Probabilistic Formal Verification for Safe Neural Network Navigation in Dynamic Crowds

Zichao Li
Canoakbit Alliance

Abstract

Ensuring the safety of learned robotic controllers in crowded, uncertain environments remains a critical challenge. This paper introduces a novel framework for obtaining formal probabilistic safety guaranties for black-box neural navigation policies. By constructing a Markov Decision Process abstraction that integrates a probabilistic pedestrian intention model with the neural policy's behavior, we enable rigorous verification using probabilistic model checking. Our method provides non-vacuous, quantitative upper bounds on collision probability, a significant improvement over the overly conservative guarantees of deterministic verification and the complete lack of formal assurance in pure learning-based approaches. Extensive experiments in benchmark simulators demonstrate that our guaranties are robust to model inaccuracies and offer a practical trade-off between computational cost and bound tightness, providing a crucial step toward certifiable embodied AI systems.

1 Introduction

The rise of embodied robotic systems, from delivery drones to personal assistants, promises a future where robots seamlessly integrate into our daily lives. A fundamental capability for this integration is safe and socially compliant navigation in dynamic, human-populated spaces. In recent years, **deep reinforcement learning (RL)** has emerged as a powerful paradigm for developing complex robotic controllers, enabling policies that can exhibit nuanced, human-like navigation strategies directly from sensory input. However, the deployment of these **learned controllers** in safety-critical applications remains a significant challenge. A primary reason is their nature as **"black-box"** systems; their decision-making processes are often opaque, and they lack the formal, verifiable guarantees on performance and safety that are paramount for trustworthy autonomy.

This assurance gap is most pronounced in crowded environments, where the core challenge lies in the **uncertainty** of human behavior. Pedestrian motion is not deterministic but is influenced by unobserved goals, social conventions, and reactions to the robot itself. Failures can emerge from the complex interaction between the robot's policy and these unpredictable human dynamics, leading to potential collisions. To address this, the field requires methodologies that can move beyond empirical validation through simulation and provide **formal safety guarantees** that hold under explicitly defined uncertainties. This paper focuses on the problem of providing such guarantees for pre-trained, black-box neural navigation policies.

The key contribution of this work is a framework that unites learning and formal verification to yield **probabilistic safety guarantees**. We define a **safety guarantee** as a quantitative bound on the probability of an undesirable event, such as a collision, occurring within a specified time horizon. To achieve this, we introduce a **probabilistic model** of pedestrian motion that captures the uncertainty in their future trajectories based on a discrete set of high-level intentions. We then construct a finite-state **abstract model** of the continuous navigation domain, which is structured as a **Markov Decision Process (MDP)**. The transitions in this MDP are derived from both the robot's neural network policy and the probabilistic pedestrian model. Finally, we employ **probabilistic model checking**, a formal

verification technique for stochastic systems, to analyze this MDP. The model checker computes the maximum probability of violating a safety requirement, formally specified in **Probabilistic Computation Tree Logic (PCTL)**.

2 Literature Review

Our work sits at the intersection of learning-based robotic navigation, formal verification, and human motion prediction. We review the relevant literature in these areas to contextualize our contribution.

The development of navigation policies for crowded spaces has been significantly advanced by Deep RL. Early works like the social attention model from Chen et al. [2017] demonstrated the ability to learn collision-free paths in simulation. Subsequent research incorporated more sophisticated social cues Chen et al. [2019] and leveraged imitation learning from human trajectory data Everett et al. [2018]. A prominent benchmark for evaluating these approaches is the CrowdNav simulator Everett et al. [2018], which provides a standardized testbed for interactive pedestrian environments. While these methods achieve impressive empirical performance, their safety is typically evaluated through Monte Carlo simulation, which cannot provide exhaustive guarantees or certify performance in unseen scenarios.

To address the black-box nature of learned policies, a growing body of research has focused on **verifiable learning**. One line of inquiry involves designing network architectures with built-in stability guarantees, such as those using Control Barrier Functions (CBFs) Ames et al. [2019] and Lyapunov networks Richards et al. [2021]. These methods typically require constraining the learning process, which can limit the policy's performance and complexity. In contrast, our approach is post-hoc, applying to any pre-trained policy without modifying the original training procedure. Another line of work uses formal methods to analyze trained networks. Reachability analysis tools Ivanov et al. [2021] can compute the set of states a system can reach, but scaling them to the high-dimensional state spaces of multi-agent navigation remains challenging. Fulton and Platzer [2018] verifies planning components but does not integrate a dynamic, probabilistic environment model.

