Data-Driven Reachability Analysis of Pedestrians Using Behavior Modes

August Söderlund[1,2], Frank J. Jiang[1], Vandana Narri[1,2], Amr Alanwar[3], and Karl H. Johansson[1]

Abstract— In this paper, we present a data-driven approach for safely predicting the future state sets of pedestrians. Previous approaches to predicting the future state sets of pedestrians either do not provide safety guarantees or are overly conservative. Moreover, an additional challenge is the selection or identification of a model that sufficiently captures the motion of pedestrians. To address these issues, this paper introduces the idea of splitting previously collected, historical pedestrian trajectories into different behavior modes for performing data-driven reachability analysis. Through this proposed approach, we are able to use data-driven reachability analysis to capture the future state sets of pedestrians, while being less conservative and still maintaining safety guarantees. Furthermore, this approach is modular and can support different approaches for behavior splitting. To illustrate the efficacy of the approach, we implement our method with a basic behavior-splitting module and evaluate the implementation on an open-source data set of real pedestrian trajectories. In this evaluation, we find that the modal reachable sets are less conservative and more descriptive of the future state sets of the pedestrian.

I. INTRODUCTION

In urban environments, pedestrians are one of the most vulnerable and challenging road users that automated vehicles need to operate around safely [1]. While pedestrians only have the right-of-way in certain scenarios, they often move around freely, as can be seen in Fig. 1, and need to be evaded regardless of how they behave. Moreover, since pedestrian motion is difficult to model accurately, their future states can be hard to predict.

To safely predict the future state sets of other road users, researchers propose using reachability analysis [2], [3]. Normally, the reachability of a system is calculated by propagating a known model of the system over a specific time horizon [2], [4]. However, for traffic scenarios involving pedestrians, there is no consensus on how to appropriately model the human and all of its different behaviors [5], [6].
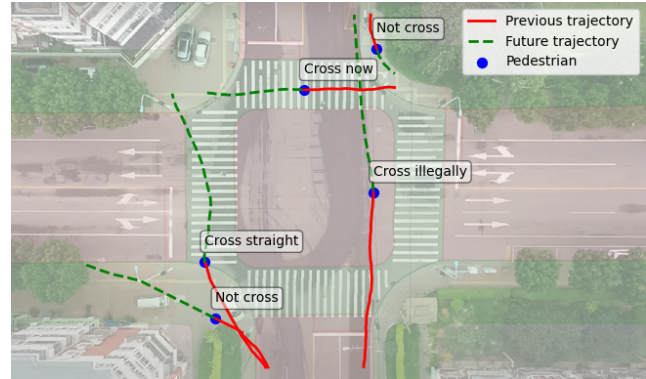
Fig. 1: An intersection in Tianjin, China, showcasing the movement of five pedestrians, indicated by the blue circles, and their respective behavior mode, described in the text box above each pedestrian, deduced from either the previous trajectory (red lines) or the future trajectory (dashed green lines).

Alternatively, using data-driven approaches for the reachability analysis simplifies the problem by removing the need for an explicit model of the system [7]–[11]. However, data-driven reachability analysis can sometimes lead to predicting overly conservative sets [8], [9]. One previous work explores the use of known side information to reduce the conservativeness of data-driven reachability analysis [7]. In this work, we will explore the use of behavior predictions to reduce the conservativeness of data-driven reachability analysis.

Several recent works have shown that pedestrians normally communicate their intention in traffic scenarios [1], [12]. For instance, in many scenarios, pedestrians cross the road after achieving eye contact with the driver. Also, other external factors that affect their behavior include different aspects in the local scene or the speed of a potential approaching vehicles. To this end, several works aim to model and predict the different behaviors expressed by humans [13]–[16]. We aim to leverage behavior predictions to reduce the conservativeness of the data-driven reachable sets by separating the collected pedestrian trajectory data into different modes, representing the different behaviors and intentions of pedestrians in traffic scenarios.

The main contribution of this paper is providing a modular framework for predicting the future set of states of pedestrians by performing data-driven reachability analysis using the historical trajectories with the same mode as the current behavioral mode of the pedestrian. The presented approach enables less conservative state predictions in conjunction with an accurate behavior prediction approach while providing similar safety guarantees to the conventional method in [8], [9]. To the extent of the authors' knowledge, similar

research has been conducted in [17]; however, our work contrasts in depending on a more computationally efficient set representation called zonotopes. Using zonotopes, we can better represent the uncertainties in inputs and noise prevalent in the process, thus rendering guarantees on safety. More specifically, the contribution of this paper is threefold:

1) We present a modular approach for incorporating behavior mode predictions into data-driven reachability analysis.
2) We propose an architecture for integrating modal data-driven reachability analysis into automated driving systems.
3) We implement and evaluate the presented approach on an open-source data set of real pedestrian trajectories. The code to recreate our findings is publicly available[1].

The remainder of the paper is outlined as follows. Section II defines the problem and the necessary prerequisites. Section III introduces the method for the modal data-driven reachability analysis. In Section IV, the setup used for producing the results is introduced. Section V presents the results and discusses the findings. Lastly, in Section VI, the conclusions are summarized along with a short discussion of future work proposals.

II. PRELIMINARIES

In this section, the necessary prerequisites are presented, along with an assumed model of the pedestrian movement system. This section ends with an explicit formulation of the problem that will be solved by using the method presented in Section III.

