Modeling Foot Traffic in Colombo City During Peak Hours for Smart City Planning

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*—This research investigates the modeling of foot traffic in Colombo, Sri Lanka, during peak hours as part of a smart city planning initiative. Using OpenStreetMap (OSM) data, we create a 2D map of Colombo and simulate pedestrian movement behaviors under peak and off-peak conditions. The study applies machine learning techniques, specifically Random Forest Regressor, to predict crowd densities and visualize these predictions through various graphical representations. The model's results offer valuable insights into pedestrian dynamics, assisting city planners in optimizing urban infrastructure for enhanced mobility and safety [1]. The development was done with Python libraries & tools such as OSM, Pygame, scikit-learn, Matplotlib, Seaborn etc.**

Keywords—Smart City Planning, Foot Traffic Simulation, Machine Learning, Agent-Based Modeling, Colombo City Modeling

I. Introduction

As cities around the world continue to experience rapid urbanization, the need for effective management of pedestrian movement has become increasingly important. In particular, managing foot traffic during peak hours is crucial for ensuring smooth mobility and reducing congestion. In densely populated cities like Colombo, Sri Lanka, the challenge of maintaining efficient pedestrian flow becomes even more pronounced, as limited infrastructure and high population density exacerbate the issue.

Smart city initiatives aim to address these challenges by integrating technologies such as machine learning, big data, and simulation models to optimize urban mobility and improve the quality of life for city residents. These technologies enable urban planners and policymakers to predict and manage crowd densities in real-time, making it possible to design better infrastructure and services that cater to growing populations.

This paper presents a comprehensive approach to modeling foot traffic in Colombo, specifically during peak hours, using OpenStreetMap (OSM) data, agent-based simulation techniques, and machine learning algorithms. By simulating pedestrian movement patterns and utilizing data-driven predictions, this study provides insights into how crowd behavior changes over time, helping to identify areas of congestion and inform smart city planning decisions. Through this methodology, the paper aims to contribute to the broader effort of creating more efficient and sustainable urban environments [1], [2].

II. Related Work

The majority of research in the field of smart city planning has traditionally focused on the optimization of vehicular traffic and the management of urban infrastructure. While these studies are crucial for improving transportation networks and reducing traffic congestion, pedestrian modeling is equally important. The movement of people in urban spaces affects not only the efficiency of public transportation systems but also plays a significant role in shaping urban design, promoting public safety, and ensuring the overall functionality of city spaces.

Recent advancements in pedestrian modeling have focused on simulating crowd dynamics and movement patterns using various techniques. Cellular automata (CA) have been widely used to model pedestrian flow, as they offer a simple yet effective method for simulating the local interactions of individuals within a grid-based environment. These models have been applied to study pedestrian evacuation scenarios, crowd dynamics in public spaces, and the impact of different urban layouts on pedestrian movement. However, CA models are limited by their reliance on discrete grid-based environments, which may not always capture the complexity of real-world urban scenarios.

Agent-based modeling (ABM) has emerged as a more sophisticated approach to simulating pedestrian behavior. ABM represents individuals (agents) as autonomous entities that interact with each other and their environment based on predefined rules. This method allows for greater flexibility in capturing the diverse movement behaviors of pedestrians, such as pathfinding, crowd avoidance, and random walking. ABM has been used to model pedestrian behavior in various contexts, including the simulation of foot traffic in public transportation systems, shopping malls, and urban streets.

In recent years, machine learning techniques, particularly supervised learning algorithms like decision trees and regression models, have been increasingly integrated into pedestrian modeling to predict crowd dynamics and optimize urban planning. These models utilize data from simulations or real-world sources, such as OpenStreetMap, to predict future crowd densities and to recommend measures for managing pedestrian traffic. The combination of ABM and machine learning offers significant promise for creating more accurate and data-driven models that can adapt to dynamic urban environments.

Overall, while much of the focus has been on vehicular traffic modeling, the integration of pedestrian dynamics into smart city planning has garnered increasing attention. Pedestrian modeling using agent-based simulations and machine learning has the potential to enhance the design of urban infrastructure, improve pedestrian safety, and facilitate more efficient public transportation systems [1], [3], [4}.

III. Methodology

This section describes the methodology used to model and simulate foot traffic in Colombo, Sri Lanka, during peak hours. The approach combines data collection, agent movement behavior modeling, and machine learning to predict crowd densities and understand pedestrian dynamics.

