CS4632 Simulation Runs and Data Collection

Department of Computer Science
College of Computing and Software Engineering
Kennesaw State University

Damien Castro, Dom Evans, Eric Legostaev

Spring 2025

1 Introduction

1.1 Project Overview

The purpose of this project is to develop a pedestrian traffic simulation model for Kennesaw State University's (KSU) Marietta campus using SUMO (Simulation of Urban Mobility). The primary goal is to predict pedestrian movement patterns across campus by leveraging estimated building occupancy data. This simulation aims to improve understanding of pedestrian flow, identify potential congestion points, and inform decision-making regarding campus infrastructure and pedestrian safety.

This model is particularly relevant in a university setting where class schedules, campus facilities, and walkway accessibility influence movement patterns. By integrating estimated building occupancy levels—based on general knowledge of building usage and historical class schedules—the simulation will generate realistic pedestrian traffic flows throughout the day.

1.2 Purpose of Simulation Runs

The simulation runs will serve as a critical validation tool to assess the accuracy of the predicted pedestrian movements. Specifically, the simulation aims to:

- Evaluate pedestrian density patterns at different times of the day based on estimated building occupancy.
- Identify peak congestion areas on campus to support infrastructure planning.
- Analyze the impact of schedule-driven fluctuations on overall pedestrian movement.
- Test different scenarios, including low-traffic periods and peak transition times.

1.3 Key Questions to Address

1. **What is the real-world problem or scenario being simulated?** The model seeks to predict pedestrian movement on the KSU Marietta campus using estimated building occupancy data.
2. **What are the primary goals of the simulation runs?** To analyze and visualize pedestrian traffic flow, congestion points, and movement patterns across campus.
3. **How does this milestone fit into the larger project context?** This milestone establishes the framework for data collection and model validation. While actual pedestrian data is not yet available, this phase will define the testing scenarios and expected outcomes to refine the simulation in future iterations.

2 Simulation Outline

2.1 Simulation Description

This simulation models pedestrian movement across the Kennesaw State University (KSU) Marietta campus, focusing on predicting traffic patterns based on estimated building occupancy data. The simulation is built using SUMO (Simulation of Urban Mobility) and integrates estimated occupancy levels for various campus buildings to generate pedestrian flows.

The simulation operates by using predefined occupancy schedules to determine the number of pedestrians entering and exiting buildings at different times of the day. These pedestrians are then assigned movement paths based on logical routing assumptions, such as travel between academic buildings and parking areas. The model considers factors such as walkway connectivity, pedestrian speed variations, and potential congestion points.

Key components of the simulation include:

- Pedestrian Entities: Individuals are represented as agents who move through the simulated campus environment based on building occupancy-driven spawn conditions.
- Building Occupancy Data: Pedestrian spawn rates are determined by estimated occupancy values, reflecting expected movement patterns at different hours of the day.
- Pathfinding and Routing: Pedestrian routes are assigned based on the most efficient paths between buildings.
- Time-Based Movement Variability: Pedestrian flows change over the course of the day, reflecting peaks during class transitions and lower activity during off-hours.

This model is designed to replicate real-world behaviors by simulating the effects of class schedules, building capacities, and pedestrian movement constraints on campus walkways. Future iterations may refine the accuracy of these predictions by incorporating real-world validation data, such as manual pedestrian counts or, if possible, sensor-based tracking.

2.2 Execution Instructions

This section provides a step-by-step guide to executing the pedestrian simulation for Kennesaw State University (KSU) Marietta campus using SUMO. The instructions include setup details, required dependencies, and expected input/output formats.

2.2.1 Prerequisites

Before running the simulation, ensure the following are installed on your system:

- **Python 3.8+**
- **SUMO (Simulation of Urban Mobility) 1.15+**
- **PostgreSQL 12+**
- **Required Python Libraries**: Install using command "pip install"
  - traci
  - sumolib
  - psycopg2
- **SUMO Configuration Files**: Ensure the following SUMO-related files are in place:
  - KSUMariettaConfig.sumocfg (Main SUMO configuration file)
  - KSUMarietta.net.xml.gz (SUMO network representation of the campus)
  - osm.pedestrian.rou.xml (Pedestrian type initialization)
  - osm.view.xml (Visualization configuration for SUMO-GUI(if used))

2.2.2 Steps to Execute the Simulation

1. **Set Up the PostgreSQL Database**: Ensure PostgreSQL is running and set up the mobinav database. Ensure the database contains the required tables (building, occupancy_schedule).

2. **Verify Configuration Files**: Check that the config.py file contains the correct paths to SUMO configuration.

   - SUMO_CONFIG_FILE = "KSUMariettaConfig.sumocfg"
   - SIMULATION STEPS = 86400 # each step equals one second

3. **Run the Simulation**: Execute the main script main.py. This script initializes the SUMO simulation, loads pedestrian data, and runs the simulation for a full day (86400 steps). To run in headless mode (no GUI) ensure simulation.py *def run_simulation* specifies "sumo". To run with GUI specify "sumo-gui".