A. Pedestrian Model

In this paper, we let the state of a pedestrian be $x = [p_{\mathrm{x}} \ p_{\mathrm{y}}]^\top \in \mathbb{R}^2$, where p_{x} and p_{y} are the global x and y positions of the pedestrian, respectively. In this paper, we represent pedestrian dynamics as a linear time-invariant (LTI) discrete model of the form

$$x(k+1) = Ax(k) + Bu(k) + w(k), \tag{1}$$

where $u = [v_{\mathrm{x}} \ v_{\mathrm{y}}]^\top \in \mathbb{R}^2$ is the x and y components of the pedestrian's velocity and $w \in \mathcal{Z}_w$ is the process noise. Note, that the true pedestrian model $[A \ B]$ is considered unknown. We aim to use data-driven reachability analysis to safely approximate this model from historical data.

B. Reachability Analysis and Set Representation

To represent the future state set prediction of a pedestrian, we use the following definition for reachable sets.

*Definition 1 (**Reachable set**):* The reachable set after N time steps $\mathcal{R}_N$ for the system given in (1), considering the set of inputs $\mathcal{U}_k$ and set of initial states $\mathcal{X}_0$, is defined as

$$\mathcal{R}_N = \{x(N) \in \mathbb{R}^{n_x} | x(k+1) = Ax(k) + Bu(k) + w(k),$$
$$x(0) \in \mathcal{X}_0, u(k) \in \mathcal{U}_k, w(k) \in \mathcal{Z}_w :$$
$$k = 0, \dots, N-1\} \tag{2}$$

[1]https://github.com/AugustSoderlund/behavioral-data-driven-reachability

We consider a reachable set with $n_x = 2$.

The reachable set is represented using a zonotope which is defined next.

*Definition 2 (**Zonotope** [18]):* Given a center $c_{\mathcal{Z}} \in \mathbb{R}^{n_x}$ and $\gamma_{\mathcal{Z}} \in \mathbb{N}$ generators vectors in a generator matrix $G_{\mathcal{Z}} = \begin{bmatrix} g_{\mathcal{Z}}^{(1)} \ \dots \ g_{\mathcal{Z}}^{(\gamma_{\mathcal{Z}})} \end{bmatrix}$, a zonotope is defined as

$$\mathcal{Z} = \left\{ x \in \mathbb{R}^{n_x} | x = c_{\mathcal{Z}} + \sum_{i=1}^{\gamma_{\mathcal{Z}}} g_{\mathcal{Z}}^{(i)} \beta_{\mathcal{Z}}^{(i)}, -1 \leq \beta_{\mathcal{Z}}^{(i)} \leq 1 \right\}.$$

We use the shorthand notation $\mathcal{Z} = \langle c_{\mathcal{Z}}, G_{\mathcal{Z}} \rangle$ for a zonotope.

The linear map [19] of a zonotope $\mathcal{Z} = \langle c_{\mathcal{Z}}, G_{\mathcal{Z}} \rangle$, given $L \in \mathbb{R}^{n'_x \times n_x}$ is defined and computed as

$$L\mathcal{Z} = \{Lz | z \in \mathcal{Z}\} = \langle Lc_{\mathcal{Z}}, LG_{\mathcal{Z}} \rangle. \tag{3}$$

Given two zonotopes $\mathcal{Z}_1 = \langle c_{\mathcal{Z}_1}, G_{\mathcal{Z}_1} \rangle$ and $\mathcal{Z}_2 = \langle c_{\mathcal{Z}_2}, G_{\mathcal{Z}_2} \rangle$, the Minkowski sum [19] is defined as $\mathcal{Z}_1 \oplus \mathcal{Z}_2 = \{z_1 + z_2 | z_1 \in \mathcal{Z}_1, z_2 \in \mathcal{Z}_2\}$, and can be computed by

$$\mathcal{Z}_1 \oplus \mathcal{Z}_2 = \langle c_{\mathcal{Z}_1} + c_{\mathcal{Z}_2}, [G_{\mathcal{Z}_1}, G_{\mathcal{Z}_2}] \rangle. \tag{4}$$

For simplicity, $+$ will be used instead of $\oplus$ to denote the Minkowski sum. Furthermore, we use $\mathcal{Z}_1 - \mathcal{Z}_2$ to denote $\mathcal{Z}_1 + -1\mathcal{Z}_2$ not the Minkowski difference.

The Cartesian product between two zonotopes $\mathcal{Z}_1$ and $\mathcal{Z}_2$ is defined and computed as

$$\mathcal{Z}_1 \times \mathcal{Z}_2 = \left\{ \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \middle| \ z_1 \in \mathcal{Z}_1, z_2 \in \mathcal{Z}_2 \right\}$$
$$\left\langle \begin{bmatrix} c_{\mathcal{Z}_1} \\ c_{\mathcal{Z}_2} \end{bmatrix}, \begin{bmatrix} G_{\mathcal{Z}_1} & 0 \\ 0 & G_{\mathcal{Z}_2} \end{bmatrix} \right\rangle \tag{5}$$

The set of models which is consistent with the noisy historical pedestrians' trajectories is represented using a matrix zonotope which is defined below.

*Definition 3 (**Matrix zonotope** [2, p. 52]):* A matrix zonotope is defined using a center matrix $C_{\mathcal{M}} \in \mathbb{R}^{n_x \times p}$ and $\gamma_{\mathcal{M}} \in \mathbb{N}$ generator matrices $\tilde{G}_{\mathcal{M}} = \begin{bmatrix} G_{\mathcal{M}}^{(1)} \ \dots \ G_{\mathcal{M}}^{(\gamma_{\mathcal{M}})} \end{bmatrix} \in \mathbb{R}^{n_x \times (p\gamma_{\mathcal{M}})}$ by

$$\mathcal{M} = \left\{ X \in \mathbb{R}^{n_x \times p} \middle| X = C_{\mathcal{M}} + \sum_{i=1}^{\gamma_{\mathcal{M}}} G_{\mathcal{M}}^{(i)} \beta_{\mathcal{M}}^{(i)}, -1 \leq \beta_{\mathcal{M}}^{(i)} \leq 1 \right\}.$$

Using similar notation as for zonotopes, the shorthand notation is used for a matrix zonotope $\mathcal{M} = \langle C_{\mathcal{M}}, \tilde{G}_{\mathcal{M}} \rangle$.