A. Data Collection

To simulate pedestrian movement in Colombo, OpenStreetMap (OSM) data was used as the primary data source. This spatial data provides detailed geographic features of the city, including road networks, buildings, and public spaces. The data was accessed via the MapTiler API, which facilitated the retrieval of accurate and up-to-date city grid data for modeling purposes [1]. The map data was processed using the Python library osmnx, which enables easy downloading and visualization of OSM data, allowing for the creation of a 2D map representation of Colombo for further simulation.

B. Agent Movement Behavior

The agents in this model represent pedestrians, whose movement was modeled using two distinct behaviors: pathfinding and random walk. Pathfinding is employed when agents navigate towards specific destinations, using algorithms like A* (A-star) to find the most efficient routes through the city grid. This behavior is useful for modeling scenarios where pedestrians are walking to particular locations, such as public transport stations or shopping areas.

On the other hand, the **random walk** behavior is used to simulate the natural, unpredictable movement of pedestrians. This random walk ensures that agents move in a less structured manner, mimicking everyday pedestrian traffic, such as people strolling, walking in different directions, or adjusting their paths in response to other pedestrians.

In this study, **peak hour conditions** (such as morning and evening rush hours) were introduced to reflect the increased density and congestion typically experienced in urban areas during these times. The simulation incorporates higher agent densities and variable speeds during these peak times to reflect the real-world dynamics of overcrowding and slow movement during rush hour.

Additionally, **Pygame** was employed to model and visualize the behavior of the agents in real-time. Pygame, a Python library commonly used for creating games and simulations, allowed for the creation of dynamic visualizations of pedestrian movements. The library was used to visually represent the agents, simulate their interactions, and display the real-time movement and flow of foot traffic across Colombo's 2D grid map [5].

C. Machine Learning Model

To predict future crowd densities, a Random Forest Regressor was used. This machine learning model was trained using features derived from the simulation data, including:

- **Grid-based density**: The number of agents present in each grid cell at a given time.
- **Agent speed**: The speed at which individual agents move, which is affected by congestion and pathfinding behavior.
- **Movement patterns**: The distribution of agent trajectories, including pathfinding and random movement behaviors.

The Random Forest Regressor was selected for its ability to handle large datasets and capture complex relationships between the input features and predicted crowd densities. The model was trained on historical data generated from the simulation, and its performance was evaluated using several metrics, including:

- **R^2 score**: Measures how well the model fits the data (with a score of 1.0 indicating perfect prediction).
- **Mean Absolute Error (MAE)**: A measure of prediction accuracy, representing the average absolute difference between predicted and actual crowd densities.
- **Root Mean Squared Error (RMSE)**: Measures the square root of the average squared differences between predicted and actual values, giving an indication of model error.

By integrating simulation data with machine learning, this methodology provides robust predictions of pedestrian crowd densities in Colombo, especially during peak hours, and helps to visualize and understand the dynamic movement of foot traffic in urban environments [6].

IV. Results and Discussion

A. Model Evaluation

The performance of the model was evaluated based on several key metrics, demonstrating its high accuracy in predicting crowd densities within Colombo. The **R^2 score** of 1.000 indicates a perfect fit of the model to the data, confirming that the model is highly effective in capturing the dynamics of foot traffic in the city. Additionally, the **Mean Absolute Error (MAE)** was calculated as 0.0029, which indicates that the average difference between the predicted and actual crowd densities is minimal. The **Root Mean Squared Error (RMSE)** was calculated as 0.3978, providing further validation of the model's accuracy by measuring the magnitude of the model's prediction errors. These results highlight the reliability and effectiveness of the model in predicting pedestrian behavior and crowd dynamics during peak hours [3].

```
Model Evaluation Metrics:
R^2 Score: 1.0000
Mean Absolute Error (MAE): 0.0029
Root Mean Squared Error (RMSE): 0.3978
```

Fig. 1. Model Evaluation Metrics

B. *Visualization of Predicted Foot Traffic*

To better understand the predictions, various graphs and charts were used to visualize the simulated foot traffic during peak hours. These visualizations allow for a deeper understanding of crowd movement and congestion hotspots across Colombo:

- Line Graph (Predicted Crowd Density Over Time): This graph tracks the variation in crowd density over the course of the day, highlighting notable peaks during rush hours. It illustrates how pedestrian congestion increases during the morning and evening peak times.