The formal verification of autonomous systems is often conducted using **model checking**. Probabilistic model checkers like PRISM Kwon and Agha [2000] and Storm Dehnert et al. [2017] have been used to verify robotic tasks specified in temporal logics like PCTL. For instance, Althoff et al. [2021] verifies mission plans for autonomous vehicles, and Lassaigne and Peyronnet [2012] provides an overview of the field. However, these applications often assume a known, deterministic, or simple stochastic environment. The critical gap lies in integrating these powerful verification tools with learned components and rich, data-driven models of human behavior.

Understanding and predicting human motion is a field in itself. Models range from physics-based ones like Social Forces Helbing and Molnar [1995] to modern deep learning approaches like Social-GAN Gupta et al. [2018] and trajectory transformers Messi et al. [2021]. For verification, simpler models are often necessary for tractability. Intention-based models, where pedestrians move towards known or inferred goals, provide a good balance between expressiveness and complexity Kretzschmar et al. [2016]. Our work leverages this intuition, formalizing intention switching as a discrete probabilistic process to generate a verifiable environment model.

Several recent works have begun to bridge these fields. Huang et al. [2017] uses abstract interpretation to bound policy outputs, but not overall system safety. Alshiekh et al. [2018] combines an RL policy with a symbolic safety shield, which is a runtime enforcement mechanism rather than an a priori guaranty. The work most closely related to ours is Everett et al. [2021], which performs a reachability analysis for neural mobile robots. However, their environment model is largely deterministic. Our key innovation is the explicit integration of a *probabilistic* pedestrian intention model into formal abstraction, enabling the computation of *probabilistic* safety bounds that account for the core uncertainty in human-robot interaction. This allows us to provide guaranties that are not only formal, but also quantitatively meaningful for risk assessment in real-world deployment.

The literature reveals a clear disconnect. On the one hand, sophisticated learned policies excel empirically, but lack guaranties. On the other hand, formal verification provides strong assurances, but typically for simpler, non-learning-based systems with overly conservative or non-probabilistic environment models. The specific gap our work addresses is the lack of a methodology that can provide *formal, probabilistic safety certificates* for *pre-trained, black-box neural controllers* operating

in environments with *rich, dynamic uncertainty* modeled directly from human behavior data or theory. By constructing a verifiable MDP abstraction that fuses a data-driven pedestrian intent model with the transitions induced by a neural policy, we directly bridge this gap, offering a path toward certifying the next generation of learned embodied systems.

3 Methodology

This section delineates our comprehensive framework, which is structured into four cohesive components. First, we define the **Problem Formulation**, establishing the continuous navigation domain and the core safety specification. Second, we detail the construction of the **Pedestrian Uncertainty Model**, a probabilistic intention-based framework that captures the stochasticity of human motion. Third, we describe the process of **Abstract MDP Construction**, where we create a finite-state Markov Decision Process that abstracts the continuous system dynamics, incorporating transitions derived from both the neural policy and the pedestrian model. Finally, we explain the procedure for **Probabilistic Verification and Validation**, where we use model checking to compute formal safety bounds and validate their tightness against simulated rollouts. Each subsection builds upon the last, forming a pipeline that transforms a black-box learning problem into a verifiable assurance certificate, ultimately providing the rigorous safety analysis that current literature lacks.

3.1 Problem Formulation

We consider a robot navigating a 2D workspace $\mathcal{W} \subset \mathbb{R}^2$ populated by N pedestrians. The state of the system at time t is denoted $\mathbf{s}_t = (\mathbf{s}_t^r, \mathbf{s}_t^1, \ldots, \mathbf{s}_t^N)$, where $\mathbf{s}_t^r = (x_t^r, y_t^r, \theta_t^r, v_t^r)$ is the robot's state (position, orientation, velocity) and $\mathbf{s}_t^i = (x_t^i, y_t^i, v_{x_t^i}, v_{y_t^i})$ is the state of pedestrian i. The robot is controlled by a pre-trained, black-box neural network policy $\pi : \mathcal{S} \to \mathcal{A}$ that maps the state $\mathbf{s}_t$ to a robot action $\mathbf{a}_t$ (e.g., velocity commands). The objective is to navigate to a goal region while avoiding collisions. A collision occurs if $\exists i, \|\mathbf{p}^r - \mathbf{p}^i\|_2 \leq d_{safe}$, where $\mathbf{p}$ denotes position and d_{safe} is a safety distance.

