Pedestrian Safety Analysis at Urban Midblock Section under Mixed Traffic Conditions Using Time to Collision as Surrogate Safety Measure

Hareshkumar Golakiya\textsuperscript{1}, Ritvik Chauhan\textsuperscript{2}, Chintaman Bari\textsuperscript{3} and Ashish Dhamaniya\textsuperscript{4}

\textsuperscript{1}Assistant Professor, Department of Civil Engineering, Dr. S. & S. Ghandhy Government Engineering College, Surat, INDIA, hdgolakiya@gmail.com

\textsuperscript{2}Research Scholar, Department of Civil Engineering, Sardar Vallabhbhai National Institute of Technology, Surat-395007, INDIA, ritvikdchauhan@gmail.com

\textsuperscript{3}Research Scholar, Department of Civil Engineering, Sardar Vallabhbhai National Institute of Technology, Surat-395007, INDIA, chintamanbari@gmail.com

\textsuperscript{4}Associate Professor, Department of Civil Engineering, Sardar Vallabhbhai National Institute of Technology, Surat-395007, INDIA: adhamaniya@gmail.com
Abstract

Pedestrians are the most vulnerable road users, and pedestrian safety has become a major concern of researchers in recent years due to the increasing number of fatalities on roads. Conflict analysis using surrogate safety measures (SSMs) are hence a helpful technique to study pedestrian safety, as there are many limitations with collision data. Moreover, it is a cost-effective technique compared to historical crash data analysis. The present paper analyse pedestrian safety at urban midblock crosswalks using Time to Collision (TTC) as a SSM. The data for the present study were collected from four different midblock pedestrian crossing locations of different cities located in western parts of India using the videographic technique. The trajectory of pedestrians and vehicles are extracted for micro-level analysis of pedestrian-vehicle interactions. The trajectory data are further used for calculating TTC at regular time intervals during the interaction of pedestrians and vehicles. Two different types of pedestrian road crossing behaviour, known as vehicle-pass-first and pedestrian-pass-first, are identified, and TTC analysis has been done differently for each scenario. The variation of TTC based on gender and vehicle category is analysed to evaluate the influence of such parameters on pedestrian safety. The generalised linear mixed model approach is used to develop linear regression models for TTC based on the empirical data. The threshold values for TTC are used to define various safety levels of pedestrians using a clustering approach.

Keywords: Pedestrian, Safety, Midblock, Time to Collision
1. Introduction

A crosswalk is an essential facility that provides a bridge between activities on either side of the road and is frequently used by pedestrians. In developing countries like India, crossing treatments such as road marking, signboard, and signals are mostly absent or disregarded by vehicle users if present. Also, the pedestrian shows unsafe behaviour while crossing the road when compared to walking on a sidewalk at such sections. Such risk-taking behaviour leads to higher accident rates in developing counties for pedestrians. Many studies highlight this scenario and show that the casualty rates of pedestrians are much higher in developing countries as compared to developed countries. In India, it is prevalent that pedestrians cross the road at undesignated midblock sections and are so habituated that they avoid using the grade-separated facility even if it is available. On the other side, vehicle drivers do not give way to pedestrians even at a marked crosswalk facility. Studies have shown that urban areas account for 60% of pedestrian fatalities and 85% of these fatalities occur at midblock crosswalks. Pedestrian crossing operation is fundamentally based on the gap acceptance process. After arriving at the kerb or median, a pedestrian examines the gap and accepts or rejects the vehicular gap based on his perception, speed, approaching vehicle class, behaviour, and experience. Mixed traffic conditions are predominant in developing countries which further complicates the process of road crossing or pedestrians. Any mistake or misjudgement committed by a vehicle driver or pedestrian may result in a collision among these road users.

Historical crash data is one of the techniques to evaluate pedestrian safety at crosswalk locations under mixed traffic conditions. However, due to many issues related to the quality and quantity of such data, the use of traffic conflict technique has been introduced and has been established for vehicle-vehicle conflict analysis. The traffic conflict technique is also used in pedestrian-vehicle conflict analysis and is also focused on by many researchers. The traffic
conflict technique can recognise close-miss chances of a conflict. During the traffic operation, some of the conflicts may result in a collision, causing a fatality or severe to non-severe injury. The traffic conflict technique can identify the severity of the pedestrian-vehicle conflict. Time to collision (TTC) is such a traffic conflict technique and can evaluate pedestrian safety. TTC is defined as “the time required for two entities to collide if they continue at their present speed on the same path” (11,12). Many researchers used TTC or TTC based surrogate safety measures in vehicle-vehicle and pedestrian-vehicle interaction analysis (6). The decrement in TTC results in an increase in pedestrian-vehicle interaction and further increases the probability of conflict with vehicles. The present study aims to evaluate the safety of pedestrians at midblock crosswalks under mixed traffic conditions using TTC as a surrogate safety measure.