C. Data-Driven Reachability Analysis

Given K input-state data trajectories of arbitrary lengths T_i for $1 \leq i \leq K$, denoted by $\{u^{(i)}(k)\}_{k=0}^{T_i-1}$ and $\{x^{(i)}(k)\}_{k=0}^{T_i}$, the input-state matrices essential for data-driven reachability analysis are constructed by

$$X = [x^{(1)}(0) \dots x^{(1)}(T_1) \dots x^{(K)}(0) \dots x^{(K)}(T_K)],$$
$$X_+ = [x^{(1)}(1) \dots x^{(1)}(T_1) \dots x^{(K)}(1) \dots x^{(K)}(T_K)],$$
$$X_- = [x^{(1)}(0) \dots x^{(1)}(T_1-1) \dots x^{(K)}(0) \dots x^{(K)}(T_K-1)],$$
$$U_- = [u^{(1)}(0) \dots u^{(1)}(T_1-1) \dots u^{(K)}(0) \dots u^{(K)}(T_K-1)].$$
$$\tag{6}$$

Note that both the first data point and the last data point in each trajectory are dropped in the case of X_+ and X_-, respectively. The total amount of data points in the shifted matrices is $T = \sum_{i=1}^{K} T_i$.

The algorithm used for data-driven reachability analysis is presented in [9, Algorithm 1]. The input-state trajectories are preferably collected offline and stored appropriately for easy extraction and usage. Also, the algorithm yields an over-approximation of the reachable set according to [9, Theorem 1] given trajectories generated by the unknown system with a structure given in (1).

D. Problem Statement

Given the prerequisites previously presented in this section, it is now possible to formulate the problem this paper aims to solve. Our overall goal is to utilize historically collected data on pedestrian trajectories to improve the real-time predictions of pedestrian future state sets, so automated driving systems can navigate with safety guarantees in urban environments. As a motivating example, Fig. 2 shows the setup where a pedestrian crosses a zebra crossing with two approaching autonomous vehicles. Conventional methods significantly over-approximate the predictions that hinder both vehicles from proceeding to avoid intersecting with the over-approximated reachable set of the pedestrian. We aim to predict more descriptive reachable sets that enable more efficient planning around pedestrians in urban situations.

More specifically, given the current behavior mode of a pedestrian, historically collected pedestrian input-state trajectories $\mathcal{D} = (U_-, X_-, X_+)$, and assumed noise captured by zonotope $\mathcal{Z}_w$, construct the modal input zonotope $\hat{\mathcal{U}}_k$ and compute the modal reachable set $\hat{\mathcal{R}}_N$, which represents the real-time prediction of the future state sets of the pedestrian.

III. METHODOLOGY

The proposed method is presented in this section, along with techniques that aim to fully utilize pedestrian behavior in computing the reachable sets.

A. Context: System Architecture

In this subsection, we contextualize the method proposed in this paper by presenting an architecture illustrated in Fig. 3
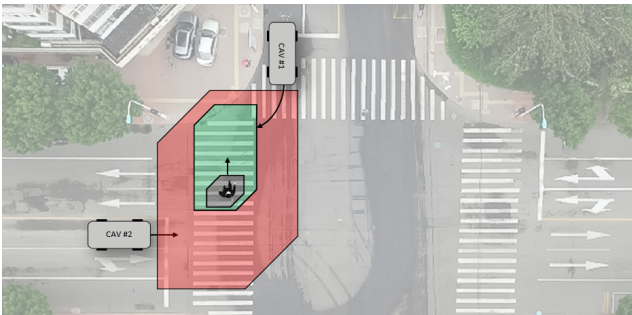

Fig. 2: Problem setup with a crossing pedestrian and two approaching connected and autonomous vehicles. The gray set is the vehicles' estimation of the location of the pedestrian. The red set and the green set are the predicted set of states using conventional methods and using proposed techniques in this paper, respectively.
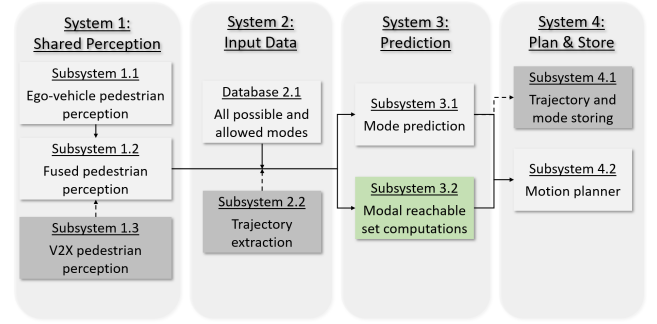

Fig. 3: Proposed architecture of systems in a CAV. The light-gray boxes represent internal subsystems, the dark-gray boxes represent external subsystems and the light-green box is the subsystem we focus on in this paper.

for a connected and automated vehicle (CAV) that leverages modal reachable sets.