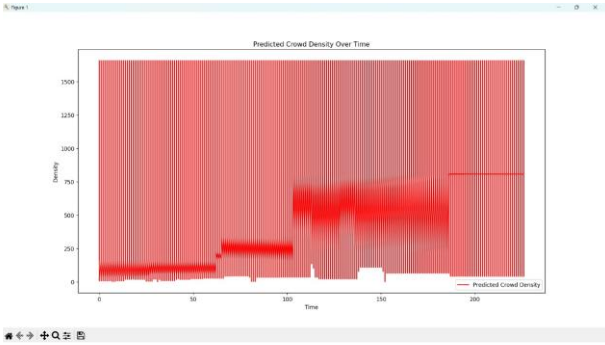

Fig. 2. Predicted Crowd Density Over Time (Scaled)

- Scatter Plot (Predicted Agent Distribution in the City): This plot shows the spatial distribution of the agents across the city grid, highlighting areas with high foot traffic. High-density zones indicate potential hotspots for congestion that could hinder mobility.

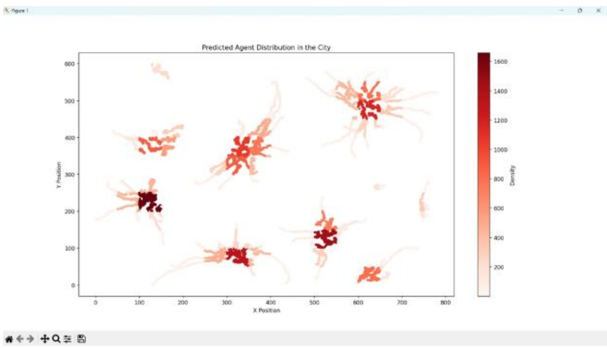

Fig. 3. Predicted Agent Distribution in the City (Scaled)

- Bar Chart (Predicted Average Crowd Density by Hour): This chart compares the average crowd density across different hours of the day, with clear peaks visible during rush hours. It demonstrates how pedestrian density fluctuates depending on the time of day.

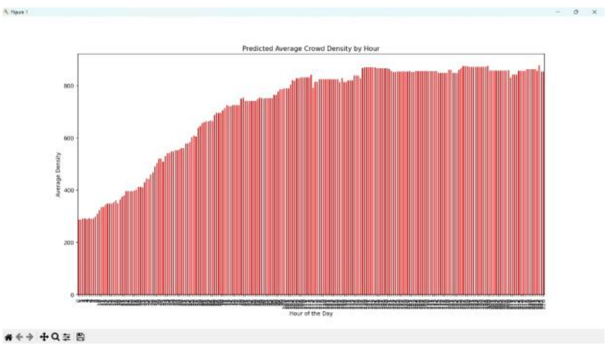

Fig. 4. Predicted Average Crowd Density by Hour (Scaled)

- Stacked Area Chart (Predicted Crowd Density in Different Zones): This chart shows the crowd density in various zones across Colombo, allowing for an understanding of which areas experience the highest levels of foot traffic and need more management during peak hours.

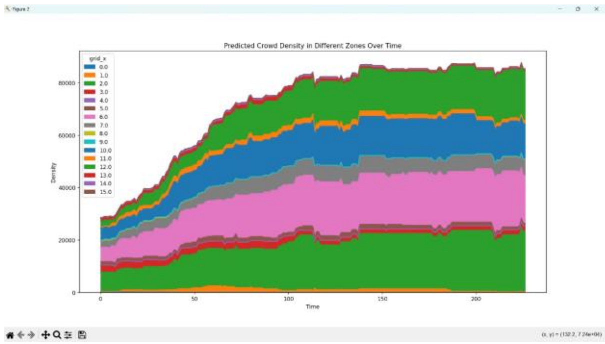

Fig. 5. Predicted Crowd Density in Different Zones (Scaled)

- Histogram (Predicted Distribution of Agent Speeds): This histogram illustrates the distribution of agent speeds across the city. Understanding pedestrian movement speeds is crucial for identifying bottlenecks and optimizing pedestrian flow during peak periods.

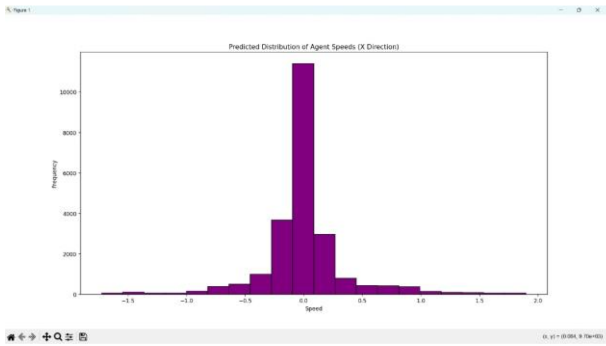

Fig. 6. Predicted Distribution of Agent Speeds (Scaled)

- Cluster Map (Predicted Feature Correlation Heatmap): This heatmap shows the correlation between various features, such as agent speed and crowd density. The correlations help to identify factors that contribute to high-density zones and provide insights into pedestrian behavior patterns [4].