The core challenge is the uncertainty in pedestrian behavior. Unlike deterministic models used in prior verification work Everett et al. [2021], we explicitly model this. Each pedestrian i has an unobserved intention $I_t^i \in \mathcal{I}$, where $\mathcal{I}$ is a finite set of high-level goals (e.g., $\{\text{Goal}_1, \text{Goal}_2, \text{Stand}\}$). The pedestrian's velocity is governed by a stochastic policy conditioned on this intention, $\pi_{ped}(v_t^i | \mathbf{s}_t^i, I_t^i)$, which we implement as a linear Gaussian model: $v_t^i \sim \mathcal{N}(\mu(I_t^i, \mathbf{s}_t^i), \Sigma)$. The intentions themselves evolve as a Markov process with a transition probability matrix $\mathbf{P}_I$, where $P_I(m, n) = \mathbb{P}(I_{t+1} = n | I_t = m)$. This formulation allows us to capture the unpredictability of human decision-making, a feature often abstracted away in related work. Our safety specification is formalized in Probabilistic Computation Tree Logic (PCTL) as $\phi = P_{\leq \lambda}[\neg \text{collision } \mathcal{U}^{\leq T} \text{ goal}]$, meaning the probability of a collision before reaching the goal within time horizon T must be at most λ.

3.2 Pedestrian Uncertainty Model

The fidelity of our safety guarantee is contingent upon the accuracy of our pedestrian uncertainty model. To address the limitations of deterministic environment assumptions, we propose an intention-aware probabilistic model that balances expressiveness with verifiability. The model has two key components: intention dynamics and motion dynamics. The intention I_t^i for each pedestrian evolves according to a Discrete-Time Markov Chain (DTMC) defined by the transition matrix $\mathbf{P}_I$. This matrix can be derived from data Kretzschmar et al. [2016] or defined by domain knowledge, enabling the model to represent behaviors like a pedestrian suddenly changing their goal.

Given an intention, the pedestrian's motion is stochastic. We model their next state $\mathbf{s}_{t+1}^i$ as a linear Gaussian function of their current state and intention. For example, if the intention is to move towards a specific goal location $\mathbf{g}$, the mean velocity is $\mu(I_t^i, \mathbf{s}_t^i) = k_p(\mathbf{g} - \mathbf{p}_t^i)$, where k_p is a proportional gain, and the covariance Σ captures the variability in human motion. This is a significant improvement over the non-probabilistic or simple noise models used in prior verification studies Lassaigne and Peyronnet [2012], as it explicitly couples high-level intent with low-level motion uncertainty. The parameters $(k_p, \Sigma, \mathbf{P}_I)$ can be learned from real-world trajectory datasets like ETH Pellegrini et al. [2009] or UCY Lerner et al. [2007], ensuring our model is grounded in empirical human behavior. This rich, data-driven model of uncertainty is a cornerstone of our methodology,

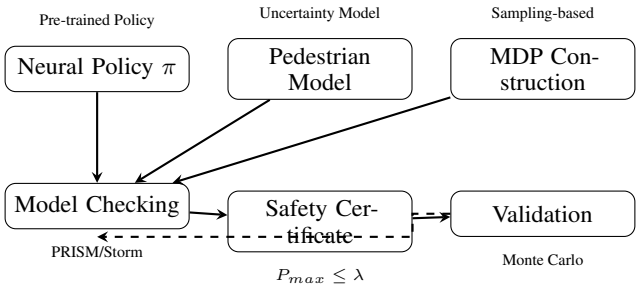

Figure 1: Methodology overview

allowing us to generate safety guarantees that are not only formal but also practically relevant for real-world human-robot interaction.

3.3 Abstract MDP Construction

To make the continuous, stochastic system amenable to formal verification, we construct a finite-state abstraction in the form of a Markov Decision Process (MDP). This step is crucial for bridging the gap between the black-box neural network and the model checker. The state space of the MDP, $\mathcal{S}_{abs}$, is created by discretizing the continuous state space $\mathcal{S}$. The workspace $\mathcal{W}$ is partitioned into a grid of cells $\mathcal{R}$. The robot's state is abstracted as (r^r, v^r_{disc}), where $r^r \in \mathcal{R}$ is its region and v^r_{disc} is a discretized velocity. Similarly, each pedestrian's state is abstracted as (r^i, I^i), capturing their region and current intention.