4. **Monitor Pedestrian Spawning**: Pedestrians are spawned dynamically based on occupancy data retrieved from:

   - get_buildings_list.py (Fetches building coordinates)
   - get_coordinates_occupancy.py (Fetches estimated occupancy schedules)

5. **Analyze Outputs**:

   - Simulation Logs:
     - summary.xml: Overall simulation summary
     - tripinfo.xml: Individual pedestrian trip details
     - simulation_stats.csv: Derived simulation statistics
       * Total Pedestrians
       * Mean Speed
       * Waiting Pedestrians
       * Average Route Time
       * Throughput
       * Max Pedestrian Density
       * Max Density Lane
       * Max Waiting Time
       * Total Distance Traveled
       * Average Distance

6. **Stopping the Simulation**: To stop the simulation early, use Ctrl + C in the terminal.

2.3 Assumptions and Limitations

2.3.1 Assumptions

This pedestrian simulation is built upon several assumptions to simplify the modeling process while maintaining realistic movement patterns. These assumptions include:

- **Building Occupancy Estimation**: The number of pedestrians entering and exiting buildings is determined based on estimated occupancy data rather than real-time data or verified data. The estimates assume general campus activity trends.

- **Uniform Walking Speeds**: Pedestrians are assigned speeds within a standard range (e.g., 0.7 m/s to 2.0 m/s) to reflect natural walking variability, but individual behavioral differences (e.g., running, stopping, crowding effects) are not explicitly modeled.

- **Direct Route Selection**: Pedestrians take the shortest, fastest, or most efficient path between their origin and destination. The model does not account for personal route preferences, detours, or environmental factors such as weather conditions affecting movement.

- **Constant Campus Layout**: The simulation assumes that pathways, sidewalks, and building locations remain static, without accounting for temporary closures due to construction or events.
- **Instantaneous Entry and Exit**: Pedestrians instantly appear at their origin and disappear at their destination, without modeling behaviors such as waiting at entrances or spending time in transition areas.
- **Continuous Flow**: The model does not simulate bottlenecks at doorways, crosswalks, or intersections beyond simple congestion calculations.

2.3.2 Limitations

While the simulation aims to provide realistic pedestrian movement predictions, certain limitations affect the accuracy and scope of the results:

- **Lack of Real-World Validation Data**: Since actual pedestrian traffic data is not available, there is no direct validation against observed movement patterns on campus. The accuracy of the model depends on the reliability of occupancy estimates.
- **Simplified Pedestrian Behavior**: The model does not simulate behaviors such as group walking, sudden stops, or interactions between pedestrians, which can influence real-world movement patterns.
- **No Road Interaction Modeling**: The simulation focuses exclusively on pedestrian movement and does not account for interactions with vehicular traffic or other mobility constraints.
- **Limited Impact of Crowding**: While density is recorded, the simulation does not dynamically alter pedestrian speeds or path choices based on congestion levels.
- **Assumed Constant Weather and Environment**: Weather conditions, lighting, and other environmental factors that could impact pedestrian movement are not considered.

Despite these limitations, the simulation provides a valuable framework for analyzing general pedestrian traffic trends on the KSU Marietta campus. Future iterations can refine accuracy by incorporating real-world validation data, adjusting behavioral models, and integrating additional environmental factors.

3 Simulation Run Summary

3.1 Overview of Runs

The simulation runs are designed to analyze pedestrian movement across the KSU Marietta campus under different conditions. By varying occupancy levels, time periods, and movement constraints, these runs provide insights into congestion, flow efficiency, and potential improvements to campus mobility.

The following scenarios will be tested:

- Baseline Scenario – Standard pedestrian movement based on estimated building occupancy.
- Peak Hour Congestion Test – Evaluates walkway congestion during peak class transition times.
- Low-Traffic Period Analysis – Simulates nighttime and early morning pedestrian activity.
- Class Schedule Variations – Compares pedestrian flow across MWF, TTh, and weekend schedules (where MWF is Monday, Wednesday, Friday and TTh is Tuesday and Thursday).

Each scenario will assess pedestrian density, movement patterns, and congestion hotspots.

3.2 Simulation Run Table

The following table summarizes the results from each simulation scenario. Key metrics such as total pedestrian count, mean speed, average route time, and congestion effects were analyzed to assess pedestrian mobility.