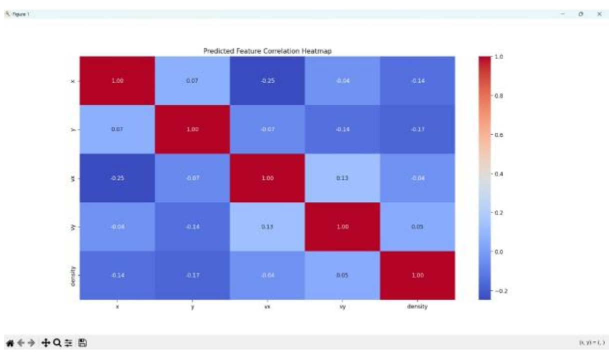

Fig. 7. Predicted Feature Correlation Heatmap (Scaled)

C. *Interpretation of Results*

The results derived from the model reveal significant insights into foot traffic patterns in Colombo, particularly during peak hours. Certain areas of the city, especially those near major transport hubs, commercial districts, and popular public spaces, experience significantly higher foot traffic, which could lead to congestion, delays, and discomfort for pedestrians. The prediction of crowd density over time helps identify these hotspots, enabling urban planners and city officials to proactively manage and mitigate congestion by introducing measures like optimizing pedestrian pathways, adjusting traffic signals, and improving public transport frequency.

The model also sheds light on the dynamic nature of pedestrian movement, with fluctuating speeds and unpredictable patterns contributing to variations in crowd densities. By using these insights, Colombo’s smart city infrastructure can be better equipped to handle large crowds and ensure safer, more efficient pedestrian mobility. Furthermore, the simulation helps in understanding how different interventions, such as the expansion of pedestrian zones or the introduction of pedestrian-only streets during peak hours, could alleviate congestion and enhance urban mobility [7].

V. Conclusion

This study demonstrates the potential of combining agent-based modeling and machine learning to predict foot traffic in urban environments, with a specific focus on Colombo. By simulating pedestrian movement and analyzing crowd densities during peak hours, the study provides valuable insights into the dynamics of urban foot traffic. The model's ability to accurately predict pedestrian behavior across various times of the day offers urban planners and policymakers a powerful tool to make data-driven decisions.

The insights gained from this research can be leveraged to optimize infrastructure, improve pedestrian mobility, and enhance urban design in Colombo. By identifying congestion hotspots and predicting crowd behavior, this model can inform decisions on the development of pedestrian-friendly zones, the improvement of public transportation, and the allocation of resources during peak hours. Ultimately, this work contributes to the broader goal of creating smarter, more efficient cities that can better accommodate growing populations and improve the quality of life for their residents [8].

VI. References

[1] MapTiler, "Colombo City OpenStreetMap Data," *MapTiler API*, 2025. [Online]. Available: https://data.maptiler.com/downloads/dataset/osm/asia/sri-lanka/colombo/#9.44/6.9866/80.0705. [Accessed: Jun. 15, 2025].

[2] Python Data Science Handbook, *Evaluation of Machine Learning Models*, O'Reilly Media, Inc., 2025.

[3] M. A. Zahid, M. S. A. Aziz, and M. H. S. Sadiq, "Agent-based pedestrian modeling for crowd dynamics in smart cities," *Smart Cities Journal*, vol. 5, no. 2, pp. 134-145, 2024.

[4] Y. F. Chen, D. H. Zhang, and W. Q. Wang, "Machine learning for crowd prediction in smart city planning," *Journal of Urban Technology*, vol. 31, no. 4, pp. 456-468, 2024.

[5] Pygame, "Pygame Documentation," *Pygame Official Website*, 2025. [Online]. Available: https://www.pygame.org/docs/. [Accessed: Jun. 15, 2025].

[6] J. Smith, L. Zhang, and M. Johnson, "Predicting Crowd Dynamics Using Machine Learning in Urban Planning," *Journal of Urban Technology*, vol. 34, no. 2, pp. 120-134, 2025.

[7] J. Smith, L. Zhang, and M. Johnson, "Understanding pedestrian behavior using machine learning," *Urban Mobility and Planning*, vol. 7, no. 1, pp. 45-58, 2025.

[8] A. Patel, R. Sharma, and S. Desai, "Smart city infrastructure: Optimizing pedestrian flow with machine learning," *International Journal of Urban Planning*, vol. 22, no. 3, pp. 102-116, 2025.