2. Literature review

The studies related to pedestrian safety assessment are mainly carried out by crash based methods such as historical crash data, users’ perception survey, or conflict technique. Many researchers attempted to study the impact of built environment on pedestrian safety based using historical crash data (13-16). Some of the researchers focused on identifying the factors affecting pedestrian safety using historical crash data (17-19). Many authors adopted the conflict based method to study the impact of pedestrian characteristics like age, gender, etc. on the safety of pedestrians during crossing manoeuvre and attempted to develop the model for safety (20-22). Users’ perception-based investigation on effect pedestrian characteristics like age gender on the safety of pedestrians has been focussed by many researchers (23-25). Many researchers adopted a driving simulator to study the safety of pedestrian and their behaviour. Wu et al. (26) used the driving simulator and designed a full factorial experiment to study pedestrian-vehicle conflict, which includes four different potential risk factors; the time of day, crosswalk
marking, roadway type, and pedestrian dressing colour. Chrysler et al. 27 used a driving simulator to examine the driver’s response to a crash-imminent situation involving a pedestrian.

Many researchers emphasised on traffic conflict technique and used different surrogate safety measures (SSM) to analyse pedestrian safety at a midblock crosswalk or intersection because of the qualitative and quantitative issues with road collision data. Kaparias et al. 28 presented a new vehicle-pedestrian conflict analysis technique based on existing vehicle-vehicle techniques for conventional roads and in shared-space environments.

Zhang et al. 29 compared one hundred groups of pedestrian and vehicle interactions based on vehicle pass first (VPF) and pedestrian pass first (PPF) cases among different safety scenes and introduced a new parameter called time difference to collision (TDTC) related to safety. Zhang et al. 30 adopted the TDTC parameter as a variation from TTC and post-encroachment time (PET). The interaction behaviour between pedestrians and vehicles were analysed and validated for the TDTC parameter indicating pedestrian safety, to co-relate the pedestrian-involved potential collisions and conflicts. Zhang et al. 31 developed a scene-based pedestrian safety performance evaluation model. Alhajyaseen and Iryo-Asano 32 developed a multinomial logit model to evaluate the probability of a pedestrian suddenly varying the speed as a function as one of the various factors affecting the safety of the pedestrian at signalised crosswalks. Hagiwara et al. 33 investigated conflicts between the right-turning vehicle and the pedestrian coming from the right (left-turning in the case of USA) on the crosswalk based on time lag. Ismail et al. 34 attempted to extract conflict indicators from an automated video analysis system that can calculate four severity conflict indicators fully automatic way at Vancouver, British Columbia. Ismail et al. 35 used video data for automated analysis for the safety evaluations and demonstrated the feasibility of conducting before-and-after (BA) scenario safety evaluations. Zheng et al. 36 explored pedestrian jaywalking behaviour and corresponding driver yielding behaviour to model vehicle-pedestrian behaviour outside the
crosswalk using the micro-simulation approach at the USA. Lorion and Persaud \(^{37}\) proposed a model to predict crash prediction based on two SSM, namely conflicts and delay and evaluate their predictively at urban intersections in Canada. Ni et al. \(^{38}\) used trajectories to assess safety by paying more attention to behavioural factors, which consider pedestrian-vehicle interaction. The authors suggest the concept of three interaction patterns using a support vector machine (SVM) in China. Chen et al. \(^{39}\) applied two SSMs of PET and relative time to collision (RTTC), characterising how spatially and temporally close the pedestrian-vehicle conflict is to a collision using the aerial sensor named Unmanned Aerial Vehicles (UAVs) urban intersection in Beijing, China. Madhumita and Ghosh \(^{40}\) proposed a new methodology over the existing one using two proximal safety indicators, PET and conflicting speed of through moving vehicles on major roads. Authors found that PET values less than the threshold do not always create critical situations when the speed of the corresponding conflicting vehicle is low and vice-versa. Babu and Vedagiri \(^{41}\) used two surrogate measures of PET and the corresponding speed of the conflicting vehicles to analyse the traffic conflict at an intersection in India. The authors used the required deceleration rate to categorise the conflicts.

