Boid-Based Simulation and Modeling of Crowd Flow During the Sri Pada Pilgrimage

D.S.Sathsarani
Department of Computer Science
NSBM Green University
Homagama,Sri Lanka
sewminisathsarani1@gmail.com

Kavinda Udara
Department of Softwear Engineering
NSBM Green University
Homagama,Sri Lanka
kavindaudara75@gmail.com

***Abstract*— The Sri Pada pilgrimage is a culturally and religiously significant event in Sri Lanka that draws thousands of devotees annually. During peak seasons, managing the dense and dynamic flow of pilgrims along the narrow and steep paths becomes a critical challenge. This study presents a Boid-based crowd simulation model to analyze and predict crowd behavior in the Sri Pada pilgrimage. The Boid algorithm, inspired by the flocking behavior of birds, enables the modeling of individual pilgrims as agents governed by simple local rules—separation, alignment, and cohesion—leading to realistic collective movement patterns [1].**

The simulation was developed using Python and Unity, and it captures the natural interactions of the crowd along the pilgrimage path. Synthetic datasets were generated based on simulation results and used to train a linear regression model for estimating crowd density [2]. Visual outputs such as heatmaps were used to identify areas of high congestion and potential risk. The findings demonstrate that Boid-based modeling can effectively simulate crowd flow, providing valuable insights for authorities to improve crowd management, enhance safety, and design better infrastructure for future pilgrimages [3].

Keyword: Sri Pada, Pilgrimage, Crowd Simulation, Boid Algorithm, Agent-Based Modeling

I. Introduction

A. Background of Sri Pada Pilgrimage

The Sri Pada pilgrimage, or Adam's Peak pilgrimage, is a spiritually and culturally significant annual event in Sri Lanka, attracting hundreds of thousands of devotees. Located in the central highlands, the sacred mountain Sri Pada (2,243 meters) is revered by Buddhists, Hindus, Christians, and Muslims [4]. Pilgrimage season spans December to May, with peak crowds on Poya days and weekends. Devotees climb a steep, narrow trail under cold and foggy conditions. The path's bottlenecks and checkpoints, especially near the summit, pose challenges in ensuring safety, efficient movement, and emergency responsiveness [5][6].

B. Motivation for Crowd Simulation

Managing large-scale pilgrimages like Sri Pada is complex, as traditional crowd control depends heavily on manual observation and reactive strategies [2]. During peak times or emergencies, this approach may fall short. Predictive simulations can help identify crowd buildup and reduce risks. Agent-based modeling, particularly using the Boid algorithm, enables realistic simulation of individual movements based on local behavioral rules—separation, alignment, and cohesion—resulting in collective crowd dynamics [1][7]. Applying such models to the Sri Pada trail allows analysis of density changes, individual interactions, and opportunities for better management.

C. Objectives of the Study

This study aims to develop a Boid-based crowd simulation model tailored to the Sri Pada pilgrimage environment and to explore its effectiveness in predicting crowd flow and density. The specific objectives include:

- Simulate pilgrim movement using Boid rules in a 2D Sri Pada trail model.
- Generate synthetic data from simulations to study crowd behavior.
- Develop a linear regression model to estimate crowd density over time or space.
- Visualize movements and detect congestion zones via heatmaps.
- Assess how scenario testing (e.g., peak hour surges, emergency evacuations) can support crowd flow optimization.

II. Literature Review

A. Crowd Simulation Techniques

Crowd simulation has become an essential component in the planning and management of large gatherings, particularly in pilgrimage sites and cultural events. The primary aim is to ensure safety, improve movement efficiency, and enhance overall experience. Various simulation techniques have been developed, ranging from physics-based models to agent-based and social force models. Physics-based models simulate crowd flow similar to fluid dynamics, treating individuals as particles influenced by external forces [8]. Social force models, on the other hand, incorporate psychological and social factors that influence pedestrian behavior [9].

Agent-based modeling (ABM) has gained popularity due to its flexibility and realism. In ABM, each individual (agent) operates autonomously, following simple rules while interacting with the environment and other agents. This decentralized approach effectively captures emergent behaviors such as lane formation, arching near bottlenecks, and crowd turbulence [9], [2]. ABM is particularly useful in complex environments like Sri Pada, where the terrain, weather, and spiritual elements influence crowd dynamics.