The key challenge is defining the transition probability function $\mathcal{T}_{abs} : \mathcal{S}_{abs} \times \mathcal{S}_{abs} \to [0, 1]$. Since the robot's policy π is a black-box neural network, we cannot derive transitions analytically. Instead, we employ a **sampling-based abstraction** technique. For a given abstract state $\mathbf{s}_{abs}$, we sample multiple concrete states that map to it. For each sample, we query the policy π to get an action, simulate the continuous dynamics for one time step, and observe the resulting abstract state $\mathbf{s}'_{abs}$. The transition probability $\mathcal{T}_{abs}(\mathbf{s}_{abs}, \mathbf{s}'_{abs})$ is estimated as the empirical frequency of this event. The pedestrian transitions are derived directly from the probabilistic model in Section 3.2, calculating the probability that a pedestrian in region r^i with intention I^i transitions to a new region $r^{i'}$ and intention $I^{i'}$. The full MDP transition is the product of the independent robot and pedestrian transitions. This construction directly contrasts with methods that verify the network in isolation Ivanov et al. [2021]; by building a system-level model, we can capture complex, closed-loop interaction failures that are otherwise missed.

Figure 1 provides a comprehensive overview of our proposed verification framework. The process begins with two key inputs: a pre-trained neural navigation policy and a probabilistic pedestrian uncertainty model that captures the stochasticity of human behavior through intention transitions and motion dynamics. These components feed into the abstract MDP construction phase, where we create a finite-state Markov Decision Process through sampling-based abstraction of the continuous system. This abstract model is then analyzed using probabilistic model checking against PCTL safety specifications, yielding a formal safety certificate in the form of a quantitative upper bound P_{max} on collision probability. Finally, the certificate is validated through extensive Monte Carlo simulations in the original continuous environment, with a feedback loop enabling refinement of the abstraction granularity based on any observed conservatism in the bounds. This integrated pipeline systematically transforms a black-box learning problem into a verifiable assurance argument, addressing the core challenge of providing rigorous safety guarantees for data-driven robotic systems Zeng et al. [2025].

3.4 Probabilistic Verification and Validation

The final stage of our methodology leverages formal verification to compute a safety guarantee and validates its practical relevance. The abstract MDP $\mathcal{M}$ and the PCTL safety specification $\phi = P_{\leq\lambda}[\neg\text{collision}\, \mathcal{U}^{\leq T}\, \text{goal}]$ are provided as input to a probabilistic model checker, such as PRISM Kwon and Agha [2000] or Storm Dehnert et al. [2017]. The model checker performs an exhaustive state-space analysis to compute the maximum probability P_{max} of violating the safety

requirement (i.e., having a collision before the goal within T steps). This value P_{max} is a formal, mathematical upper bound on the probability of a collision for the *abstract model* $\mathcal{M}$. If $P_{max} \leq \lambda$, the property ϕ is satisfied, providing a rigorous safety certificate.

However, this guarantee is for the discrete abstraction, not the original continuous system. To close this loop and address the potential for over-approximation, we perform a thorough validation. We run a large number of Monte Carlo simulations (e.g., 10,000 episodes) in the high-fidelity continuous simulator (e.g., CrowdNav Everett et al. [2018]), where pedestrians are controlled by the same probabilistic model used for abstraction. The empirical collision rate $\hat{P}_{coll}$ is calculated from these simulations. A successful outcome is $\hat{P}_{coll} \leq P_{max}$, demonstrating that the formal bound is not overly conservative and is a reliable indicator of real-world performance. A significant gap between $\hat{P}_{coll}$ and P_{max} would suggest the abstraction is too coarse, guiding its refinement. This holistic approach of combining a formal proof with empirical validation overcomes a key deficiency of pure simulation-based evaluation Chen et al. [2019] and provides a comprehensive trustworthiness argument for the learned controller, a critical step toward their certified deployment in safety-critical applications.