Researchers also focused on the safety of pedestrians during the crossing at the midblock section. Jiang et al. \(^{5}\) estimated the differences between TTC and TTC related parameters between China and Germany using road user trajectory. Chen and Wang \(^{42}\) proposed a cellular automata (CA) model to simulate the interaction between vehicle flow and pedestrian crossing. Traffic parameters related to pedestrian and vehicle flow were investigated in China. Cafiso et al. \(^{6}\) carried out a before-after analysis to assess the safety performance of newly installed traffic calming devices using Pedestrian Risk Index (PRI) as SSM at the urban midblock section in Spain. Chandrappa et al. \(^{43}\) examined pedestrian-related safety facets in the urban road in India by assessing the PET and the threshold wait time (TWT) for pedestrians during the crossing. Kadali and Vedagiri\(^{44}\) used pedestrian safety margin (PSM) examined the
pedestrian safety at unprotected midblock crosswalks in India. Further, the authors carried out regression and develop a binary logit model to find the factors influencing to the PSM. The developed models can predict the probability of avoiding conflict with an approaching vehicle at unprotected midblock crosswalks. Chen et al. 45 applied evolutionary game theory and cumulative prospect theory to consider the decision process of vehicle drivers and pedestrians during the interaction for addressing the crossing decision behaviour under bounded rationality and risk. Rankavat and Tiwari 46 studied the perception of risk for identifying the potential crash risk for Indian mixed traffic conditions. The results showed that the four-legged intersections below the flyovers were the critical locations of risk. Pawar and Patil47 found that the critical gaps for the pedestrians’ crossings at uncontrolled mid-block sections are found to be lower than the Highway Capacity Manual (HCM). The investigation by Chaudhari et al. 48 found that the average value of 6.2 seconds is recommended for designing the crossing facility with pedestrian safety. In another study by Chaudhari et al. 49 a multilinear regression model is developed for evaluating pedestrian safety margin for Indian traffic conditions for crossing pedestrians. Chen and Fan 50 developed a multinomial logit model of pedestrian-vehicular crash severity and found that the drivers physical condition, vehicle category such as motorcycle and trucks, pedestrian age, etc., are the most significant factors causing the crashes. Danaf et al. (2020) 51 developed a methodology for finding out the interactions of pedestrians with vehicles in the mixed traffic conditions in the presence of a crosswalk. The results showed that the presence of a crosswalk decreases the pedestrian waiting time and reduces the speed of vehicles in the section before the crosswalks. Golakiya et al. 52 suggested distance based safety index namely safe distance (SD) at urban midblock crosswalk under mixed traffic environment and threshold value for safe crossing was suggested. Golakiya and Dhamaniya 53 adopted as SSM to evaluate safety of crossing pedestrians. Safety index threshold value was developed for two
separate cases namely, PPF and VPF for different category of vehicles based on vehicle speed as a variable.

The literature review presented above reveals that the safety of pedestrians is a major concern of researchers. Researchers used different methods to evaluate pedestrian safety such as historical data, 'users’ perception, conflict technique method, and simulation-based method. However, the approach based on conflict technique is more rational and cost-effective. Researchers used various surrogate safety parameters to evaluate pedestrian safety. However, the majority of studies are carried out at intersections. Some works have also been reported at the midblock location to study vehicle-pedestrian interaction. However, a few studies have been reported for heterogeneous traffic without lane discipline, similar to the Indian condition. Hence, further research in this direction can be useful to focus the pedestrian safety. None of the reported studies have modelled the SSM parameters for mixed traffic conditions. The present study is carried out in this direction to analyse the safety of the pedestrian and model the TTC at unprotected urban midblock crosswalks under mixed traffic conditions.

3. Research objectives

Pedestrian crossing at urban midblock is a common phenomenon observed in developing countries. This pedestrian crossing is a compound aspect and has profound safety implications. The prime objective of the study is to examine the safety aspects of crossing pedestrians using the TTC as the surrogate safety measure and to model TTC by generalised linear model approach. Moreover, to define the threshold value of TTC using a clustering technique to categorise pedestrian risk.
4. Site selection

To meet the objectives of the present study, four different locations at uncontrolled (no right of way to pedestrian) midblock sections were selected on six-lane urban arterials. These sections were chosen in four different cities (Surat, Vadodara, Ahmedabad, and Jaipur) in the western part of India. The criteria for selecting the section are that it should be free from side frictions other than crossing pedestrians such as on-street parking, stopped vehicles, hawkers, curb-side bus-stop, etc. The sections not should be under the influence of intersection or grade. Moreover, the sections should have uniform geometry. The selected survey locations contain diverse traffic volumes, motor vehicle speeds, pedestrian crossflow, and pedestrian behavioural characteristics. At some of the locations, zebra crossing is provided for pedestrian crossing. However, it was observed that vehicle drivers rarely give way to pedestrians. Hence, the pedestrian road crossing operation is the same as the undesignated pedestrian crossing section as shown in Fig. 1. Hence, the data collection was done at locations without any designated crosswalks and also the locations where crosswalks were present, but the road markings for crosswalks are completely faded and there is an absence of traffic signs and signals informing the vehicular traffic of a crosswalk. This diverseness of pedestrian as well as vehicular characteristics is appropriate to obtain a wide range of TTC values and is useful to develop the generic model.

[Figure 1 here]

5. Data collection

The data collection program supplementing the objectives of the study was conducted at the selected midblock sections using videographic survey with the help of a camera of high resolution and magnification. The data was collected on a dry weather day from morning 7:00 AM to evening 7:00 PM which includes morning and evening peak hours and off-peak hours to ensure the safety assessment of pedestrian at all possible vehicular and pedestrian flow. To
record the simultaneous movement of both vehicles and pedestrians, the camera was installed on a 15 m high vantage point. During the videographic survey, the marking was done on the roadway at regular intervals to prepare the grid based on the real dimension (Fig. 2(b)).