The architecture starts off with a system that allows measurements from both internal and external sensors (e.g., information from a camera or LiDAR from another CAV) using sensory data from Subsystems 1.1 and 1.3, as seen in Fig. 3. This information is internally fused in Subsystem 1.2 according to the method presented in [20]. Then, from database 2.1, we obtain the full set of allowed modes. To support the modal reachable set computations we extract historical trajectory data from the cloud using the management system in Subsystem 2.2 proposed in [21]. The extracted trajectories may be either labeled or unlabeled, since Subsystem 3.2 can perform trajectory labeling before predicting the modal reachable sets. In Subsystem 3.1, we perform the mode prediction of a detected pedestrian, as proposed in [14]. This mode prediction is used by the motion planner to evaluate how to use the modal reachable sets and not for the modal reachability analysis. In Subsystem 3.2, we compute the modal reachable set computations based on the method presented in this paper. Then, the motion planner in Subsystem 4.2 uses the modal reachable sets as unsafe regions that the CAV is not allowed to traverse. Finally, this modal trajectory data, together with the mode of the pedestrian, is then sent to the cloud for storage in Subsystem 4.1, using the same cloud-management system previously mentioned. We leave it as future work on how to retrieve mode prediction certainties in Subsystem 3.1 and how the motion planner should combine these certainties with the modal reachable sets in Subsystem 4.2 to still provide safety guarantees.

For the rest of this section, we will detail the implementation of Subsystem 3.2: the modal reachable set computations.

B. Trajectory Modal Split

To perform modal reachable set computations, we need to take the extracted trajectories from Subsystem 2.2 and split the historical data into chunks of data that we will label with corresponding behavior modes. This procedure assumes that we have previously collected historical trajectories $\mathcal{D}$ from multiple pedestrians, ideally from a large data set for different road-crossing scenarios.

We start by dividing these trajectories into smaller chunks of length c_s, since the behavior along a long trajectory can vary. We also need to specify a labeling oracle $\mathcal{O}(\cdot)$. The intention of this oracle is to take each small chunk of trajectory and label it in accordance with which behavior mode the pedestrian expresses, or is predicted to express. For the sake of evaluation in this work, we implement a simple labeling oracle that has conditional statements on the pedestrian's position and initial heading. Due to our modular approach, the labeling oracle can be easily replaced with other, more advanced implementations.

The result of performing the data division is three lists of length N_c: $\mathcal{C}$, α and $\mathcal{L}$, where at index c, $\mathcal{C}^{(c)}$, $\alpha^{(c)}$ and $\mathcal{L}^{(c)} = \mathcal{O}(\mathcal{C}^{(c)})$ denotes a trajectory chunk, the heading of the pedestrian throughout the chunk, and the behavior label given by the labeling oracle, respectively. We denote the ith value in $\mathcal{C}^{(c)}$ and $\alpha^{(c)}$ with $\mathcal{C}_i^{(c)}$ and $\alpha_i^{(c)}$.

C. Modal Reachable Set Computation

Now that the historical trajectories are divided into chunks and labeled, it is possible to perform the modal data-driven reachability analysis, which we summarize in Algorithm 1. Note, Algorithm 1 is designed to output the reachable sets for the possible N_m modes found in Database 2.1 in Fig. 3.

The inputs of the algorithm are the split lists, $\mathcal{C}$, α, and $\mathcal{L}$, information about the pedestrian's current position and heading, and a heading interval limit κ that we will use for finding relevant chunks of trajectory data. For the current position of the pedestrian, we use the estimated state set $\mathcal{X}_p$ of the locations the pedestrian could be in, which is given to us by the fused perception in Subsystem 1.2 in Fig. 3. This allows us to start our reachability analysis from a realistic initial condition, where we assume the location of the pedestrian is not known but has to be estimated from sensors. Moreover, we also input the current heading of the pedestrian α_p. Then, in lines 2-6, we perform a selection to find the chunks of trajectory data that are relevant for each possible behavior mode in Database 2.1.

We start by initializing a list for chunks to keep in line 2. Then, in lines 3-6, we iterate through each chunk to check if the chunk is relevant or not. Our criteria for determining if a chunk is relevant or not are:

(1) The labeled mode of the chunk must be the same as the possible mode of the detected pedestrian.
(2) The starting state x_c of the chunk $\mathcal{C}_0^{(c)}$ must be inside $\mathcal{X}_p$.
(3) The initial heading $\alpha_0^{(c)}$ must be within the specified heading interval limit κ, with respect to α_p.

We illustrate the selection process in Fig. 4, where only one out of four trajectories is kept.

From lines 7-14, we perform the reachability analysis for the current mode. In lines 7-8, we initialize the reachability analysis. In lines 10-12, we use the input trajectories from the kept chunks to compute the input zonotope from the mean and max deviation in all of the relevant inputs. We compute the input zonotope from the relevant historical data; however,

---

Algorithm 1 Modal Data-Driven Reachability Analysis

Input: chunks list $\mathcal{C}$, heading list α, labels list $\mathcal{L}$, process noise zonotope $\mathcal{Z}_w$, estimated set $\mathcal{X}_p$, current heading α_p, heading interval limit κ

Output: Modal reachable sets $\hat{\mathcal{R}}_k^m, \forall k = 1, \dots, N$ and $\forall m = 1, \dots, N_m$