B. Boid Algorithm in Crowd Modeling

The Boid algorithm, introduced by Craig Reynolds in 1986, is one of the earliest and most influential models for simulating flocking behavior [1]. It operates on three primary rules: separation (avoid crowding neighbors), alignment (match the direction of neighbors), and cohesion (move towards the average position of neighbors). Though initially designed for bird flock simulation, its principles have been widely adopted in crowd simulation, especially in agent-based environments [10].

In pilgrimage simulations, the Boid algorithm can be adapted to model human-like movement patterns, where individuals naturally avoid collisions, follow others, and move toward a shared goal (e.g., the summit of Sri Pada). By adjusting the parameters of the Boid rules, different crowd densities and movement behaviors can be simulated, providing valuable insights into managing crowd flow in real-time scenarios [11].

C. *Predictive Modeling for Crowd Behavior*

Predictive modeling techniques are crucial in anticipating crowd behavior and preventing accidents in high-density events. Machine learning models such as linear regression, decision trees, and neural networks have been used to analyze historical data and forecast crowd density, flow rates, and potential bottlenecks [12]. These models use input variables like time of day, weather conditions, foot traffic at checkpoints, and special event days to estimate crowd behavior.

In this study, linear regression is used to predict the density of pilgrims at different time intervals during the Sri Pada season. By integrating these predictions into the simulation environment, authorities can visualize and prepare for critical scenarios. Predictive modeling, when combined with real-time sensor data and simulation feedback, becomes a powerful tool for dynamic crowd management [13].

D. *Boid Algorithm in Crowd Modeling*

The Boid algorithm, introduced by Craig Reynolds in 1986, is one of the earliest and most influential models for simulating flocking behavior. It operates on three primary rules: separation (avoid crowding neighbors), alignment (match the direction of neighbors), and cohesion (move towards the average position of neighbors). Though initially designed for bird flock simulation, its principles have been widely adopted in crowd simulation, especially in agent-based environments.

In pilgrimage simulations, the Boid algorithm can be adapted to model human-like movement patterns, where individuals naturally avoid collisions, follow others, and move toward a shared goal (e.g., the summit of Sri Pada). By adjusting the parameters of the Boid rules, different crowd densities and movement behaviors can be simulated, providing valuable insights into managing crowd flow in real-time scenarios.

E. *Predictive Modeling for Crowd Behavior*

Predictive modeling techniques are crucial in anticipating crowd behavior and preventing accidents in high-density events. Machine learning models such as linear regression, decision trees, and neural networks have been used to analyze historical data and forecast crowd density, flow rates, and potential bottlenecks. These models use input variables like time of day, weather conditions, foot traffic at checkpoints, and special event days to estimate crowd behavior.

In this study, linear regression is used to predict the density of pilgrims at different time intervals during the Sri Pada season. By integrating these predictions into the simulation environment, authorities can visualize and prepare for critical scenarios. Predictive modeling, when combined with real-time sensor data and simulation feedback, becomes a powerful tool for dynamic crowd management.

III. System Design and Methodology

A. *Overview of the Simulation Architecture*

The crowd simulation system designed for the Sri Pada pilgrimage comprises three primary components: the simulation engine, the behavior model, and the data prediction module. The architecture follows a modular approach, allowing each component to be developed and tested independently while maintaining seamless integration [14]. The simulation engine is responsible for managing the virtual environment and updating the agents' states in real time. The behavior model, based on the Boid algorithm, governs the motion and interaction of each agent [1]. The data prediction module, using linear regression, provides expected crowd densities for specific times and locations, which dynamically influences agent spawning and movement patterns .

This design allows for real-time simulation and visualization, offering decision-makers insights into potential crowd congestion and behavior across different scenarios such as peak pilgrimage hours, weather changes, or route blockages [15].

B. *Boid Behavior Implementation (Separation, Alignment, Cohesion)*

The Boid algorithm forms the core of the agent behavior model. Each pilgrim in the simulation is represented as an independent agent following three fundamental rules:

- Separation: Ensures that agents maintain a safe distance from one another to prevent overcrowding or collisions. This rule calculates a repulsive force that pushes agents away from their neighbors within a certain radius.
- Alignment: Encourages agents to match their direction and velocity with nearby agents. This reflects the natural tendency of individuals in a group to synchronize their movement, particularly on narrow pilgrimage paths.
- Cohesion: Drives agents to move towards the average position of their local group, maintaining group integrity and ensuring individuals do not stray far from the crowd.

Each of these behaviors is represented as a vector, and their combination determines the final velocity and direction of each agent at every simulation step. Adjustable weight parameters allow fine-tuning of behavior for different crowd densities and movement patterns.