6. Data extraction and trajectory plotting

To study pedestrian safety, it is necessary to study the vehicle and pedestrian movement at the micro-level, and hence, the trajectory approach is adopted in the present study. The trajectory data ensures the in-depth study of pedestrian-vehicle interaction to examine the safety of pedestrians. The trajectories of the crossing pedestrian(s) and approaching vehicles were plotted using a two-dimensional coordinates system. A grid of size 50-m X 10.5-m was generated using AutoCAD 2016 software by importing the study location image containing the marking done during data collection in the software (Fig. 2(a)). The size of blocks in the grid was kept 1.25-m X 1.25-m. The grid image was overlaid over the captured video, such that the grid accurately and exactly fit over study location video using the Ulead Video Studio 11 software, as shown in Fig. 2(b). The overlaid video replayed on a large screen monitor in the laboratory using AVIDEMUX 2.6 software. The software can convert every 1 second of the video in 25 frames, i.e., capture a frame after every 0.040-s. The exact position of crossing pedestrian after every 0.48-s in the grid was observed and recorded manually in an excel sheet. Same way, the position of interacted vehicles with the crossing pedestrian(s) during the crossing manoeuvre was accurately observed and noted. The utmost care has been taken that the timeline of crossing pedestrian and vehicle position kept the same by replaying the video number of times for each interacting vehicle.

All the vehicles are classified into five categories, as shown in Table 1, along with the physical size of vehicles. There are several models of the same category of car running on roads in India. Therefore, cars have been divided into two separate categories as small or standard
cars and big car. The classification of the car is carried out according to its size and engine power, as illustrated in Table 1. The small car is identified with engine power up to 1400 c, and the big car is categorised with engine power more than 1400 cc with size as per Table 1. The average dimension of the vehicle is taken if more than one type of vehicle is included in the same category (motorised two-wheelers).

Trajectory data of various interactions between the pedestrian(s) and different categories of vehicles were extracted for further analysis. From the extracted data, trajectories for the pedestrian and interacting vehicles are plotted. Fig. 2(c) shows the trajectory of pedestrian and vehicle movement at the section (without time). To plot trajectories, the length of the grid, i.e., 50-m has been taken on X-axis, the width of the grid, i.e., 10.5-m has been taken on Y-axis, and time has been taken on Z-axis as shown in Fig. 2 (d). The figure shows a sample trajectory plot presenting the interaction of crossing pedestrians with different vehicles. The trajectories along the width of the road are pedestrian trajectories, whereas the trajectories along the length of the road are vehicle trajectories. The trajectories are expedient to understand the relative movement of crossing pedestrians and vehicles.

7. Computation of time to collision (TTC)

In this study, TTC has been used as a surrogate safety measure. TTC is the time taken by vehicle-pedestrian to collide if they continue on their present trajectory at the same speed. A higher TTC value indicates lower probability of conflict to occur; on the contrary, a lower TTC value suggests high probability of conflict. Thus lower TTC means unsafe condition that arises due to the risky or aggressive behaviour of vehicle driver or pedestrian. To evaluate TTC, two cases have been considered, Vehicle Pass First (VPF) and Pedestrian Pass First (PPF). In the case of VPF, the vehicle reaches the theoretical conflict point first, crosses the point, and
in due course of time, the pedestrian arrives at the conflict point, as shown in Fig. 3 (a). In
addition to this, the figure also indicates the PPF case in which the pedestrian reaches the
conflict point before the vehicle.

All kinematic parameters have been calculated using the trajectories data, where the
conflicting vehicle is traced using a point marking system. However, in terms of collision
analysis, vehicle geometry should be considered. At a definite time in the interaction process,
the current state of the vehicle and the pedestrian can be defined by their speed vehicle crossing
speed (VS), Pedestrian Speed (PS) and vehicle distance to the conflict point (VDC), and
pedestrian distance to the conflict point (PDC). VDC and PDC are the distance of vehicle and
pedestrian to travel in order to reach the conflict point, respectively. Initially, the central point
of the front bumper of the conflicting vehicle is considered as a reference point during the data
extraction process. However, during conflict analysis, vehicle dimensions need to be
considered. To consider the effect of the physical dimensions of the vehicle in VPF and PPF
cases, VDC and PDC are calculated, as shown in Fig. 3 (b).

[Figure 3 here]

In the VPF case, conflict may occur before the pedestrian reach the conflict point (when
pedestrian reaches point A in Fig. 3(b)) by the distance equal to the half-width of the conflicting
vehicle. In addition to it, conflict may be possible until the entire length of the vehicle does not
cross the conflict point. Hence, in the VPF case, PDC is considered up to point A and VDC is
calculated by including the length of the vehicle, as shown in Fig. 3(b). In the PPF case, conflict
may be possible, even when the pedestrian crosses the conflict point (till pedestrian does not
cross point B in Fig. 3(b)). So, in the PPF case, PDC is measured up to point B.