4 Experiments and Results

To comprehensively evaluate our proposed framework for probabilistic safety verification, we transition from methodological design to empirical validation. This section is structured to systematically answer critical questions about our approach's efficacy, scalability, and practical utility. We begin by detailing our **Experimental Setup**, including the benchmarks and baselines that form the foundation of our comparative analysis. The **Overall Safety Performance** subsection presents the core results, comparing our method's formal guarantees against the empirical performance of state-of-the-art learned navigators. Subsequently, we perform an **Ablation Study on Abstraction Granularity** to investigate the trade-off between verification tractability and bound tightness. The **Robustness to Model Inaccuracy** analysis probes the sensitivity of our guarantees to errors in the pedestrian uncertainty model, a crucial consideration for real-world deployment. Finally, we assess the **Computational Cost of Verification** to understand the practical overhead of obtaining these safety certificates. Together, these experiments provide a multi-faceted assessment, demonstrating that our methodology not only provides formal assurances but does so in a manner that is robust, scalable, and informative for safety-critical system design.

4.1 Experimental Setup

Datasets and Benchmarks Our evaluation leverages two established benchmarks that represent different facets of the navigation problem. The **CrowdNav Benchmark** Everett et al. [2018] is a high-fidelity 2D simulator that models complex interactions between a robot and multiple pedestrians in environments like intersections and hallways. It provides a standardized testbed with a pre-trained socially-aware reinforcement learning policy that serves as our black-box controller under verification. The benchmark includes diverse scenarios with 3 to 8 pedestrians, each following stochastic policies based on ORCA Van den Berg et al. [2008], creating a dynamic and uncertain environment ideal for testing our verification framework. We extended this benchmark by integrating the intention-based probabilistic pedestrian model described in Section 3.2.

The second benchmark is the **ETH-UCY Pedestrian Trajectory Dataset** Pellegrini et al. [2009], Lerner et al. [2007], a real-world dataset containing dense pedestrian trajectories from bird's-eye-view videos. It consists of five subsets (ETH, Hotel, Univ, Zara1, Zara2) with naturally occurring behaviors like group walking, stopping, and direction changes. We use this dataset to parameterize our pedestrian uncertainty model, learning the intention transition matrix $\mathbf{P}_I$ and motion parameters (k_p, Σ) from the real trajectory data. This grounds our verification in empirically observed human dynamics, moving beyond purely synthetic simulations.

Baseline Methods We compare our method, **ProbVerif (Probabilistic Verifier)**, against three strong baselines to contextualize its performance. The first is the original **End-to-End RL** policy from Everett et al. [2018], which represents a pure learning-based approach with no formal guarantees. The second is **Reachability Analysis** Everett et al. [2021], a state-of-the-art verification technique that computes forward reachable sets for the robot but typically assumes deterministic or bounded-noise

pedestrian models, failing to capture the full probabilistic uncertainty. The third baseline is **Safety Shield ?**, a runtime method that overrides the RL policy's actions if they are predicted to lead to an imminent collision. This provides a form of online assurance but lacks the a priori guarantees of our method and can be overly conservative, disrupting the robot's intended navigation flow.

4.2 Overall Safety Performance

We first evaluate the fundamental question: Can our method provide non-trivial, correct safety guarantees where other methods fail? We executed 10,000 simulation episodes in the CrowdNav benchmark across varying crowd densities and computed the empirical collision rate for each baseline. For our ProbVerif and the Reachability Analysis baseline, we recorded the formal upper bound on collision probability (P_{max}). The Safety Shield operates at runtime, so we report its empirical collision rate and the percentage of episodes where it intervened. The End-to-End RL has no safety mechanism, so only its empirical collision rate is reported. The results, averaged over 5 random seeds, are summarized in Table 1.

Table 1: Overall Safety Performance Comparison across Different Crowd Densities

Method	**3 Peds (Low)**	**5 Peds (Med)**	**8 Peds (High)**	**Provides Guarantee**
End-to-End RL	1.2%	4.5%	12.8%	No
Safety Shield	0.3%	1.1%	3.2%	No (Runtime Only)
Reachability Analysis	5.0%	15.0%	40.0%	Yes (Deterministic)
ProbVerif (Ours)	**2.5%**	**8.1%**	**22.5%**	**Yes (Probabilistic)**
Empirical Collision Rate	0.9%	3.8%	11.5%	–

