 show that most crashes mainly occur on straight and flat roads with no traffic control. We identified that items namely {Vehicle maneuver: going ahead}, {Vehicle damage: minor}, and {Collision type: hit-pedestrian} were most common among the rules generated.

Severe injury crash outcomes observed on PNLOFF roads are generally associated with items such as {Weather: clear}, {Median: present}, {Vehicle damage: extensive}, and {Collision type: hit-pedestrian}. The rule with the highest lift value (rule 2.1) reveals that those crashes involving mid-aged drivers on median separated PNLOFF roads are associated with severe outcomes. Besides, the rule with the second-highest lift value (rule 2.2) suggests that two-vehicle-involved crashes at curved sections of the roadway are likely to have severe outcomes. Rules 2.5 and 2.9 include items relating to hit-pedestrian crashes involving younger drivers and minibuses. Since median separated roads typically have higher speed limits, drivers are likely to overspeed and engage in risky driving maneuvers during the night when the weather is clear. Besides, drivers may lose control of their buses/minibuses when they drive fast at curved sections in the night with low visibility – thus, leading to crashes which can have serious consequences. Our results align with previous studies that identified that median separations, curved sections, and poor visibility significantly increase the severity of nighttime crashes (Liu et al. 2019; Zhang and Hassan 2019).

Rule 3.1 reveals that run-off/overturn crashes involving minibuses at roadway segments are associated with injured-but-not-hospitalized/minor injury outcomes. Most importantly, rule 3.2 shows that those minibus-involved crashes involving mid-aged drivers who committed no errors are likely to result in minor injury. A previous study also identified that, relative to mid-aged drivers (aged 35–50), younger drivers (age <35) have a lower chance of being involved in minor injuries crashes at night. The authors mentioned that mid-aged drivers are likely to be involved in minor injury crashes due to excessive mental stress and reduced sight/ability relative to younger drivers (Zhang and Hassan 2019). Further, rules 3.3–3.5 show that minor crash outcomes are expected when run-off/overturn crashes occur at inclined roadway segments even when the vehicle has no defects.

Concerning rules for PDO crashes that occurred PNLOFF, items such as {Collision type: sideswipe}, {Road description: curve and incline}, and {Shoulder type: Present – tarred} were frequent. A careful look at the rule with the highest support reveals that 1.3% of rear-end collisions during the weekday on PNLOFF roads are associated with PDO outcomes (rule 4.8).

Discussion

Identifying the groups of factors affecting the severity of crashes in DC's has become necessary, given the increased number of casualties recorded in these jurisdictions. Although a substantial number of daily trips in low- and middle-income countries are made by buses/minibuses, and the severity outcomes associated with them are significantly high (Barua and Tay 2010; Sam, Daniels, et al. 2018), research directed toward identifying factors affecting the severity of bus/minibus crash is limited. Among the few studies from DC's, research focused on using more robust tools for assessing the contributory factors leading to each crash severity outcome while considering the different combinations of pavement and lighting conditions has not been explored. Given that most developing nations have infrastructural challenges, we posit that understanding the hidden groups of factors leading to each crash severity outcome based on the different combinations of poor roadway pavement and lighting conditions would provide valuable insights into which groups of factors can be controlled to improve traffic safety in such nations.

This study's main contribution to the bus/minibus safety literature is that it is the first to employ the robust nonparametric data mining technique (ARM) to extract meaningful relationships in bus/minibus crash dataset considering the mechanism underpinning the simultaneous effect of both pavement and lighting conditions in a DC. In particular, this study is the first to discover the combination of risk factors that are likely to influence the severity of crashes on poor roads given different lighting conditions. The most significant strength in the selected methodology is that it has superior performance regarding prediction accuracy and the ability to handle big data more effectively without employing predetermined assumptions, as does its parametric counterparts. The study results would deepen the understanding of how crash-risk factors influence the severity levels of bus/minibuses in DC's and guide in making appropriate policies targeted at ensuring the safety of pedestrians.