TTC\textsubscript{PED} and TTC\textsubscript{VEH} are defined as the time required by pedestrian and vehicle
respectively to reach the conflict point if they both continue at the same speed. TTC\textsubscript{VEP} and
TTC\textsubscript{PED} are worked out as under using Equation 1 and 2.
\[ \text{TTC}_{VEP} = \frac{V_{DC}}{V_S} \]  

(1)

\[ \text{TTC}_{PED} = \frac{P_{DC}}{P_S} \]  

(2)

The TTC is stated as,

\[ \text{TTC} = \{\text{TTC}_{VEH}, \text{TTC}_{VEH} \geq \text{TTC}_{PED}\} \]

Or

\[ \text{TTC} = \{\text{TTC}_{PED}, \text{TTC}_{VEH} < \text{TTC}_{PED}\} \]  

(3)

For each interaction, the TTC is worked out using the above approach as given in Equation 3.

8. Analysis of TTC

In the present study, for the analysis purpose, all the vehicles have been grouped into different categories, as shown in Fig. 1. The traffic composition observed at study locations is shown in Fig. 1. At most of the locations, the proportion of 2W is higher. The proportion of HV is minimum at all the locations. The vehicular flow and pedestrian flow observed in the study sections is presented in Fig. 1. As the sample size of BC is less, the samples of BC are merged into SC for analysis purpose after calculation of TTC.

[Table 2 here]

It is observed that the pedestrian interacts with several vehicles during the crossing manoeuvre. In such instances, the pedestrian may have to wait due to an impending conflict with another vehicle, or the pedestrian decides to be cautious and waits for a suitable opportunity. In such cases, it is observed that the pedestrian spends considerable time in the carriageway or on the fringe position. Although he ultimately reaches a conflict point shared by another vehicle, can it be called a conflict? The vehicle might have passed long before the pedestrian(s) arrival. To avoid such instances and to measure the risk with higher reliability,
only those interactions are considered where the difference between the arrival time of the predecessor and successor in a conflict is less than 2.5-s. The reason for selecting the limiting time of 2.5-s is because if a pedestrian travels with an average crossing speed of 1.3-m/s, it will take him 2.5-s to cross a lane of 3.3-m width, after which he might be interacting with other vehicles. TTC for total, all interactions have been worked out fulfilling the above criteria.

The analysis has been carried out based on the vehicle category and gender in both VPF and PPF cases. The descriptive statistics for the same has been illustrated in Table 2.

Table 2 shows the vehicle category wise and gender-wise descriptive statistics of TTC in VPF and PPF. The variation in TTC value is observed in VPF and PPF cases that show the interaction behaviour between pedestrian and vehicle is different in both cases. In the VPF case, the TTC is higher than PPF, which shows pedestrians are at higher risk in the PPF case as compared to the VPF case. From the table, it can be noticed that the TTC values are sensitive to the vehicle category. TTC is least for 2W and highest for HV. TTC for 3W is higher than 2W and lower than a car. So, it can be said that TTC increases with vehicle size. It is detected by field observations that pedestrians mostly avoid crossing the road when HV is present in the traffic stream. Hence, very few samples have been found at all study locations. Minimum TTC value ranges from 0.12-s for the car to 0.41-s for HV in the VPF case, whereas in the PPF case, the minimum TTC value varies from 0.17-s for 2W to 0.52-s for HV. The maximum value of TTC ranges from 17.54-s to 21.97-s. The table also shows the gender-wise descriptive statistics for TTC, and it can be observed that the mean TTC value for the male is least followed by female and group in both VPF and PPF, which indicates male take more risk than female. In group, TTC value is highest in both VPF and PPF cases that interprets as it is safer to cross the road in a group.
Fig. 4 shows the box plot and the cumulative frequency of TTC. Fig. 4(a) shows the variation in TTC in VPF and PPF conditions, whereas Fig. 4(b) and Fig. 4(c) shows such variation based on the vehicle class and pedestrian gender, respectively. The figure shows the TTC values are higher in the VPF condition. From the plot, it is being confirmed that the TTC value is sensitive to vehicle class and increases with an increase in the size of vehicles as pedestrians perceive higher risk with large sized of approaching vehicles. Moreover, it can also discern that in both VPF and PPF, TTC is least for male and highest for the group.

9. Statistical distribution of TTC

The distribution of TTC is fitted with three types of hypothesised distribution, and Kolmogorov Smirnov (K-S) and Anderson Darling (AD) tests are performed to determine the goodness of fit. General Extreme Value (GEV) distribution is found to be best fitted for TTC in both types of interaction and as a whole data set. The values of K-S statistics for VPF, PPF, and combined data are observed as 0.0108, 0.0142, and 0.0015 in the comparison of critical values as 0.0111, 0.0155, and 0.0090, respectively. The K-S statistics values for all three cases are less than the respective critical values, which shows that TTC follows the GEV distribution. Same way, the critical value in the AD test for VPF, PPF, and overall data is 2.5018, which is higher than AD statistics values, 1.904, 1.7422 and 1.3717, respectively. Hence, it is strongly concluded that TTC follows the GEV distribution.