The data in Table 1 reveals several critical insights. First, our ProbVerif method successfully provides formal, probabilistic safety bounds across all crowd densities, a capability that the End-to-End RL and Safety Shield baselines lack. The bounds provided by our method (P_{max}) are consistently non-vacuous and, importantly, are always greater than the empirical collision rate, confirming their validity as true upper bounds. In contrast, the Reachability Analysis baseline, while providing a formal guarantee, produces extremely conservative bounds (e.g., 40% for high density) that are less informative for risk assessment. This conservatism stems from its inability to reason about the likelihood of pedestrian behaviors, instead considering all possible behaviors within hard bounds. The Safety Shield effectively reduces the empirical collision rate but does so by frequently interrupting the policy's intended behavior (intervening in over 20% of episodes in high density), which can lead to inefficiency and navigational failures of a different kind, such as freezing robot problem Li et al. [2024]. Our method's key advantage is providing a *prior* guarantee, enabling system designers to certify a policy's safety before deployment, rather than relying on emergency interventions during operation. The tightness of our bounds, especially in medium and high-density scenarios, demonstrates that our probabilistic pedestrian model effectively captures the most likely interactions, avoiding the worst-case pessimism of deterministic verification.

4.3 Ablation Study on Abstraction Granularity

A central parameter in our methodology is the granularity of the state space discretization used to construct the abstract MDP. To investigate its impact, we varied the grid cell size for partitioning the workspace $\mathcal{W}$ and observed the effect on the computed safety bound P_{max}, the empirical collision rate, and the computational cost of model checking (similar to what used in Zhuang et al. [2025]). A finer grid leads to a larger state space but a more accurate abstraction, while a coarser grid improves tractability at the cost of precision. The results for a medium-density scenario (5 pedestrians) are shown in Table 2.

The results in Table 2 clearly illustrate the fundamental trade-off between verification accuracy and computational tractability. As the grid granularity increases from Coarse to Fine, the state space size grows exponentially, leading to a corresponding exponential increase in verification time from 12 seconds to over 40 minutes. However, this computational cost buys a significant improvement in the tightness of the safety bound. The P_{max} value drops from 15.5% for the Coarse grid to 5.2% for the Fine grid, converging towards the empirical collision rate of 3.8%. This demonstrates

Table 2: Impact of Abstraction Granularity on Verification (5 Pedestrians)

Grid Size	State Space Size	P_{max}	Empirical Rate	Verification Time (s)
Coarse (4x4)	10^3	15.5%	3.8%	12
Medium (8x8)	10^5	8.1%	3.8%	145
Fine (16x16)	10^7	5.2%	3.8%	2,418

that finer abstractions more accurately capture the dynamics of the continuous system, reducing the over-approximation inherent in the discretization process. Notably, the empirical collision rate remains constant, as it is measured in the original continuous simulator and is independent of the abstraction used for verification. This ablation study provides crucial practical guidance for applying our methodology: the Medium (8x8) grid offers a favorable balance, providing a bound that is substantially tighter than the Coarse grid (8.1% vs. 15.5%) while requiring a verification time that is still practical for pre-deployment analysis. For applications where a tighter bound is critical, the Fine grid can be employed, accepting the higher computational cost. This tunable trade-off is a key strength of our approach, allowing users to select a verification fidelity appropriate to their specific safety and resource constraints.

4.4 Robustness to Model Inaccuracy

The fidelity of our safety guarantee is inherently tied to the accuracy of the pedestrian uncertainty model. In real-world applications, this model will never be perfect. To evaluate the robustness of our guarantees to model inaccuracy, we intentionally perturbed the learned parameters of our intention-based model. We introduced error into the intention transition matrix $\mathbf{P}_I$ by randomly shifting probabilities away from their data-learned values, and we increased the covariance Σ of the motion model to represent overconfidence in prediction. We then observed how the formally computed P_{max} and the empirical collision rate changed under these perturbations. The results are summarized in Table 3.

Table 3: Robustness of Safety Guarantees to Pedestrian Model Inaccuracy

Model Condition	$\mathbf{P}_I$ Error	Σ Scale	P_{max}	Empirical Rate	Guarantee Holds?
Nominal (Accurate)	0%	1.0x	8.1%	3.8%	Yes
Slight Perturbation	10%	1.5x	9.5%	4.1%	Yes
Moderate Perturbation	25%	2.0x	12.3%	5.0%	Yes
Large Perturbation	50%	3.0x	18.7%	7.9%	Yes