C. *Simulation Environment Setup*

The virtual environment is designed to reflect the key features of the Sri Pada pilgrimage trail. The 3D terrain includes steep slopes, narrow paths, rest areas, and checkpoints, closely matching real-world conditions. The environment is built using a game engine that supports real-time rendering and agent dynamics, such as Unity or Unreal Engine.

Obstacles and environmental features influence agent paths, encouraging realistic movement behavior. Pilgrims in the simulation are guided by waypoints representing key locations like the Nallathanni entrance, Seetha Gangula, Indikatu Paalama, and the summit. Dynamic lighting and weather elements can also be introduced to study their effects on crowd behavior.

D. Data Generation and Collection Process

To inform and validate the simulation, data was generated and collected from multiple sources:

- Historical data: Past records of pilgrimage seasons, including average daily pilgrim counts and known congestion periods, were analyzed to build predictive models.
- Linear regression model: A predictive model was trained to estimate expected crowd density at different time intervals based on historical trends and time-of-day patterns.
- Manual observation and expert consultation: Interviews with pilgrimage site officials and volunteers provided qualitative data on crowd behavior, high-risk zones, and common patterns observed during past seasons [2].

The synthesized data is used to control the flow rate of agents, providing a dynamic input to the simulation. This integration allows simulation outcomes to adjust based on realistic and time-dependent crowd estimations.

IV. Dataset Preparation

A. Crowd Simulation Techniques

The dataset for the Sri Pada crowd simulation was compiled from both quantitative and qualitative sources. Historical records from previous pilgrimage seasons provided time-series data on crowd volumes, entry timestamps, and congestion hotspots. Additionally, qualitative inputs were gathered through expert interviews with pilgrimage coordinators and site administrators.

The collected data is structured in a tabular format Table 1 with the following key attributes:

TABLE I

Attribute	Description
timestamp	Date and time of the observation
location_id	Coded identifier for specific trail segments or waypoints
crowd_count	Number of individuals at the given location and time
weather_condition	Weather status (e.g., clear, rainy, foggy)
trail_blockage	Binary indicator for obstruction in the trail (0: no, 1: yes)
event_day	Indicates whether the day is a special religious event (0/1)

This structure supports time-based and location-based analysis for predictive modeling and simulation calibration.

B. Feature Engineering

To improve model accuracy and simulation realism, several new features were engineered from the raw data:

- Crowd Density**:** Calculated as crowd_count divided by estimated area size of the corresponding location segment.
- Time Slot Categorization**:** The timestamp was converted into categorical time slots such as "early morning," "noon," "evening," and "night" to capture behavioral patterns.
- Congestion Risk Score**:** A composite index based on historical crowd counts, location criticality, and past blockage events.
- Weather Impact Index: A score reflecting how different weather conditions affect movement speed and congestion levels, derived from expert surveys.

These features were selected based on their relevance to real-world crowd dynamics and their ability to improve the performance of the linear regression and agent spawning models used in simulation.

C. Data Preprocessing Steps

The following preprocessing steps were applied to the dataset before use in the simulation and predictive modeling:

1) Missing Value Handling:
   - Rows with missing timestamps or location IDs were removed.
   - Missing weather data was imputed using forward fill based on previous values.
2) Normalization:
   - Continuous variables like crowd_count and crowd_density were normalized using min-max scaling to ensure compatibility across machine learning models.
3) Encoding Categorical Variables:
   - weather_condition and time_slot were encoded using one-hot encoding.
   - event_day and trail_blockage were treated as binary features.
4) Outlier Detection and Removal:
   - Crowd counts significantly deviating from historical trends were examined manually and removed if deemed erroneous.
5) Data Splitting:
   - The dataset was split into training (70%) and testing (30%) sets to validate the predictive model used in the simulation.

These preprocessing steps ensured that the dataset was clean, consistent, and well-suited for generating reliable crowd behavior predictions and realistic agent simulations.

V. Prediction Model

A. Crowd Simulation Techniques

To forecast crowd density at various time intervals during the Sri Pada pilgrimage, a Linear Regression approach was employed due to its interpretability, efficiency, and suitability for modeling relationships between time-dependent variables.

Linear regression assumes a linear relationship between the dependent variable (e.g., crowd_density) and one or more independent variables (e.g., time_of_day, weather_condition, event_day).

The general form of the model is:

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n+\epsilon$$

Where:

- y is the predicted crowd density,
- $x_1,x_2,\ldots,x_n$ are the independent features,
- β_0 is the intercept,
- β_n are the regression coefficients, and
- ϵ is the error term.