1: **for** $m = 1 : N_m$ **do**
2:    $\mathcal{C}_{\text{keep}} = \emptyset$
3:    **for** $c = 0 : N_c - 1$ **do**
4:       **if** the label of the chunk $\mathcal{L}^{(c)} = m$, starting state x_c of the chunk $\mathcal{C}_0^{(c)}$ is inside $\mathcal{X}_k$, and the initial heading $\alpha_0^{(c)} \in \{\alpha_p + \phi, -\kappa < \phi \leq \kappa\}$ **then**
4:          Append $\mathcal{C}^{(c)}$ to $\mathcal{C}_{\text{keep}}$
5:       **end if**
6:    **end for**
7:    $\hat{\mathcal{R}}_0^m = \mathcal{X}_p$
8:    $\hat{U}_- \leftarrow$ inputs trajectories from $\mathcal{C}_{\text{keep}}$
9:    **for** $k = 0 : N - 1$ **do**
10:      $\mu_k = \text{mean}\left(\hat{U}_-^{(k)}\right)$
11:      $\sigma_k = \max\left(\left|\hat{U}_-^{(k)} - \mu_k\right|\right)$
12:      $\hat{\mathcal{U}}_k = \langle \mu_k, \text{diag}(\sigma_k) \rangle$
13:      $\hat{\mathcal{R}}_{k+1}^m \leftarrow$ [9, Algorithm 1] with $(\mathcal{C}_{\text{keep}}, \hat{\mathcal{R}}_k^m, \mathcal{Z}_w, \hat{\mathcal{U}}_k)$
14:    **end for**
15: **end for**

---

it's also possible to use a predetermined input zonotope for more conservative results, as is done in [9]. Finally, in line 13, the data-driven reachable set for the given mode is computed using [9, Algorithm 1].

Assumption 1: We assume the modal input data $\hat{U}_-$ contains the bounds of the true, unknown modal input zonotope $\mathcal{U}_k$, $\forall k = 0, \dots, N - 1$. Thus, we compute the mean and shifted max value of $\hat{U}_-$, as detailed in line 10-12 of Algorithm 1 where the max and mean functions are performed row-wise, and construct $\hat{\mathcal{U}}_k = \langle \mu_k, \text{diag}(\sigma_k) \rangle$, $\forall k = 1, \dots, N - 1$.

Theorem 1: Given the modal input-state trajectories $\mathcal{D} = (U_-, X_-, X_+)$ generated by system (1) with a full row rank of $\begin{bmatrix} X_-^\top & U_-^\top \end{bmatrix}^\top$, the process noise zonotope $\mathcal{Z}_w$, and assum-
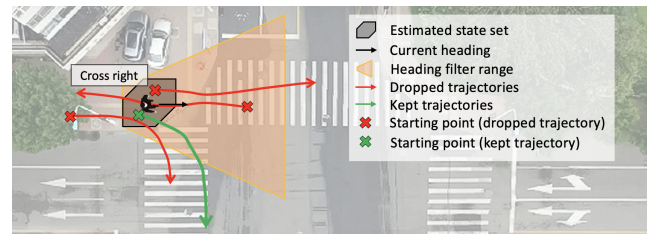
Fig. 4: Illustration of selection technique for a particular scenario. The green trajectory is the only trajectory fulfilling the selection criteria where its starting point is contained in the estimated state set of the pedestrian location, its initial heading is within the given heading range, and it has the same mode as the pedestrian. The red trajectories do not fulfill either one criterion or a combination of the criteria, hence not being used in the reachable set predictions.

ing that we have a correct mode classifier, a correct labeling oracle, and $\hat{\mathcal{U}}_k, \forall k = 1, \ldots, N-1$ fulfills Assumption 1, the reachable set computed from Algorithm 1 over-approximates the exact reachable set, hence $\hat{\mathcal{R}}_k \supseteq \mathcal{R}_k$.

Proof: The exact reachable set, computed using an explicit model of the system, can be established by

$$\mathcal{R}_{k+1} = [A_\mathrm{m}\ B_\mathrm{m}]\,(\mathcal{R}_k \times \mathcal{U}_k) + \mathcal{Z}_w, \qquad (7)$$

where $[A_\mathrm{m}\ B_\mathrm{m}]$ is the true model of the system for a behavior mode. Assuming that the mode of the pedestrian is correctly predicted and that the labeling oracle correctly labels the input-state trajectories, then we know the intention of the pedestrian exactly and the exact actions of pedestrians in previous equivalent conditions. Hence, the behavioral actions observed in the input data are precisely represented in $\hat{\mathcal{U}}_k$, describing the only permissible actions. Also, since $[A_\mathrm{m}\ B_\mathrm{m}] \in \mathcal{M}_\Sigma$ according to [9, Lemma 1] and since both sets $\mathcal{R}_k$ and $\hat{\mathcal{R}}_k$ are computed starting from the same estimated state set $\mathcal{X}_p$, it holds that $\mathcal{R}_{k+1} \subseteq \hat{\mathcal{R}}_{k+1}$. ∎

Remark 1: Bounding $\hat{\mathcal{U}}_k$ to only permit actions prevalent in the data is our method to represent previous behavior. Realistically, a human in traffic scenarios is normally permitted to take actions that are constrained to legal actions (e.g., only walking on sidewalks and crosswalks). However, our implementation constricts the inputs of a pedestrian by prohibiting actions that result in the pedestrian entering regions described by another mode. For example, assuming that the mode classifier accurately predicts that the pedestrian will cross to the right and that historical trajectories are all labeled correctly, we prohibit inputs that result in the reachable set entering the crosswalk straight ahead. As long as these assumptions are satisfied, we can provide safety guarantees. While these assumptions seems restrictive, we believe this enables a principled and less conservative approach to reasoning about the safety of motion around pedestrians.

IV. EXPERIMENTAL SETUP

In this section, the setup for producing the results is presented. This includes specifications of parameters, scenario descriptions, and a summary of the data set.

For the evaluating purposes of this paper, the SIND data set [22] is used. This data set contains trajectory data for different types of road users (e.g., vehicles, bicycles, pedestrians) from a large signalized intersection in Tianjin, China. We focus on the pedestrians from this data set that consists of their location, velocity, and acceleration for all

TABLE I: General parameters

Description	Symbol	Value	Unit
Chunk size	c_s	90	[—]
Heading interval limit	κ	$\pi/4$	[rad]
Process noise (for $\mathcal{Z}_w$)	ω	0.005	[—]
Time horizon	N	9	[s]
Estimated set generator matrix	$G_{\mathcal{X}_{k'}}$	$\begin{bmatrix} 0.5 & 0 & 0.25 \\ 0 & 0.5 & 0.15 \end{bmatrix}$	[m]

time steps. The data set is also split into training data and testing data. This is done by reserving about 100 seconds for testing and using the rest of the data set for training.