10. Relation between TTC, vehicle speed and pedestrian speed

The observed value of TTC is plotted against VS and PS to analyse the relationship among these parameters. Fig. 5 shows the relation between TTC and VS. The TTC is following negative relation with pedestrian speed. Pedestrian requires less time to reach conflict points if the speed of pedestrian increases. A negative logarithmic relationship is observed between TTC and pedestrian speed. The TTC is also plotted with the corresponding speed of vehicles and
shown in Fig. 5. The TTC decreases with the increase in vehicles speed which is observed to follow the negative logarithmic trend.

11. Development of generalised linear model for TTC

The extracted trajectory data used for calculating TTC along with different parameters used for calculating TTC like vehicle speed, pedestrian speed. The other factors affecting TTC are also identified, such as the gender of pedestrian, group of crossing pedestrians, type of interaction, which includes VPF and PPF, and vehicle category. These factors are taken as independent variables to develop the GLM model. The TTC value VS and PS after every 0.48s, are considered for the model formation.

[Figure 5 here]

Total 21,110 valid data are considered for the analysis. Out of the total dataset, 70 percent of data is used for model formation, and the remaining 30 percent of data used for validation. The GLM model is developed in R software (version 3.5.3). The proposed model for TTC is given in Equation 4.

\[
TTC = 0.802 \times GEN + 0.155 \times GRP + 0.543 \times VEH + 1.140 \times CON - 0.462 \times PS - 0.0607 \times VS
\]

\[
(26.86) \quad (7.78) \quad (22.81) \quad (30.34) \quad (12.56) \quad (11.29)
\]

Where,

- GEN is the gender of pedestrian, which is the discrete variable (1-male, 2-female 3-group),
- GRP is the group size which is the discrete variable (1-single, 2-two, 3-three, 4-four 5-five and more than five),
- VEH is the vehicle class which is the discrete variable (1-2W, 2-3W, 3-Car, 4-HV),
- CON is the conflict type which is the discrete variable (1-VPF, 2-PPF),
- PS is the pedestrian crossing speed in m/s, which is the continuous variable,
VS is the vehicle speed in m/s, which is the continuous variable.

For the developed model, the value of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are obtained as 69,421 and 69,474, respectively. Here, the value of BIC is more than AIC, which indicates that the model is best fitted. In the developed model, the negative sign of VS and PS shows that TTC decreases with increasing VS and PS. Same way, the sign of gender, group, vehicle class, and conflict type also fits logically. In the above equation, the ‘t’ value of all coefficients is more than 1.96, which shows that all coefficients are significant at 95 percent confidence level.

12. Model validation

To assess the accuracy of the proposed model, the TTC predicted by the model is compared with field observed values. The data kept aside for model validation are used for this purpose. The input parameters for the proposed model are observed from the field data. The developed model is used to predict TTC based on the input parameters. The predicted TTC values are compared with the TTC values calculated using trajectory data. In order to compare the two datasets, a t-test has been carried out to at a 5% significant level with 6333 degrees of freedom for statistical validation. The test result showed that the t-statistics value (1.11) is less than the t-critical value (1.98) at a 5% significant level. Hence, there is no statistically significant difference between the predicted and observed TTC.

13. Clustering analysis of TTC

The TTC data sets were classified into different groups to identify the severity of conflict using a k-means clustering technique which is also adopted by many researchers.\textsuperscript{55–60} The clustering analysis is carried out using ‘MATLAB’ software. Classification in the MATLAB tool was conducted for different numbers of k-values, which resulted in two-cluster, three-cluster, four-cluster, five-cluster, and six-cluster. Silhouette analysis is carried out to
identify the optimum number of clusters for a given data set range and variation. The Silhouette value is calculated for TTC data sets and used for comparison of scenarios involving the various number of clusters. Fig. 6 shows the average Silhouette values and Silhouette plot for different clusters. Based on the average Silhouette value, two clusters are optimum for the TTC dataset. Based on this result, the TTC value is classified into two clusters.

[Figure 6 here]

In the present study, TTC has been worked out based on pedestrian-vehicle interaction with high to moderate risk for the pedestrian. During data extraction, only such interactions are considered where the pedestrian-vehicle interaction is observed. The same is verified by silhouette analysis. As the data set is within the range of high to moderate risk two clusters, cluster-1 identifies the pedestrians at high risk, and cluster-2 specifies moderate risk of collision to pedestrians.

Based on the cluster analysis, the threshold value has been worked out, which implies the boundary between high risk to moderate risk of collision. Fig. 6 shows that when TTC is less than 3.60-s, the pedestrian is at high risk of collision.