Upon segregating the data based on the lighting condition, certain interesting discoveries were made. Hit-pedestrian crashes were likely to be fatal on dark roads (lighted or unlighted) and severe during daylight conditions. Regarding crashes in the PD category, hit-pedestrian, shoulder presence, and driver inattentiveness were key factors leading to severe crashes. For crashes on PNLO roads, the hit-pedestrian variable was also common among the rules for severe crashes. However, they were likely to occur on roads with no shoulders. Further, PNLO roads with no shoulders, no medians, and no traffic control were associated with minor injuries. Finally, PNLOFF crashes associated with curved sections, median and shoulder presence, and driver errors were mostly severe. Also, weekend crashes were more likely to be severe, and they mainly occurred on PNLOFF roadways.

Detailed discussions about the results obtained from this study are presented in Text D1 in Appendix D of the supplementary material. Based on the key findings of the study, the following countermeasures are recommended to improve bus/minibus safety are suggested.

- The finding that the hit-pedestrian variable occurs frequently among groups of risk factors leading to fatal/severe crashes during the day and night is indicative of high pedestrian-traffic conflicts on the roadways. These crashes were found to be likely to occur in "going ahead" maneuvers on straight main roads where drivers are likely to relax and overspeed due to the elevated safety perception of driving on straight (no curve) roads (Hong et al. 2020a). DC's lack proper pedestrian road space planning, and the design of infrastructure for pedestrians (Mukherjee and Mitra 2021). Thus, improving the built roadway environment through the implementation of pedestrian priority street projects which focuses on enhancing walkability and providing enough infrastructure for pedestrians (overhead pedestrian crossings, pedestrian guardrails, and pedestrian refuge islands) would effectively separate traffic from pedestrians and help minimize the hit-pedestrian crash fatalities. In the case of Seoul, South Korea, these projects were effective at clearly separating pedestrians from vehicles, reducing vehicle speeds where necessary, and increasing the overall safety of pedestrians (Lee and Kim 2019).
- Inattentiveness on the part of drivers and no traffic control were found to be critical causes of hit-pedestrian crashes during daylight conditions on poor roads. This finding is plausible as bus/minibus drivers mostly disregard safety rules and actively focus on finding new passengers as they drive (Boateng 2020). The results also show that hit-pedestrian crashes mostly occur at shoulder present segments during the day, which are the only spaces where pedestrians walk. This supports the idea that inattentiveness of drivers and disregard to road safety rules could lead to hitting pedestrians (mostly street hawkers; see Figure B1) who are likely to be using the road's shoulders (Damsere-Derry et al. 2019). While the provision of adequate pedestrian protection infrastructure could help address this issue, punishment of reckless drivers, education through imbibing the pedestrian-first culture in drivers and other roadway users, and enacting policies that take hawkers off the would be an important step toward addressing the bus/minibus-pedestrian crash problem. Pedestrians are part of the most vulnerable road users – thus, when drivers learn to give priority to pedestrians, particularly on roads with no traffic control, several preventable crashes could be avoided. Based on the case of South Korea, the number of vehicle-pedestrian/hawker interactions could reduce drastically, leading to a reduced chance of pedestrian-involved crashes (Lee and Kim 2019).
- Run-off road and incline/curved road indicators were found to be common factors in evident injury crashes particularly on dark roads with no lighting. This finding is likely attributable to the poor visibility on the unlit roads, the poor pavement condition, and the impatience of these drivers causing them to overspeed and engage in improper turning and overtaking instead of concentrating on safely maneuvering these curved/inclined segments (Sam, Daniels, et al. 2018; Boateng 2020). Increasing traffic surveillance and control of public transport driver behavior through the strict enforcement of traffic regulations would ensure that drivers comply with the roadway code. There is the need to provide an adequate budget for the provision of lighting facilities on roadways to help improve visibility at night. Legible roadway signs coated with retroreflective materials should be placed at vantage points to inform drivers of impending danger at curved sections. Further, to improve the safety at these sections, visual guidance in