Run ID	Objective	Parameters/Configurations	Expected Outcome	Actual Outcome
Run 1	Baseline Scenario	Standard estimated occupancy for a weekday.	Normal pedestrian movement patterns with peaks at class transitions.	Pedestrians appeared to move normally with the expected peaks at class transitions and no movement between classes.
Run 2	Low-Traffic Period	Estimated occupancy based overnight hours (10PM to 6AM).	Greatly reduced pedestrian movement patterns during overnight hours.	No pedestrian slow downs. No congestion.
Run 3	MWF Scenario	Estimated occupancy with pedestrian shifts at standard MWF class changes.	High pedestrian density following class end at the 50 minute mark of the hour during the typical college day.	High density shortly following end of classes with densities tapering off within 10-minutes. Low densities as evening hours are entered.
Run 4	TTh Scenario	Estimated occupancy with pedestrian shifts at standard TTh class changes.	High pedestrian density following class end after each 75-minute class starting at 8AM.	High density shortly following end of classes with densities tapering off within 10-minutes. Low densities as evening hours are entered.
Run 5	SSu (Weekend) Scenario	Estimated occupancy with reduced pedestrians due to less students on campus.	Normal pedestrian movement patterns without any significant congestion or pedestrian slowdowns.	Low densities. No or minimal pedestrian slowdowns.
Run 6	Peak MWF Scenario	Estimated occupancy with pedestrian shifts at standard MWF class changes. Adds additional pedestrians at peaking hours of 8AM to 9AM and 12PM to 1PM.	High pedestrian density following class end at the 50 minute mark of the hour during the typical college day. Additional high density and more extreme slowdowns during peaking hours.	High density shortly following end of classes with densities tapering off within 10-minutes. Highest densities with high pedestrian wait times in the peaking hours.
Run 7	Peak TTh Scenario	Estimated occupancy with pedestrian shifts at standard TTh class changes. Adds additional pedestrians at peaking hours of 8AM to 9AM and 12PM to 1PM.	High pedestrian density following class end after each 75-minute class starting at 8AM. Additional high density and more extreme slowdowns during peaking hours.	High density shortly following end of classes with densities tapering off within 10-minutes. Highest densities with high pedestrian wait times in the peaking hours. Wait times approached 300 seconds.
Run 8	Peak SSu (Weekend) Scenario	Estimated occupancy with reduced pedestrians due to less students on campus. Adds additional pedestrians at peaking hours of 8AM to 9AM and 12PM to 1PM.	Normal pedestrian movement patterns without any significant congestion or pedestrian slowdowns. Increased density but no true slow downs due to density as pedestrian population is lower on weekends.	Low densities. Minimal pedestrian slowdowns aroudn peaking hours.

Figure 1: Simulation Run Table

3.3 Key Insights

1. Peak Pedestrian Density Exceeds Comfortable Walkway Limits
   - The highest pedestrian congestion occurred in the Peak MWF and Peak TTh scenarios.
   - Maximum observed density exceeded 50 pedestrians per lane, indicating potential infrastructure limitations.
   - The longest waiting times reached 300 seconds, suggesting major bottlenecks in high-traffic corridors.
2. Pedestrian Speeds Drop Significantly in Congested Conditions
   - Peak-hour speeds averaged 0.19 m/s, compared to 0.29 m/s in lower-density scenarios.
   - Average route times increased in high-density scenarios, confirming that walkway congestion affects movement efficiency.
3. MWF vs. TTh Schedule Differences
   - MWF schedules distributed pedestrian flow more evenly throughout the day, leading to moderate but sustained congestion.
   - TTh schedules had fewer transitions but sharper peaks, resulting in higher localized congestion at specific times.
4. Low-Traffic Periods Show Minimal Movement
   - Pedestrian counts dropped significantly in the late-night and early-morning simulations, confirming reduced foot traffic outside of class hours.

4 Data Collection and Storage

4.1 Data Collection Methodology

The simulation collects data on pedestrian movement across the KSU Marietta campus to analyze traffic patterns and identify congestion points. The data is gathered through SUMO's TraCI interface and recorded at regular intervals to capture dynamic pedestrian behavior.

The key data points collected include:

- Total Pedestrians – The number of pedestrians present at each time step.
- Mean Speed – The average walking speed of all pedestrians.
- Average Route Time – The mean duration required for pedestrians to complete their routes.
- Throughput – The number of pedestrians successfully completing their routes.
- Max Pedestrian Density – The highest density recorded on any walkway segment.
- Max Density Lane – The specific walkway segment experiencing the highest pedestrian density.
- Max Waiting Time – The longest time a pedestrian remains stationary.
- Average Distance – The mean distance traveled per pedestrian.
- Total Distance Traveled – The cumulative distance covered by all pedestrians.

Data Collection Process

- Pedestrian movement data is logged every 60 time steps (1 minute) to reduce unnecessary storage overhead and balance granularity and storage efficiency.
- SUMO's built-in functions (TraCI API) retrieve pedestrian IDs, locations, and speeds as the simulation runs.
- PostgreSQL database integration stores building and occupancy data for later use.

4.2 Data Formats and Storage

4.2.1 Data Formats

Collected data is stored in CSV files and XML logs, with an option to store key metrics in a PostgreSQL database. The formats include:

- Pedestrian Statistics (CSV)
  - File: simulation_stats.csv
  - Columns: Time, Total Pedestrians, Mean Speed, Average Route Time, Throughput, Max Pedestrian Density, Max Density Lane, Max Waiting Time, Total Distance Traveled, Average Distance
  - Logged every 60 steps (1 minute).
- Simulation Summary (XML)
  - File: summary.xml
  - Contains high-level statistics such as total pedestrian count and movement trends.
- Trip Information (XML)
  - File: tripinfo.xml
  - Logs detailed pedestrian routes, departure times, and travel durations.
- PostgreSQL Database
  - Table: building
    * Stores building information, including building ID, name, latitude, and longitude.
    * Used to determine pedestrian origins and destinations.
  - Table: occupancy_schedule
    * Stores estimated occupancy schedules per building.
    * Columns: building ID, day, start time, end time, occupancy value.
    * Used to spawn pedestrians based on predicted building usage patterns.