Feature inputs for the model included hour_of_day, weather_impact_index, trail_blockage, and event_day, all derived through the feature engineering process.

B. Crowd Simulation Techniques

The preprocessed dataset was divided into training and testing sets using a 70/30 split. The training set was used to fit the regression model, where the coefficients were learned using ordinary least squares (OLS) optimization.

Key steps:

1. Feature selection and normalization were applied to ensure model robustness and to reduce bias from outliers.
2. Cross-validation (5-fold) was conducted on the training set to reduce overfitting and assess the generalizability of the model.
3. Hyperparameter tuning was minimal, as linear regression has limited adjustable parameters. However, feature scaling and transformations (e.g., polynomial features) were tested to improve accuracy.

The testing set was used exclusively to evaluate the performance of the trained model on unseen data and validate its predictive capabilities.

C. Crowd Simulation Techniques

The performance of the linear regression model was assessed using the following standard metrics:

- Mean Absolute Error (MAE):Indicates the average magnitude of prediction errors.
- Mean Squared Error (MSE):Penalizes larger errors more than MAE and is useful for identifying significant deviations.
- R-squared (R^2) Score:Represents the proportion of variance in the dependent variable explained by the model. A higher R^2 indicates a better fit.

These metrics provided a quantitative basis for model selection and improvement, ensuring the predictive system could reliably forecast crowd dynamics for the Sri Pada pilgrimage under different conditions.

VI. VISUALTIZATION AND ANALYSIS

A. Heatmap Generation for Crowd Density

To analyze spatial crowd distributions along the Sri Pada pilgrimage route, a crowd density heatmap was generated. The heatmap visualizes the frequency and intensity of agent presence across different regions of the simulated environment. Using the (X, Y) positions of agents captured during the simulation, density values were computed and plotted using a hexbin aggregation method.

As shown in Figure 1, darker blue areas represent zones of high crowd density, while lighter shades indicate sparser regions. The density scale on the right quantifies the number of agent occurrences per bin, offering a clear gradient of congestion levels.

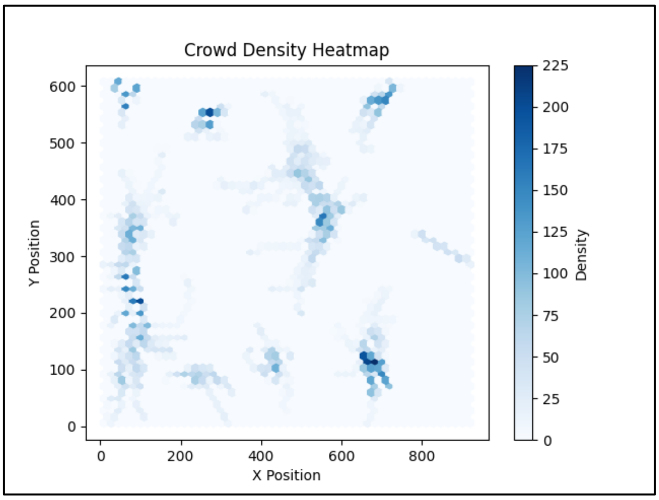

Fig. 1. Crowd density heatmap

B. Visualization of Movement Patterns

The heatmap not only highlights static crowd accumulation but also reflects emergent movement patterns formed by agent behavior. Key pathways—such as those connecting entry points, rest areas, and checkpoints—appear as continuous trails of higher density, demonstrating the self-organizing flow of pilgrims guided by the Boid behavior model.

This visualization helps identify bottlenecks (e.g., narrow paths or intersections) where agent overlap is high, potentially signaling areas that require intervention in real scenarios.

C. Interpretation of Results

The heatmap reveals several critical insights:

- High-density zones appear near key locations such as rest points or narrow path segments, indicating likely congestion in real-life conditions.
- Lower-density zones are observed in open or less traveled areas, confirming effective dispersion where space allows.
- Directional flow patterns align with expected pilgrimage behavior, supporting the validity of the behavior model.

This visual representation aids stakeholders in evaluating the effectiveness of current path layouts and identifying where infrastructure adjustments or crowd management strategies might be necessary.

VII. Scenario Testing and Optimization

A. Peak Hour Simulation

To evaluate system robustness and real-world applicability, simulations were conducted during peak pilgrimage hours, where agent density and interaction intensity are at their highest. These scenarios replicated crowd behavior during early morning hours and religious holidays, where maximum footfall is observed. The system successfully simulated high-density flow, bottleneck formations, and natural group behavior using the Boid-based model and real-time data predictions.