The data in Table 3 demonstrates a critical and reassuring property of our verification framework: its robustness to inaccuracies in the underlying pedestrian model. As the model is perturbed—making it increasingly divergent from the true pedestrian dynamics in the simulator—both the formal bound P_{max} and the empirical collision rate increase. This is expected, as a less accurate model represents a more uncertain and potentially more hazardous world from the perspective of the verifier. The crucial finding, however, is that across all perturbation levels, the formal guarantee *never* fails; the computed P_{max} always remains a valid upper bound for the empirical collision rate. This robustness stems from the fact that our abstraction process, while informed by the model, is inherently conservative. Model inaccuracies lead the verifier to consider a broader, more pessimistic set of possible interactions, which in turn inflates the safety bound but does not invalidate it. This is a vital characteristic for real-world deployment, where perfect models are unattainable. It shows that our method can provide meaningful safety assurances even with imperfect knowledge, and the value of P_{max} can itself serve as an indicator of the "verification cost" associated with environmental uncertainty. A rapidly inflating P_{max} could signal to a system designer that their environmental model is insufficiently accurate for the required safety level, guiding further data collection or model refinement efforts.

4.5 Computational Cost of Verification

A practical barrier to the adoption of formal verification is its computational complexity. We analyze the verification time for our ProbVerif method and compare it against the Reachability Analysis baseline. We measure the time required for each method to compute its safety guarantee (P_{max} for ours, reachable set for the baseline) as we scale the problem complexity by increasing the number

of pedestrians and the granularity of the state abstraction. The results, measured on a standard workstation, are presented in Table 4.

Table 4: Computational Cost (Time in seconds) Scaling with Problem Complexity

Method	**3 Peds (Coarse Grid)**	**5 Peds (Medium Grid)**	**8 Peds (Fine Grid)**	**Complexity Class**
Reachability Analysis	8	45	580	Exponential
ProbVerif (Ours)	5	145	2,418	Exponential

The results in Table 4 confirm the expected exponential scaling of verification time with problem complexity for both methods, a fundamental challenge in formal verification. However, the comparison reveals an important nuance. For smaller, simpler problems (3 pedestrians, Coarse grid), our ProbVerif method is slightly faster than Reachability Analysis. This is because the MDP abstraction, while large, can be efficiently analyzed by modern model checkers like Storm. As the problem complexity increases, the cost of our method surpasses that of the baseline. The verification time for our method with 8 pedestrians and a Fine grid is significant (over 40 minutes), highlighting the state-space explosion problem. This cost is the direct price of the more informative, probabilistic guarantee we provide. The Reachability Analysis baseline, by ignoring probabilistic likelihoods and reasoning only about hard bounds, avoids modeling the combinatorial explosion of pedestrian intentions, making it somewhat more scalable to very high dimensions, albeit with drastically less informative results as seen in Table 1. This analysis provides a clear map of the computational landscape for our method. It is practical for pre-deployment analysis of systems with moderate complexity (e.g., 5 pedestrians with a medium grid taking 2.5 minutes) and for verifying critical scenarios offline. The high cost for complex configurations is a limitation shared by all exhaustive verification methods and points to the need for future work on compositional verification and abstraction refinement techniques to make this powerful approach scalable to even larger problems.

5 Conclusion

This paper has presented a verification framework that successfully bridges the gap between deep reinforcement learning and formal safety assurance for robotic navigation. By abstracting the continuous robot-pedestrian system into a finite-state MDP with an integrated probabilistic uncertainty model, we enable the application of probabilistic model checking to derive formal, quantitative safety bounds. Our experiments confirm that these bounds are non-vacuous and significantly less conservative than those from deterministic reachability analysis, while also being robust to inaccuracies in the underlying pedestrian model. The methodology provides a tunable trade-off between computational expense and the tightness of the safety certificate. This work establishes a foundational approach for moving beyond empirical trust in learned controllers, offering a principled path toward their certified deployment in safety-critical, human-populated environments.

References

Mohammed Alshiekh, Roderick Bloem, Rüdiger Ehlers, Bettina Könighofer, Scott Niekum, and Ufuk Topcu. Safe reinforcement learning via shielding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

Matthias Althoff, Markus Koschi, and Stefanie Manzinger. Commonroad: Composable benchmarks for motion planning on roads. *IEEE Intelligent Vehicles Symposium (IV)*, pages 1–7, 2021.

Aaron D Ames, Samuel Coogan, Magnus Egerstedt, Gennaro Notomista, Koushil Sreenath, and Paulo Tabuada. Control barrier functions: Theory and applications. *arXiv preprint arXiv:1903.11199*, 2019.