4.2.2 Storage and Access

- CSV files are stored locally in the working directory.
- XML logs are generated automatically by SUMO for debugging and in-depth analysis.
- Database storage ensures structured access for querying occupancy-based pedestrian flow.

5 Data Analysis

5.1 Analysis Approach

The collected simulation data is analyzed to evaluate pedestrian traffic patterns on the KSU Marietta campus. The analysis focuses on key performance indicators that measure pedestrian density, movement efficiency, and congestion levels.

5.1.1 Metrics Used for Analysis

The following metrics are extracted from the simulation data to assess pedestrian behavior:

- Total Pedestrians – The number of pedestrians present at each time step, indicating fluctuations in pedestrian traffic throughout the day.
- Mean Speed – The average walking speed of all pedestrians, used to assess movement efficiency and congestion effects.
- Average Route Time – The mean duration pedestrians take to complete their trips, helping identify delays.
- Throughput – The number of pedestrians who successfully complete their routes, measuring efficiency.
- Max Pedestrian Density – The highest density recorded on any walkway segment, used to detect congestion hotspots.
- Max Density Lane – The specific walkway experiencing the highest pedestrian density.
- Max Waiting Time – The longest time a pedestrian remains stationary, signaling potential bottlenecks.
- Total Distance Traveled – The cumulative distance covered by all pedestrians, helping analyze movement across campus.
- Average Distance – The mean distance traveled per pedestrian, providing insight into typical trip lengths.

5.1.2 Methodology

The analysis process follows these steps:

1. Data Aggregation:
   - Simulation data is logged every minute to track pedestrian trends over time.
   - The occupancy schedule is used to correlate pedestrian presence with estimated building populations.
2. Trend Analysis:
   - Peak hours are identified by analyzing spikes in pedestrian counts.
   - Congestion zones are detected using maximum pedestrian density values.
3. Comparative Analysis:
   - Results are compared between different time intervals (e.g., low-occupancy interval and common college/university hours traffic).
   - Walkway congestion is evaluated based on density variations.

5.1.3 Tools Used

- SUMO TraCI API – Extraction of real-time pedestrian movement statistics.
- Pandas - Data processing, extraction, and aggregation before generating visualizations.
- Matplotlib – Visualization.

The analysis aims to provide insights into pedestrian congestion patterns and validate the simulation model's effectiveness in estimating real-world pedestrian traffic.

5.2 Results

The simulation runs produced detailed statistics on pedestrian movement patterns across different scenarios, including baseline, low-traffic periods, peak hours, and schedule-based variations. The results highlight key trends in pedestrian density, route efficiency, and congestion points.

5.2.1 Summary of Key Findings

Total Pedestrian Counts and Flow Patterns

- Baseline Scenario:
  - Average pedestrian count: 85 per time step
  - Peak pedestrian count: 2,190 at busiest period
- Low-Traffic Periods (LTP):
  - Minimal pedestrian movement: 0.2 pedestrians per step
  - Peak pedestrian count: 35 pedestrians
- MWF Schedule:
  - Average pedestrian count: 185 per step
  - Peak pedestrian count: 3,727 at highest class transition
- TTh Schedule:
  - Average pedestrian count: 109 per step
  - Peak pedestrian count: 2,844 pedestrians
- Weekend (SSu):
  - Average pedestrian count: 72 per step
  - Peak pedestrian count: 887 pedestrians
- Peak Hour Scenario:
  - Peak pedestrian count:
    * MWF: 3,727 pedestrians
    * TTh: 2,844 pedestrians
    * SSu: 887 pedestrians

Speed and Route Efficiency

- Mean Speed Variations:
  - Baseline: 0.23 m/s
  - Peak MWF: 0.25 m/s
  - Peak SSu: 0.29 m/s
  - Peak TTh: 0.19 m/s (slowest due to congestion effects)
- Average Route Time:

- Baseline: 362 seconds
- Peak MWF: 419 seconds
- Peak SSu: 428 seconds
- Peak TTh: 435 seconds

- Throughput (Completed Trips):
  - Baseline: 11 pedestrians per step
  - Peak MWF: 22 pedestrians per step
  - Peak TTh: 13 pedestrians per step
  - Peak SSu: 9 pedestrians per step

Congestion and Waiting Time

- Max Pedestrian Density:
  - Baseline: 39 pedestrians per lane
  - Peak MWF: 51 pedestrians per lane
  - Peak TTh: 58 pedestrians per lane (highest recorded congestion)
  - Peak SSu: 16 pedestrians per lane
- Max Waiting Time:
  - Baseline: 240 seconds
  - Peak MWF: 300 seconds
  - Peak TTh: 300 seconds
  - Peak SSu: 120 seconds

Distance Traveled

- Total Distance Traveled:
  - Baseline: 1.9 million meters
  - Peak MWF: 3.8 million meters
  - Peak TTh: 2.2 million meters
  - Peak SSu: 1.9 million meters
- Average Distance Per Pedestrian:
  - Baseline: 249 meters per pedestrian
  - Peak MWF: 281 meters per pedestrian
  - Peak TTh: 292 meters per pedestrian
  - Peak SSu: 325 meters per pedestrian