the form of longitudinal road markings should be considered. Rolling guardrails could also be installed to reduce the impact on vehicle-facility collision at curved sections. Besides, it would also be worthwhile to consider treatments such as rumble strips placed at both sides of the centerline and the road's shoulders to guide the driver's lateral position at curved sections and alert inattentive drivers of potential danger. In France, these interventions were found to have a net positive effect on reducing run-off road crashes in terms of increasing visibility and controlling driver's trajectory (Rosey et al. 2008).
- The finding that drivers aged between 30 to 49 years were mostly involved in fatal/severe injuries during the night irrespective of the lighting condition is plausible as, compared to their younger counterparts, those mid-aged drivers may have reduced eyesight and increased work pressure leading to overwork and exhaustion. Besides, younger drivers (<30 years) were mostly likely to be involved in minor injury crashes. The remuneration system of the owners of these commercial vehicles (low wages and high level of job insecurity) puts pressure on these drivers who are mostly the breadwinners of their families to overwork, placing them and their passengers at risk of crashes (Peltzer 2011; Boateng 2020). As previous studies have shown that allowing the driver to have break times for resting reduced the risk of crashes (Garbarino et al. 2016; Kwon et al. 2019), enforcing regulations that guarantee that bus/minibus drivers have adequate resting periods and instituting management measures geared at checking and improving driver mental health and work conditions would ensure that drivers drive attentively in the right state of mind and reduce the incidence of drowsy driving. It would be worthwhile to consider installing digital tachographs for monitoring abnormal bus driver behavior and time of vehicle operation. These drivers could be identified for further training, education, or punishment. Evidence from South Korea, where legislation requires that these devices be installed on all commercial vehicles, showed positive results in that the data from this device was used for identifying and targeting drivers for interventions such as safety education (Kim et al. 2016).
- The results of the study identified that, during daylight conditions, shoulder absence was found to be a critical factor causing hit parked vehicle crashes with no injuries. Besides, median absence also was likely to lead to fatal crashes. Similar to other studies, hit parked vehicles involving buses are less serious (Barua and Tay 2010). Again, median presence reduces the risk of head-on collisions with oncoming traffic which would be otherwise severe (Smith et al. 2021). Hit parked vehicle crashes are likely to be a result of aggressive driving and illegal on-street parking during the day due to the increased traffic volume, absence of road shoulders, and indiscipline on the part of drivers (refusing to park at approved bus stops but preferring to stop indiscriminately to pick or search for passengers) which is common in DC's. These crashes are more likely to result in a large quota of side-swipes (Barabino et al. 2021). In the long term, this issue can be addressed by expanding the roadways, providing adequate bus stops and shoulders to improve the maneuverability of buses. In the short term, it is necessary to increase physical police presence to reduce the menace of illegal on-street parking.

We acknowledge that, even though the problems in DC's are not uniform, this study could be used as a basis for the improvement of bus/minibus safety in low- and middle-income countries. The study identified risk factors that influence the severity of bus/minibus crashes on poor roads. Thus, besides the suggested countermeasures, policymakers can take measures that focus on eliminating the effect of variables identified to affect crash severity. The limitations of this study are primarily related to the quality of the data. As with all crash datasets, there were some missing data and underreporting issues regarding the dataset used for this analysis. Besides, the location of the crashes (rural/urban) was not provided in the data. Thus, this information could not be included in the analysis. In the future, it would be worthwhile to include this information for a deeper understanding of risk factors influencing bus safety in DC's. Finally, in ensuring that stronger rules are selected, some studies employed the lift information criterion (LIC) condition following the traditional support, confidence and lift thresholds (Montella et al. 2020; Samerei et al. 2021). However, we used only the traditional measures in this study while ensuring that rules with higher lift values were selected. Although our study compares well with others conducted by renowned researchers in the transport safety field (Pande and Abdel-Aty 2009; Das and Sun 2014; Verma et al. 2014; Weng et al. 2016; Das et al. 2018, 2019; Yu et al. 2019; Hong et al. 2020b), we plan to incorporate the LIC thresholds future studies.

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable comments, which helped improve the quality of the article.

Disclosure statement

None of the authors have any relevant competing interests to declare.

Funding

The authors report that there is no funding associated with the work featured in this article.