### 14. Conclusions

The present study is an attempt to evaluate the safety of crossing pedestrians at urban midblock locations under mixed traffic conditions. Time to collision (TTC) is taken as the surrogate safety measure for the conflict analysis. For TTC measurement at any other location, traffic trajectory data is pivotal. Using traditional methods, trajectory data can be acquired from the filed video data following the methodology used in the present study or using other methods. Alternatively, an automated video analysis tool can also be used to extract surrogate safety measure values like TTC more quickly. However, bot both approaches require analysis of
traffic trajectory data for TTC estimation. The method and formula of TTC calculation shall remain the same for each approach.

The data was collected using the videographic technique from the urban midblock section influenced by crossing pedestrians at six-lane arterials. To assess the TTC, the trajectory-based approach is adopted. Two different types of interaction, namely vehicle pass first (VPF) and pedestrian pass first (PPF) has been considered for the analysis purpose. More than 21,000 instances of interactions between the pedestrian and vehicle are used for the analysis. The TTC values are sensitive to vehicle categories and pedestrian gender. The TTC values increase with the increase in the size of vehicles. Moreover, TTC values of 2W are the least among all classes of vehicles, whereas TTC values of HV are the highest. The statistical distribution of TTC shows that TTC was found to follow a Generalised Extreme Value distribution. The TTC follows the decreasing logarithmic trend with pedestrian and vehicle speed. The generalised linear model is proposed to predict TTC under mixed traffic conditions. The k-means clustering analysis is carried out for identifying the classification of the interaction in different categories for risk. The optimum number of clusters were identified using silhouette analysis. It is found that two clusters are optimum based on the Silhouette value for the present data. Moreover, the threshold value for pedestrian-vehicle interactions has been worked and is found as 3.60-s.

**15. Contribution of the Present Study**

The models developed in the present study are based on the parameters that can be easily observed and quantified in the field. This reduces the dependability on traffic trajectory data for safety analysis. The present study may use to assess the safety of pedestrians at midblock crosswalks under mixed traffic conditions to improve planning and to design such a traffic facility to create a safe environment for vulnerable road users. Using the parameters
observed in the field, TTC can be estimated using the model and real-time safety aspects measured. The parameters can be identified using traditional or modern Intelligent transport system based tools, gender of pedestrian, group size, vehicle class, speed of the vehicle, pedestrian speed, and resulting conflict type to measure the TTC of the interaction at any other location. Additionally, the effect of different policies and their corresponding changes on the model parameters can be recorded to observe the changes in TTC values and risk levels. This also creates robust and practical means to test the effects of different policies on the safety aspects of the traffic infrastructure.

Additionally, the present study assesses the surrogate safety aspects of pedestrian and vehicular interaction at designated and undesignated urban midblock crossings by overcoming the limitations of dependency on underreported crash data \(^{65,66}\) by implementing a proactive approach. Numerous studies have utilized TTC for the safety assessment of pedestrian and vehicular interactions. However, such studies are limited for the mixed traffic conditions observed in India. The present study attempts to bridge the gaps in the literature and also provide means and methodology for real-time proactive safety assessment.

### 16. Limitations and Future Scope

In the present study the limited trajectory data is used. The same study can be evaluated by using the semi-automated and automated trajectory extraction tools. Further, in the present study, only TTC as a surrogate safety measure is considered. The different SSMs can be evaluated and compared with the present study results. The models proposed in the present study can be incorporated in the infrastructure and vehicles to aid vehicle to vehicle and vehicle to infrastructure communications to perform real-time safety analysis, driver assistance system and collision avoidance systems. Based on the risk estimated using the models, measures and policies can be tested that support mitigation of risk levels in real-time hence enhancing the
safety levels. This also forms future applications, scope, and research areas for further exploration.

References


10. Ismail, K., Sayed, T., Saunier, N., and Lim, C., Automated analysis of pedestrian-vehicle


40. Madhumita, P. and Ghosh, I., A novel approach of safety assessment at uncontrolled


632
633 61. Chauhan, R., Dhamaniya, A., and Arkatkar, S., Spatiotemporal Variation of Rear-End 
634
       Multiple Vehicular Trajectory Data for Traffic Modeling. J. Transp. Eng., 2005, 131, 
       496–505.
636
637 63. Suzuki, K. and Nakamura, H., Traffic analyzer: the integrated video image processing 
638
639 64. Ismail, K., Sayed, T., Saunier, N., and Lim, C., Automated Analysis Of Pedestrian- 
640
641 65. Dandona, R., Kumar, G. A., Ameer, M. A., Reddy, G. B., and Dandona, L., Under- 
       reporting of road traffic injuries to the police: Results from two data sources in urban 
642
       traffic accidents in traffic police records- a cross sectional study from North India. Int. 
Figures