5.2.2 Summary Visualizations

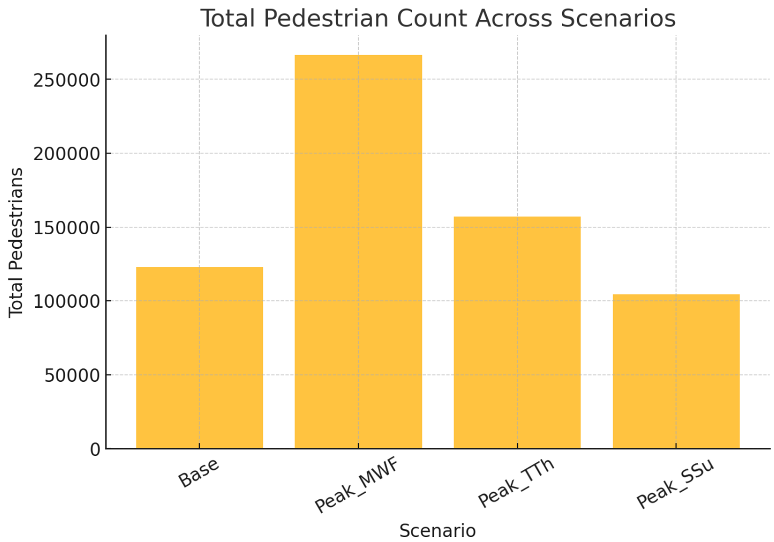

Figure 2: Total Pedestrian Count Across Scenarios

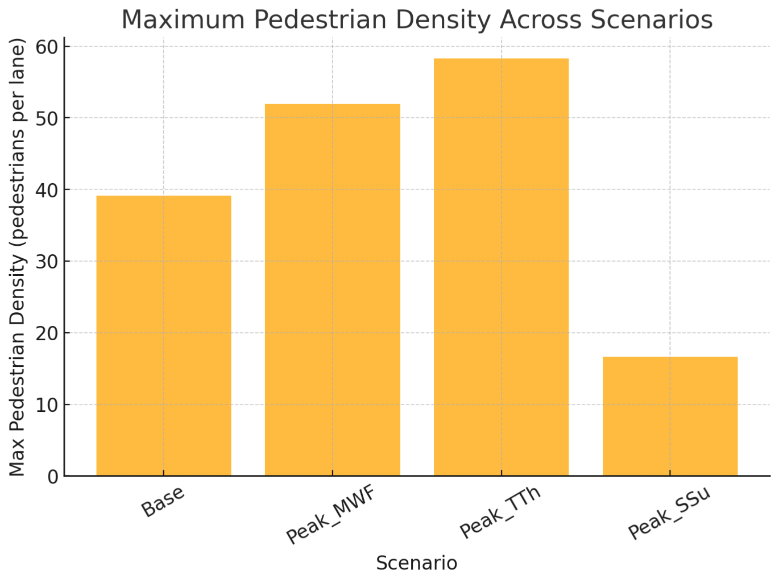

Figure 3: Maximum Pedestrian Density Across Scenarios

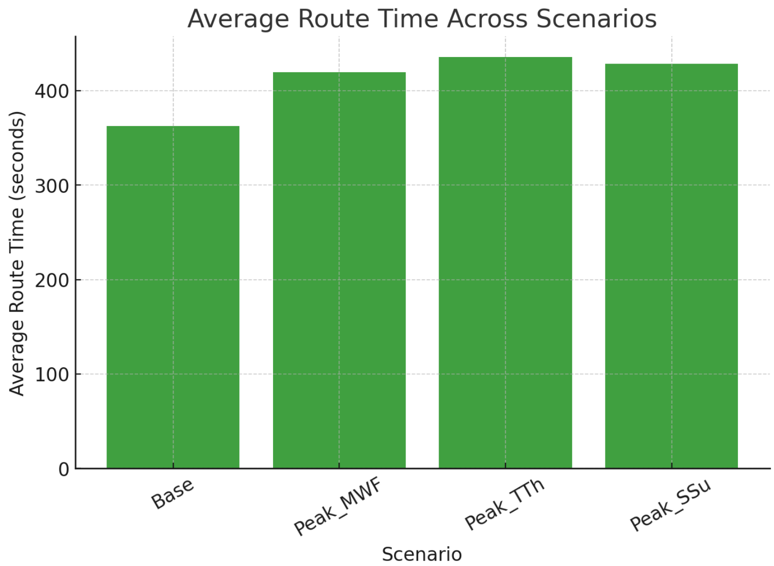

Figure 4: Average Route Time Across Scenarios

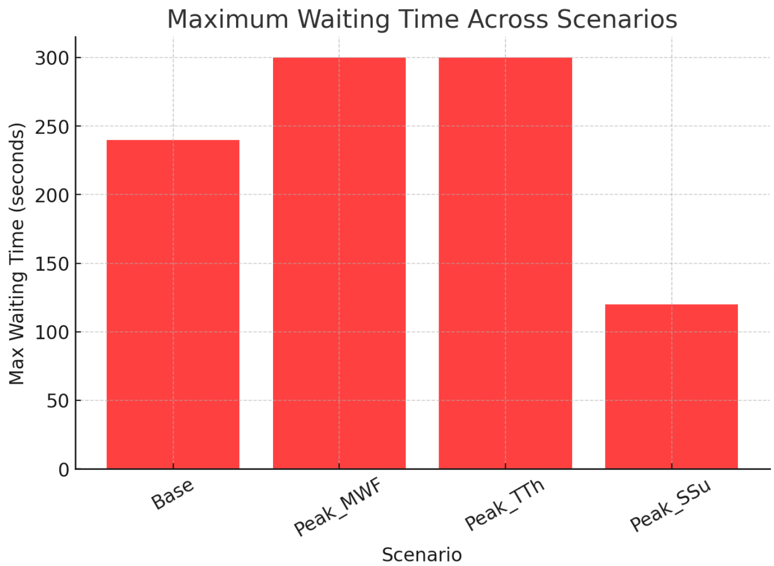

Figure 5: Maximum Waiting Time Across Scenarios

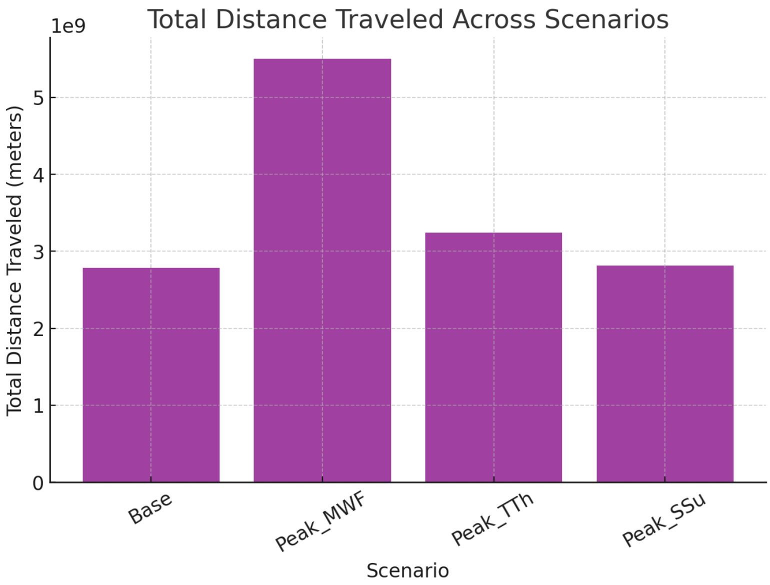

Figure 6: Total Distance Traveled Across Scenarios

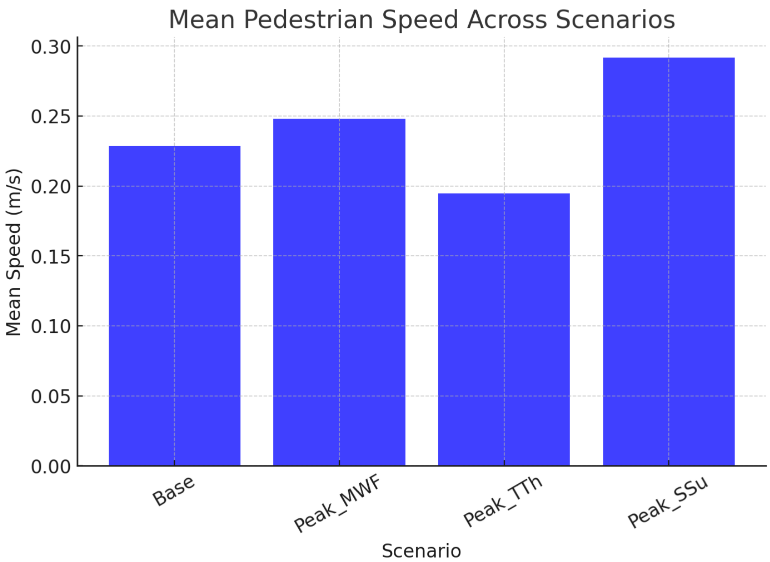

Figure 7: Mean Pedestrian Speed Across Scenarios

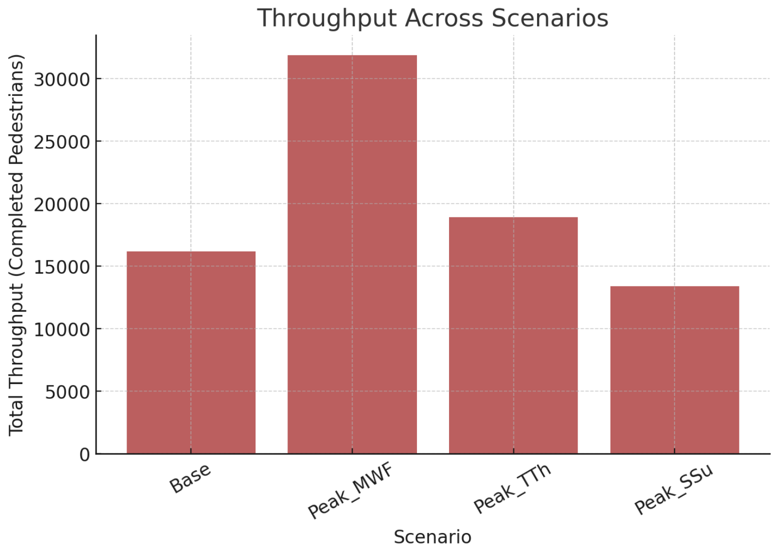

Figure 8: Throughput Across Scenarios

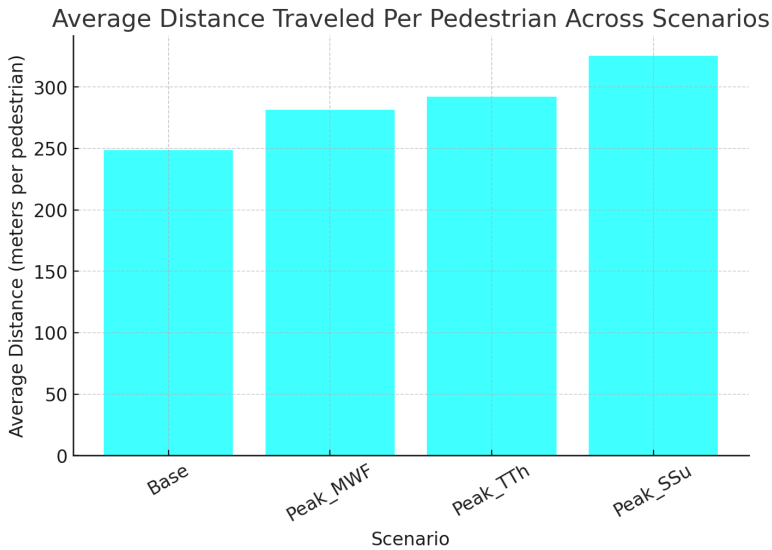

Figure 9: Average Distance Traveled Per Pedestrian Across Scenarios

5.3 Discussion

The pedestrian traffic simulation results provide key insights into campus mobility patterns, congestion areas, and differences across various scenarios. This discussion interprets the findings, assesses their implications, and identifies areas for future improvements.

5.3.1 Key Observations

1. Pedestrian Flow and Campus Activity Trends
    - Peak congestion is highest on MWF schedules due to the larger number of overlapping class periods.
    - Weekend activity remains significantly lower, but the peak of 887 pedestrians suggests recreational or dorm-based movements.
2. Impact of Peak Hours and Congestion Patterns