The considered state-space is $x = (p_\mathrm{x}, p_\mathrm{y}) \in \mathbb{R}^2$, bounded by the spatial limit of the map for the data set, and the input-space is $u = (v_\mathrm{x}, v_\mathrm{y}) \in \mathbb{R}^2$. The trajectory data are filtered for velocities below $\|u\|_2 < 0.5$ m/s, which remove nearly stationary detections beside crosswalks. The modes used in this paper are categorized as crossing events and not crossing events. Specifically, we use *cross left*, *cross right*, *cross straight*, *cross illegal*, *crossing now*, *not crossing* and *unknown*, and other parameters are presented in Table I.

The labeling oracle $\mathcal{O}(\cdot)$ forms polygons from given relations in the data set (e.g., crosswalk, road, etc), and a sidewalk is approximated. In this paper, we implement a simple version of a labeling oracle where it uses initial headings of the chunks of trajectories and intersections with the specific polygons and labels the chunk appropriately, according to a set of conditional statements. The crossing modes left, right, straight, and not crossing are all predictions of future events and are labeled when future parts of the chunk intersect legal regions for pedestrians, such as sidewalks and crosswalks. The mode *cross now* is a mode where the pedestrian is currently located inside the crosswalk, and the mode *cross illegal* is both a prediction of a future event and a currently expressed mode and is labeled when the trajectory intersects illegal regions, such as the road. Furthermore, connecting chunks within the same mode along the same full trajectory is also done to remove bias towards areas with a large prevalence of chunks. Also, after the labels have been assigned, some trajectories are dropped to have an even distribution between the different modes.

For comparison, a baseline method is adopted. This baseline also uses the same parameters given in Table I, and it uses Algorithm 1 but without forcing the modes of the trajectory chunks and the mode of the pedestrian to be equal, nor do we use a heading filter for the baseline. This enables all chunks that begin inside the initial set to be used in the baseline reachable set predictions.

Specifically, three scenarios will be investigated where the most probable modes can be logically concluded.

(i) The initial set is located at the upper-left corner of the sidewalk with an initial heading in the positive x-direction, that is, to the right.

(ii) The initial set is located at the lower-left corner of the sidewalk with an initial heading in the positive x-direction.

(iii) The initial set is located in the middle of the lower crosswalk with an initial heading parallel to the crosswalk, which is in the negative x-direction.

These scenarios will be generated 20 times, and then we compute the mean and standard deviation of the modal zonotope volumes and present it in Table II, as well as visualizing the modal zonotopes in Fig. 5. Finally, we run a simulation of the test data where the modal sets and the baseline sets are computed. From this simulation, the

(a1) Cross right   (b1) Cross left   (c1) Crossing now

(a2) Cross straight   (b2) Cross straight   (c2) Cross illegal
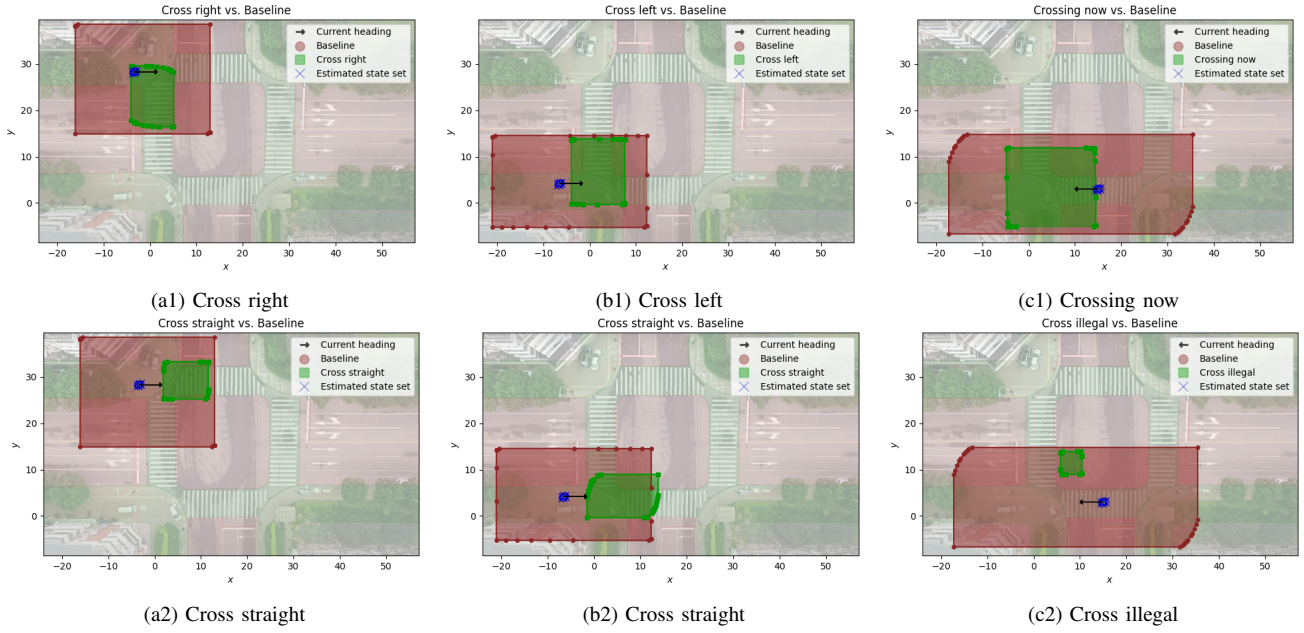
Fig. 5: Visualization of the two most descriptive modal reachable sets for Scenario (a), Scenario (b), and Scenario (c).