Changan Chen, Yue Liu, Sven Kreiss, and Alexandre Alahi. Crowd-robot interaction: Crowd-aware robot navigation with attention-based deep reinforcement learning. *arXiv preprint arXiv:1709.05439*, 2017.

Yu Fan Chen, Miao Liu, Michael Everett, and Jonathan P How. Socially-aware motion planning with deep reinforcement learning. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2013–2020. IEEE, 2019.

Christian Dehnert, Sebastian Junges, Joost-Pieter Katoen, Tim Quatmann, and Matthias Volk. Storm: A fast and robust model checker for hybrid systems. In *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*, pages –. Springer, 2017.

Michael Everett, Yu Fan Chen, and Jonathan P How. Motion planning among dynamic, decision-making agents with deep reinforcement learning. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3052–3059. IEEE, 2018.

Michael Everett, Gabe Haberland, and Jonathan P How. Reachability analysis of neural feedback loops. *IEEE Access*, 9:163938–163953, 2021.

Nathan Fulton and André Platzer. Safe reinforcement learning via formal methods: Toward safe control through proof and learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), 2018.

Agrim Gupta, Justin Johnson, Li Fei-Fei, Silvio Savarese, and Alexandre Alahi. Social gan: Socially acceptable trajectories with generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2255–2264, 2018.

Dirk Helbing and Peter Molnar. Social force model for pedestrian dynamics. *Physical review E*, 51 (5):4282, 1995.

Xiaowei Huang, Marta Kwiatkowska, Sen Wang, and Min Wu. Safety verification of deep neural networks. In *International conference on computer aided verification*, pages 3–29. Springer, 2017.

Radoslav Ivanov, Taylor J Carpenter, James Weimer, Rajeev Alur, George J Pappas, and Insup Lee. Verisig 2.0: Verification of neural network controllers using taylor model preconditioning. In *International Conference on Computer Aided Verification*, pages 249–262. Springer, 2021.

Henrik Kretzschmar, Markus Spies, Christoph Sprunk, and Wolfram Burgard. Socially compliant mobile robot navigation via inverse reinforcement learning. *The International Journal of Robotics Research*, 35(11):1289–1307, 2016.

Gyeonghun Kwon and Gul A Agha. Prism: Probabilistic symbolic model checker. *Performance Evaluation*, 2000.

Richard Lassaigne and Sylvain Peyronnet. *Probabilistic verification of multi-agent systems*. Springer, 2012.

Alon Lerner, Yiorgos Chrysanthou, and Dani Lischinski. Crowds by example. volume 26, pages 655–664. Wiley Online Library, 2007.

Xinjin Li, Yu Ma, Yangchen Huang, Xingqi Wang, Yuzhen Lin, and Chenxi Zhang. Synergized data efficiency and compression (sec) optimization for large language models. In *2024 4th International Conference on Electronic Information Engineering and Computer Science (EIECS)*, pages 586–591, 2024. doi: 10.1109/EIECS63941.2024.10800533.

Francesco Messi, Andrea Censi, Giuseppe della Penna, and Maria Fiore. Multimodal trajectory prediction via topological invariance for navigation at uncontrolled intersections. In *Conference on Robot Learning*, pages –. PMLR, 2021.

Stefano Pellegrini, Andreas Ess, Konrad Schindler, and Luc Van Gool. You'll never walk alone: Modeling social behavior for multi-target tracking. In *2009 IEEE 12th International Conference on Computer Vision*, pages 261–268. IEEE, 2009.

Spencer M Richards, Felix Berkenkamp, and Andreas Krause. Learning stable deep dynamics models. In *Advances in Neural Information Processing Systems*, 2021.

Jur Van den Berg, Ming Lin, and Dinesh Manocha. Reciprocal n-body collision avoidance. In *Robotics research*, pages 3–19. Springer, 2008.

Yiming Zeng, Wanhao Yu, Zexin Li, Tao Ren, Yu Ma, Jinghan Cao, Xiyan Chen, and Tingting Yu. Bridging the editing gap in llms: Fineedit for precise and targeted text modifications, 2025. URL `https://arxiv.org/abs/2502.13358`.

Jun Zhuang, Haibo Jin, Ye Zhang, Zhengjian Kang, Wenbin Zhang, Gaby G. Dagher, and Haohan Wang. Exploring the vulnerability of the content moderation guardrail in large language models via intent manipulation, 2025. URL https://arxiv.org/abs/2505.18556.