- The TTh schedule experiences slower pedestrian movement speeds due to higher density in key corridors.
- Long waiting times (up to 5 minutes) remain an issue in high-density areas, indicating bottleneck locations.
- The longest waiting times (300 seconds) were recorded in multiple scenarios, indicating potential bottlenecks at key intersections.

3. Route Efficiency and Movement Constraints

- Mean pedestrian speed is lower in high-density scenarios and was lowest during peak hours (0.19 m/s in TTh), reflecting congestion-related delays.
- The total distance traveled has increased significantly in peak hour scenarios, showing higher pedestrian volume across campus.
- The average distance per pedestrian varied across scenarios, with weekend pedestrians traveling longer distances, likely due to fewer people making short intra-campus trips.

5.3.2 Interpretation of Findings

1. Increased Congestion in High-Activity Periods:

- The simulations confirm that MWF schedules experience the highest pedestrian congestion, reaching a peak of 3,727 pedestrians in a single time step.
- TTh schedules exhibit fewer but more intense transitions, with congestion peaking at 2,844 pedestrians at class changeovers.
- Weekend traffic remains low overall, but the peak of 887 pedestrians suggests a significant presence of residential or event-based foot traffic on campus.

2. Impact of Density on Movement Efficiency:

- Pedestrian movement speeds decrease significantly in high-density environments.