TABLE II: Volume of modal zonotopes for the scenarios (empty fields correspond to infeasible modes for the scenario setup)

	Cross left	Cross right	Cross straight	Cross illegal	Cross now	No cross	Unknown	Baseline
Scenario (a)	—	140.6 ± 40.72	63.52 ± 37.04	42.21 ± 0	—	93.62 ± 92.82	18.51 ± 2.019	652 ± 247.9
Scenario (b)	178.8 ± 62.65	—	100.3 ± 45.91	102.2 ± 77.43	—	469.2 ± 142	343 ± 39.93	867.7 ± 83.85
Scenario (c)	—	80.88 ± 60.02	217.6 ± 61.06	750 ± 2242	414.8 ± 202.5	151.8 ± 36.54	185.2 ± 19.45	1345 ± 183.8

state inclusion accuracy is determined. We define the state inclusion accuracy as the portion of times that the true future state of the pedestrian is contained in the predicted reachable set.

V. EVALUATION

In this section, the appropriate results are presented and discussed. In Fig. 5, we present the modal zonotopes compared to the baseline zonotope for the two most justifiable modes for each scenario. In both analyzed modes in Scenario (a) (shown in Fig. 5a1- 5a2), the modal zonotope only intersects parts of the road that it reasonably should (e.g., the modal zonotope for crossing right only intersects the crosswalk that is to the right of the pedestrian). This shows that the state predictions are more descriptive of reality because a pedestrian would reasonably not walk into

Fig. 6: State inclusion accuracy depending on the time horizon.

disallowed areas, such as the middle of the intersection, if they intend to cross the road. In certain driving scenarios for an autonomous vehicle whose starting point is at the top of the intersection, the vehicle would not need to slow down for the pedestrian in Fig. 5a1 since their modal zonotope does not intersect any path originating from this position, allowing the vehicle to be more efficient regarding brake wear and fuel consumption. This is, on the contrary, untrue when considering the baseline zonotope.

In Scenario (b), similar advantages for autonomous vehicles with a planned path that ends below the bottom part of the intersection hold for the pedestrian crossing left, visualized in Fig. 5b1. However, in Fig. 5b2, the modal zonotope intersects the road in which a pedestrian is not allowed to be located. Upon further inspection of the trajectories used for computing the reachable set, we found that many trajectories originate far down in the initial set, biasing the average velocity across the time steps more homogeneously to the right. This causes the reachable set to propagate into the road. Solving this requires more trajectories in different places across the intersection. Similar reasoning explains the modal zonotope in Fig 5c1. Overall, the modal zonotopes are more descriptive than the baseline because they follow the reasonable direction of the mode.

Based on these findings, we can justifiably establish that the more data available for the specific detection of a pedestrian, the more descriptive and logical the reachable set becomes. Furthermore, given sufficiently many data points,
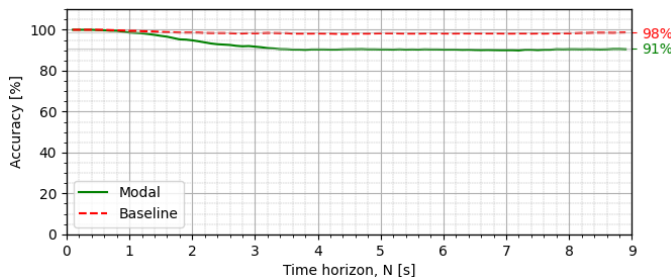
the final reachable set becomes less over-approximate of the predictions, although more importantly, the final reachable set becomes more descriptive of the real outcome.

We show the mean and standard deviation of the volumes for the zonotopes computed for the scenarios in Table II. The fields with a solid line indicate modes for the specific scenarios that are illegitimate modes. It means that it is nearly impossible to be within these modes given the surrounding and initial values; thus, consequentially, there are no trajectories within the same mode as the pedestrian. We find that all modes visualized in Fig 5 are smaller in volume than the baseline, hence being less conservative. The outlier in Table II is the mode *cross illegal* in Scenario (b) with an exceptionally large deviation. Upon further investigation, we found that a probable reason for this large deviation is that there are either too few or too different trajectories used in computing the reachable set. However, for all other modes, the modal volume is smaller than the baseline volume, indicating less conservative predictions.

Fig. 6 shows the accuracy of the modal and baseline set predictions for varying time horizons. Increasing the time horizon for the predictions lowers the accuracy slightly. Although, the modal accuracy stagnates at around 91% for $N > 3$ seconds, whereas the baseline accuracy stagnates at around 98% for $N > 6$ seconds. This shows that it is possible to predict states further in the future using the proposed method with satisfactory accuracy, which would enable more efficient planning for the ego-vehicle. Furthermore, these results use a simplified labeling oracle. Improving the validity of the labeling oracle, in addition to utilizing a mode classifier and combining the modal reachable sets based on the certainties for each mode, would increase the accuracy of the modal reachable sets significantly.

VI. CONCLUSIONS

The proposed method of using mode-splitter trajectories for data-driven reachability analysis leads to less conservativeness whilst still providing sufficient safety accuracy. Also, the modular implementation allows for easy modifications to the labeling oracle, improving the results further. In future work, a mode classifier will have to be designed and trained and also reasoned on how to combine the mode confidences with the modal reachable sets. Incorporating temporal logic side information similar to [7] to further reduce conservativeness and include more modes that allow for signalized crosswalks. Furthermore, formulate some metric that better describes the conservativeness and efficiency of the zonotopes relating to the ego-vehicle's planning. Finally, since many intersections are similar, a labeling oracle that generalizes all intersections should be researched.