ORCID

Reuben Tamakloe http://orcid.org/0000-0002-9605-8439
Enoch F. Sam http://orcid.org/0000-0002-2514-2959
Madiha Bencekri http://orcid.org/0000-0002-0456-5880
Subasish Das http://orcid.org/0000-0002-1671-2753
Dongjoo Park http://orcid.org/0000-0002-6330-4437

References

Agrawal R, Imieliński T, Swami A. 1993. Mining association rules between sets of items in large databases. ACM SIGMOD Rec. 22(2): 210–216. doi:10.1145/170035.170072

Barabino B, Bonera M, Maternini G, Olivo A, Porcu F. 2021. Bus crash risk evaluation: an adjusted framework and its application in a real network. Accid Anal Prev. 159:106258. doi:10.1016/j.aap.2021.106258

Barua U, Tay R. 2010. Severity of urban transit bus crashes in Bangladesh. J Adv Transp. 44(1):34–41. doi:10.1002/atr.104

Boateng FG. 2020. Indiscipline in context: a political-economic grounding for dangerous driving behaviors among Tro-Tro drivers in Ghana. Humanit Soc Sci Commun. 7(1):8. doi:10.1057/s41599-020-0502-8

Damsere-Derry J, Ebel BE, Mock CN, Afukaar F, Donkor P, Kalowole TO. 2019. Evaluation of the effectiveness of traffic calming measures on vehicle speeds and pedestrian injury severity in Ghana. Traffic Inj Prev. 20(3):336–342. doi:10.1080/15389588.2019.1581925

Das S, Dutta A, Avelar R, Dixon K, Sun X, Jalayer M. 2019. Supervised association rules mining on pedestrian crashes in urban areas: identifying patterns for appropriate countermeasures. Int J Urban Sci. 23(1):30–48. doi:10.1080/12265934.2018.1431146

Das S, Dutta A, Jalayer M, Bibeka A, Wu L. 2018. Factors influencing the patterns of wrong-way driving crashes on freeway exit ramps and median crossovers: exploration using 'Eclat' association rules to promote safety. Int J Transp Sci Technol. 7(2):114–123. doi:10.1016/j.ijtst.2018.02.001

Das S, Sun X. 2014. Investigating the pattern of traffic crashes under rainy weather by association rules in data mining. In: Transportation research board 93rd annual meeting. Washington DC: Transportation Research Board.

Das S, Sun X, Goel S, Sun M, Rahman A, Dutta A. 2020. Flooding related traffic crashes: findings from association rules. J Transp Saf Secur. 14(1):111–129.

Garbarino S, Lanteri P, Durando P, Magnavita N, Sannita W. 2016. Co-morbidity, mortality, quality of life and the healthcare/welfare/social costs of disordered sleep: a rapid review. IJERPH. 13(8):831. doi:10.3390/ijerph13080831

Guobadia EK. 2021. Statistical application of regression techniques in modeling road accidents in Edo State, Nigeria. Sch J Physics, Math Stat. 8(1):14–18.

Hong J, Tamakloe R, Park D. 2020a. Application of association rules mining algorithm for hazardous materials transportation crashes on expressway. Accid Anal Prev. 142:105497. doi:10.1016/j.aap.2020.105497

Hong J, Tamakloe R, Park D. 2020b. Discovering insightful rules among truck crash characteristics using apriori algorithm. J Adv Transp. 2020:1–16. doi:10.1155/2020/4323816

Kaplan S, Prato CG. 2012. Risk factors associated with bus accident severity in the United States: a generalized ordered logit model. J Safety Res. 43(3):171–180. doi:10.1016/j.jsr.2012.05.003

Kim D-G, Lee C, Park B-J. 2016. Use of digital tachograph data to provide traffic safety education and evaluate effects on bus driver behavior. Transport Res Rec. 2585(1):77–84. doi:10.3141/2585-09

Kwon S, Kim H, Kim GS, Cho E. 2019. Fatigue and poor sleep are associated with driving risk among Korean occupational drivers. J Transp Heal. 14:100572. doi:10.1016/j.jth.2019.100572

Lee T, Kim T. 2019. Shared space and pedestrian safety: empirical evidence from pedestrian priority street projects in Seoul, Korea. Sustainability. 11(17):4645. doi:10.3390/su11174645

Liu J, Li J, Wang K, Zhao J, Cong H, He P. 2019. Exploring factors affecting the severity of night-time vehicle accidents under low illumination conditions. Adv Mech Eng. 11(4):168781401984094. doi:10.1177/1687814019840940