  - The Peak TTh scenario recorded the lowest mean pedestrian speed (0.19 m/s), indicating significant congestion effects.
  - The Peak MWF scenario had a slightly higher mean speed (0.25 m/s), though congestion remains an issue.
- Longer average route times in peak scenarios confirm delays due to congestion.
  - Peak MWF: 419 seconds per pedestrian
  - Peak TTh: 435 seconds per pedestrian

3. Bottlenecks and Walkway Capacity Issues:

- The simulation highlights specific walkway segments where pedestrian density exceeds comfortable levels.
- Peak MWF and TTh scenarios recorded pedestrian densities exceeding 50 pedestrians per lane, far above the congestion threshold.
- Maximum waiting times in these scenarios reached 5 minutes, suggesting inefficiencies in walkway distribution or pedestrian routing behavior.

4. Increased Total Distance Traveled in Peak Scenarios:

- Pedestrians in peak hour conditions traveled significantly greater distances compared to the baseline scenario.
  - Peak MWF Total Distance: 3.8 million meters
  - Peak TTh Total Distance: 2.2 million meters
- This increase in total travel distance suggests that more pedestrians are using campus pathways simultaneously, reinforcing the observed congestion effects.

5. Validation of Simulation Against Expected Trends:
   The correlation between peak pedestrian traffic and expected schedule changes validates the simulation model's ability to predict movement trends accurately.