REFERENCES

[1] F. Schneemann and I. Gohl, "Analyzing driver-pedestrian interaction at crosswalks: A contribution to autonomous driving in urban environments," in *2016 IEEE Intelligent Vehicles Symposium (IV)*, pp. 38–43, June 2016.

[2] M. Althoff, *Reachability Analysis and its Application to the Safety Assessment of Autonomous Cars*. PhD thesis, Technische Universität München, July 2010.

[3] M. Althoff, G. Frehse, and A. Girard, "Set Propagation Techniques for Reachability Analysis," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, no. 1, pp. 369–395, 2021.

[4] S. Bansal, M. Chen, S. Herbert, and C. J. Tomlin, "Hamilton-Jacobi reachability: A brief overview and recent advances," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pp. 2242–2253, Dec. 2017.

[5] Q. Fan, H. Zhang, Y. Sun, Y. Zhu, X. Zhuang, J. Jia, and P. Zhang, "An Optimal Enhanced Kalman Filter for a ZUPT-Aided Pedestrian Positioning Coupling Model," *Sensors*, vol. 18, p. 1404, May 2018.

[6] M. Meuter, U. Iurgel, S.-B. Park, and A. Kummert, "The unscented Kalman filter for pedestrian tracking from a moving host," in *2008 IEEE Intelligent Vehicles Symposium*, pp. 37–42, June 2008.

[7] A. Alanwar, F. J. Jiang, M. Sharifi, D. V. Dimarogonas, and K. H. Johansson, "Enhancing Data-Driven Reachability Analysis using Temporal Logic Side Information," in *2022 International Conference on Robotics and Automation (ICRA)*, pp. 6793–6799, May 2022.

[8] A. Alanwar, A. Koch, F. Allgöwer, and K. H. Johansson, "Data-Driven Reachability Analysis Using Matrix Zonotopes," in *Proceedings of the 3rd Conference on Learning for Dynamics and Control*, pp. 163–175, PMLR, May 2021.

[9] A. Alanwar, A. Koch, F. Allgöwer, and K. H. Johansson, "Data-Driven Reachability Analysis from Noisy Data," *IEEE Transactions on Automatic Control*, vol. 68, no. 5, pp. 3054–3069, 2023.

[10] A. Devonport, F. Yang, L. El Ghaoui, and M. Arcak, "Data-Driven Reachability Analysis with Christoffel Functions," in *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 5067–5072, Dec. 2021.

[11] A. Chakrabarty, A. Raghunathan, S. Di Cairano, and C. Danielson, "Data-Driven Estimation of Backward Reachable and Invariant Sets for Unmodeled Systems via Active Learning," in *2018 IEEE Conference on Decision and Control (CDC)*, pp. 372–377, Dec. 2018.

[12] A. Rasouli, I. Kotseruba, and J. K. Tsotsos, "Understanding Pedestrian Behavior in Complex Traffic Scenes," *IEEE Transactions on Intelligent Vehicles*, vol. 3, pp. 61–70, Mar. 2018.

[13] A. K.-F. Lui, Y.-H. Chan, and M.-F. Leung, "Modelling of Destinations for Data-driven Pedestrian Trajectory Prediction in Public Buildings," in *2021 IEEE International Conference on Big Data (Big Data)*, pp. 1709–1717, Dec. 2021.

[14] A. Rasouli, T. Yau, M. Rohani, and J. Luo, "Multi-Modal Hybrid Architecture for Pedestrian Action Prediction," in *2022 IEEE Intelligent Vehicles Symposium (IV)*, pp. 91–97, June 2022.

[15] R. Peddi, C. D. Franco, S. Gao, and N. Bezzo, "A Data-driven Framework for Proactive Intention-Aware Motion Planning of a Robot in a Human Environment," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5738–5744, Oct. 2020.

[16] A. Rasouli, I. Kotseruba, and J. K. Tsotsos, "Pedestrian Action Anticipation using Contextual Feature Fusion in Stacked RNNs," May 2020.

[17] M. Hartmann, A. Ferrara, and D. Watzenig, "Data-based Reachability Analysis for Movement Prediction of Pedestrians and Motion Planning," in *2018 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, pp. 1–7, Sept. 2018.

[18] W. Kühn, "Rigorously Computed Orbits of Dynamical systems Without the Wrapping Effect," *Computing*, vol. 61, pp. 47–67, Mar. 1998.

[19] A. Alanwar, J. J. Rath, H. Said, K. H. Johansson, and M. Althoff, "Distributed Set-Based Observers Using Diffusion Strategies," *Journal of the Franklin Institute*, vol. 360, no. 10, pp. 6976–6993, 2023.

[20] V. Narri, A. Alanwar, J. Mårtensson, C. Norén, L. Dal Col, and K. H. Johansson, "Set-Membership Estimation in Shared Situational Awareness for Automated Vehicles in Occluded Scenarios," in *2021 IEEE Intelligent Vehicles Symposium (IV)*, pp. 385–392, July 2021.

[21] R. Li, S. Ruan, J. Bao, and Y. Zheng, "A Cloud-Based Trajectory Data Management System," in *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '17, (New York, NY, USA), pp. 1–4, Association for Computing Machinery, Nov. 2017.

[22] Y. Xu, W. Shao, J. Li, K. Yang, W. Wang, H. Huang, C. Lv, and H. Wang, "SIND: A Drone Dataset at Signalized Intersection in China," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2471–2478, 2022.