Fig. 1. Interruption to vehicles due to crossing pedestrians
Fig. 2. Trajectory and its plotting (a) Grid (b) Overlaid Grid (c) Trajectories in X-Y plane and (d) Trajectories in X-Y-Z plane
Fig. 3. Types of interactions (a) VVF and PPF (b) VDC and PDC calculation in VPF and PPF cases
Fig. 4. Box plot and cumulative distribution of TTC for (a) VPF and PPF (b) vehicle categories (c) pedestrian gender
Fig. 5. Variation of TTC with vehicle speed and pedestrian speed
Fig. 6. Silhouette analysis for optimum cluster and the threshold value for TTC
Tables

Table 1 Category of Vehicle with its Size

<table>
<thead>
<tr>
<th>Vehicle Type</th>
<th>Vehicles included</th>
<th>Length (m)</th>
<th>Width (m)</th>
<th>Rectangular Plan Area (m²)</th>
<th>Figure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-wheeler (2W)</td>
<td>Scooter, Motorcycle</td>
<td>1.87</td>
<td>0.64</td>
<td>1.20</td>
<td></td>
</tr>
<tr>
<td>Three-wheeler (3W)</td>
<td>Auto-rickshaw</td>
<td>3.20</td>
<td>1.43</td>
<td>4.48</td>
<td></td>
</tr>
<tr>
<td>Small Car (SC)</td>
<td>Car</td>
<td>3.72</td>
<td>1.44</td>
<td>5.36</td>
<td></td>
</tr>
<tr>
<td>Big Car (BC)</td>
<td>Big utility vehicle</td>
<td>4.58</td>
<td>1.77</td>
<td>8.11</td>
<td></td>
</tr>
<tr>
<td>Heavy Vehicle (HV)</td>
<td>Standard bus</td>
<td>10.10</td>
<td>2.43</td>
<td>24.54</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 Vehicle Category wise and Gender wise Descriptive Statistics of TTC

<table>
<thead>
<tr>
<th>Particulars</th>
<th>Type of Interaction</th>
<th>Mean (s)</th>
<th>Maximum (s)</th>
<th>Minimum (s)</th>
<th>Total</th>
<th>Standard Deviation</th>
<th>Percentile Values (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15th</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50th</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>85th</td>
</tr>
<tr>
<td>2W</td>
<td>VPF</td>
<td>3.41</td>
<td>21.97</td>
<td>0.18</td>
<td>8230</td>
<td>2.45</td>
<td>1.18</td>
</tr>
<tr>
<td></td>
<td>PPF</td>
<td>3.08</td>
<td>20.65</td>
<td>0.17</td>
<td>4072</td>
<td>2.22</td>
<td>1.27</td>
</tr>
<tr>
<td>3W</td>
<td>VPF</td>
<td>3.6</td>
<td>18.36</td>
<td>0.15</td>
<td>4457</td>
<td>2.47</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>PPF</td>
<td>3.51</td>
<td>18.200</td>
<td>0.22</td>
<td>2436</td>
<td>2.22</td>
<td>1.55</td>
</tr>
<tr>
<td>Car</td>
<td>VPF</td>
<td>3.69</td>
<td>20.18</td>
<td>0.12</td>
<td>1634</td>
<td>2.51</td>
<td>1.35</td>
</tr>
<tr>
<td></td>
<td>PPF</td>
<td>3.21</td>
<td>17.54</td>
<td>0.27</td>
<td>873</td>
<td>1.91</td>
<td>1.53</td>
</tr>
<tr>
<td>HV</td>
<td>VPF</td>
<td>4.11</td>
<td>20.03</td>
<td>0.41</td>
<td>563</td>
<td>2.39</td>
<td>1.96</td>
</tr>
<tr>
<td></td>
<td>PPF</td>
<td>3.67</td>
<td>19.30</td>
<td>0.52</td>
<td>288</td>
<td>1.99</td>
<td>1.98</td>
</tr>
<tr>
<td>Male</td>
<td>VPF</td>
<td>3.30</td>
<td>18.83</td>
<td>0.12</td>
<td>5464</td>
<td>2.34</td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>PPF</td>
<td>3.02</td>
<td>19.31</td>
<td>0.17</td>
<td>3211</td>
<td>2.16</td>
<td>1.21</td>
</tr>
<tr>
<td>Female</td>
<td>VPF</td>
<td>3.35</td>
<td>20.03</td>
<td>0.18</td>
<td>4078</td>
<td>2.21</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>PPF</td>
<td>3.20</td>
<td>20.65</td>
<td>0.22</td>
<td>2230</td>
<td>1.94</td>
<td>1.41</td>
</tr>
<tr>
<td>Group</td>
<td>VPF</td>
<td>3.87</td>
<td>21.97</td>
<td>0.18</td>
<td>5342</td>
<td>2.73</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td>PPF</td>
<td>3.63</td>
<td>20.64</td>
<td>0.29</td>
<td>2228</td>
<td>2.41</td>
<td>1.60</td>
</tr>
</tbody>
</table>