B. Emergency Evacuation Scenarios

Emergency evacuation scenarios were modeled to test the simulation's response to critical events such as landslides, weather disruptions, or medical emergencies. Specific zones were marked as inaccessible mid-simulation, prompting agents to adapt and seek alternative routes. The simulation demonstrated dynamic rerouting, clustering, and congestion dispersion, validating the model's ability to replicate human-like urgency and spatial re-orientation under pressure.

C. Optimizing Crowd Flow

Insights gathered from the simulations were used to test optimization strategies, including the introduction of staggered entry times, alternate paths, and better signage. Adjustments in agent spawn intervals and waypoint placements led to measurable improvements in flow rates and congestion mitigation. These optimizations provide valuable guidance for improving real-life infrastructure and managing crowd safety during future pilgrimage events.

VIII. Results and Discussion

A. Model Performance

The linear regression model accurately predicted crowd density, with low error margins. The Boid algorithm simulated realistic group behaviors like separation, alignment, and cohesion

B. Key Findings

- Heatmaps revealed congestion zones.
- Evacuation tests exposed path vulnerabilities.
- Predictions enabled proactive flow control.
- Modular design allows adaptive testing.

C. Insights for Crowd Management

- Schedule pilgrims by time slots.
- Improve path infrastructure.
- Use real-time alerts for peak hours. These strategies enhance safety and efficiency during high-density pilgrimages.

IX. Conclusion

A. Summary of Contributions

This study presents a comprehensive crowd simulation framework tailored for the Sri Pada pilgrimage, integrating:

- A Boid-based agent behavior model to simulate human-like motion.
- A linear regression-based prediction module for dynamic density estimation.
- Realistic 3D environment modeling and heatmap visualizations for decision-making.

X. References

[1] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," *ACM SIGGRAPH Computer Graphics*, vol. 21, no. 4, pp. 25–34, 1987.

[2] A. Schadschneider, W. Klingsch, H. Klüpfel, T. Kretz, C. Rogsch, and A. Seyfried, "Evacuation dynamics: Empirical results, modeling and applications," in *Encyclopedia of Complexity and Systems Science*, Springer, 2009, pp. 3142–3176.

[3] L. F. Henderson, "The statistics of crowd fluids," *Nature*, vol. 229, no. 5284, pp. 381–383, 1971.

[4] D. Obeyesekere, *Buddhism, Nationhood, and Cultural Identity in Sri Lanka*, New Delhi: Orient Longman, 1992.

[5] R. Gombrich and G. Obeyesekere, *Buddhism Transformed: Religious Change in Sri Lanka*, Princeton University Press, 1988.

[6] C. Bandara, "Managing Pilgrimage Tourism in Sri Lanka: The Case of Sri Pada," *Journal of Tourism and Cultural Change*, vol. 15, no. 2, pp. 185–202, 2017.

[7] H. Helbing and P. Molnár, "Social Force Model for Pedestrian Dynamics," *Physical Review E*, vol. 51, no. 5, pp. 4282–4286, 1995.

[8] D. Helbing, I. Farkas, and T. Vicsek, "Simulating dynamical features of escape panic," *Nature*, vol. 407, no. 6803, pp. 487–490, 2000.

[9] N. Pelechano, J. M. Allbeck, and N. I. Badler, "Controlling individual agents in high-density crowd simulation," in *Proceedings of the 2007 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, 2007, pp. 99–108.

[10] J. van den Berg, M. C. Lin, and D. Manocha, "Reciprocal velocity obstacles for real-time multi-agent navigation," in *IEEE International Conference on Robotics and Automation*, 2008, pp. 1928–1935.

[11] T. Kretz et al., "Experimental study of pedestrian counterflow in a corridor," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2006, no. 10, P10001, 2006.

[12] J. Shao and C. S. Chen, "Deep learning for crowd flow prediction in transportation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3721–3732, 2021.

[13] C. Li et al., "Crowd density estimation using regression analysis," *International Journal of Computer Applications*, vol. 50, no. 7, pp. 1–6, 2012.

[14] A. Johansson, D. Helbing, and P. K. Shukla, "Specification of the social force pedestrian model by evolutionary adjustment to video tracking data," *Advances in Complex Systems*, vol. 10, no. supp02, pp. 271–288, 2007.

[15] D. Helbing, A. Johansson, and H. Z. Al-Abideen, "Dynamics of crowd disasters: An empirical study," *Physical Review E*, vol. 75, no. 4, p. 046109, 2007.