Montella A. 2011. Identifying crash contributory factors at urban roundabouts and using association rules to explore their relationships to different crash types. Accid Anal Prev. 43(4):1451–1463. doi:10.1016/j.aap.2011.02.023

Montella A, de Oña R, Mauriello F, Rella Riccardi M, Silvestro G. 2020. A data mining approach to investigate patterns of powered two-wheeler crashes in Spain. Accid Anal Prev. 134:105251. doi:10.1016/j.aap.2019.07.027

Mukherjee D, Mitra S. 2020. A comprehensive study on identification of risk factors for fatal pedestrian crashes at urban intersections in a developing country. Asian Transp Stud. 6:100003. doi:10.1016/j.eastsj.2020.100003

Mukherjee D, Mitra S. 2021. Investigating the fatal pedestrian crash occurrence in urban setup in a developing country using multiple-risk source model. Accid Anal Prev. 163:106469. doi:10.1016/j.aap.2021.106469

Pande A, Abdel-Aty M. 2009. Market basket analysis of crash data from large jurisdictions and its potential as a decision support tool. Saf Sci. 47(1):145–154.

Peltzer K. 2011. Road use behavior in Sub-Saharan Africa. In: Handbook of traffic psychology. San Diego, CA: Elsevier; p. 503–518.

Rosey F, Auberlet J-M, Bertrand J, Plainchault P. 2008. Impact of perceptual treatments on lateral control during driving on crest vertical curves: a driving simulator study. Accid Anal Prev. 40(4):1513–1523. doi:10.1016/j.aap.2008.03.019

Sam EF, Brijs K, Daniels S, Brijs T, Wets G. 2018. Public bus passenger safety evaluations in Ghana: a phenomenological constructivist exploration. Transp Res Part F Traffic Psychol Behav. 58:339–350. doi:10.1016/j.trf.2018.06.031

Sam EF, Daniels S, Brijs K, Brijs T, Wets G. 2018. Modelling public bus/minibus transport accident severity in Ghana. Accid Anal Prev. 119:114–121. doi:10.1016/j.aap.2018.07.008

Samerei SA, Aghabayk K, Mohammadi A, Shiwakoti N. 2021. Data mining approach to model bus crash severity in Australia. J Safety Res. 76:73–82. doi:10.1016/j.jsr.2020.12.004

Smith J, Hosseinpour M, Mains R, Hummel N, Haleem K. 2021. Investigating head-on crash severity involving commercial motor vehicles in Kentucky. Transport Res Rec. 2675(10):133–147. doi:10.1177/03611981211010803

Tamakloe R, Das S, Nimako Aidoo E, Park D. 2022. Factors affecting motorcycle crash casualty severity at signalized and non-signalized intersections in Ghana: insights from a data mining and binary logit regression approach. Accid Anal Prev. 165:106517. doi:10.1016/j.aap.2021.106517

Tamakloe R, Hong J, Park D. 2020. A copula-based approach for jointly modeling crash severity and number of vehicles involved in express bus crashes on expressways considering temporal stability of data. Accid Anal Prev. 146:105736. doi:10.1016/j.aap.2020.105736

Verma A, Khan SD, Maiti J, Krishna OB. 2014. Identifying patterns of safety related incidents in a steel plant using association rule mining of incident investigation reports. Saf Sci. 70:89–98. doi:10.1016/j.ssci.2014.05.007

Wanvik PO. 2009. Effects of road lighting: an analysis based on Dutch accident statistics 1987–2006. Accid Anal Prev. 41(1):123–128. doi:10.1016/j.aap.2008.10.003

Weng J, Zhu JZ, Yan X, Liu Z. 2016. Investigation of work zone crash casualty patterns using association rules. Accid Anal Prev. 92:43–52. doi:10.1016/j.aap.2016.03.017

Yu S, Jia Y, Sun D. 2019. Identifying factors that influence the patterns of road crashes using association rules: A case study from Wisconsin. United States. Sustain. 11(7):1–14.

Zhang K, Hassan M. 2019. Crash severity analysis of nighttime and daytime highway work zone crashes. PLoS One. 14(8):e0221128. doi:10.1371/journal.pone.0221128