The findings highlight several key implications for infrastructure and mobility management on campus:

- Need for Walkway Expansion in High-Density Areas: The data suggests that certain campus pathways experience extreme congestion, potentially requiring wider walkways or alternate routes to manage peak-hour pedestrian flow.
- Strategic Adjustments to Class Scheduling: The high density recorded during MWF transitions suggests that minor schedule adjustments (e.g., staggered class start times) could alleviate peak-hour congestion.
- Improved Routing Strategies for Pedestrians: The introduction of adaptive routing algorithms could help reduce pedestrian buildup in congested areas by dynamically directing foot traffic toward lower-density pathways.

5.3.3 Potential Model Improvements and Future Considerations

1. Refining Congestion and Walkway Interactions
   - Incorporating pedestrian interactions (e.g., groups, overtaking) would improve congestion realism.
   - Adding constraints on walkway capacity could simulate bottlenecks more effectively.
2. Incorporating Dynamic Routing Strategies
   - Implementing adaptive pedestrian routing based on real-time congestion levels may provide more realistic movement behaviors.
3. Accounting for Environmental Factors
   - Weather-based route selection (covered walkways, avoiding open areas in rain) could be tested.
   - Lighting conditions at night may impact movement speeds.
4. Validation with Real-World Data
   - Manual pedestrian counts or sensor-based tracking would help refine the model.
   - Comparing SUMO outputs to observed traffic flow could identify calibration adjustments.

6 Conclusion

This project developed a pedestrian traffic simulation for the Kennesaw State University (KSU) Marietta campus using SUMO, with the objective of predicting pedestrian flow based on estimated building occupancy. The simulation runs analyzed different scenarios, including baseline conditions, peak hours, schedule-based variations, and low-traffic periods.

6.1 Summary of Findings

The simulation results provided valuable insights into pedestrian density trends, congestion points, and movement efficiency across campus:

- High-Density Walkways and Congestion Effects:
  - Peak hour congestion significantly impacts movement speed, with pedestrian density exceeding 50 pedestrians per lane in key locations.
  - The lowest pedestrian speeds (0.19 m/s) were recorded during Peak TTh scenarios, indicating severe congestion effects.
- Weekend and Off-Peak Activity Remains Lower:

  - SSu schedules experience significantly lower foot traffic but still show peak occupancy of 887 pedestrians, likely due to dormitory residents and event-based movement.

- Travel Distance and Route Efficiency Adjustments:
  - Pedestrians in high-density conditions travel farther overall, with Peak MWF pedestrians covering 3.8 million meters in total compared to 1.9 million meters in the baseline scenario.
  - Longer route times in congested scenarios (419+ seconds per pedestrian) indicate movement inefficiencies due to walkway crowding.

6.2 Implications for the Project

The simulation models pedestrian movement trends on the KSU Marietta campus, highlighting congestion areas, peak movement times, and differences across schedules. These findings suggest several key implications:

- **Campus Infrastructure Planning**: The high-density pedestrian zones indicate a need for walkway expansion, improved crosswalk timing, or alternate pedestrian paths to alleviate congestion.
- **Class Scheduling Adjustments**: The data suggests that peak congestion occurs during specific class transition periods (e.g., MWF mid-morning shifts). Staggering class start times slightly could reduce peak density and improve pedestrian flow efficiency.
- **Pedestrian Safety Considerations**: Identified high-waiting-time zones may require adjustments to traffic control measures or improved pedestrian crossings.
- **Adaptive Pedestrian Routing Strategies**: Congestion-aware pathfinding algorithms could improve pedestrian movement efficiency, reducing bottlenecks in key areas.

6.3 Next Steps

While the current model provides useful insights, further refinements can enhance accuracy and applicability. The following next steps are recommended:

- **Refining Congestion Modeling**: Incorporate pedestrian interactions, group movement behavior, and overtaking dynamics to improve congestion representation.
- **Implement Dynamic Routing to Address Congestion**: Current pedestrian routing is static, meaning that pedestrians do not adjust their paths based on real-time congestion levels. Future work should integrate adaptive routing logic that dynamically redirects foot traffic toward less congested pathways.
- **Exploring Environmental Factors**: Extend the model to consider weather effects, nighttime lighting conditions, and temporary walkway closures to assess their impact on movement patterns.
- **Validating with Real-World Data**: Compare simulation results with manual pedestrian counts or sensor-based tracking to ensure the model accurately reflects real pedestrian behavior.
- **Expanding Scenario Testing**: Future simulations can test emergency evacuations, accessibility-focused pedestrian routing, or policy-based interventions (e.g., restricted zones).

7 Documentation Updates

Throughout the simulation development process, several modifications were made to the model, tools, and simulation plan to enhance efficiency and accuracy. This section outlines key updates to the simulation model, tools used, and adjustments made to the simulation plan.

7.1 Model Changes

Several changes were implemented in the pedestrian simulation to improve data management, network connectivity, and computational efficiency:

- Transition to Gaussian and Agent-Based Occupancy Modeling
  - The original occupancy data used static estimates, which did not capture fluctuations in pedestrian traffic.
  - Gaussian-based estimation now accounts for natural variations in occupancy levels throughout the day, making movement patterns more realistic.
  - Agent-based occupancy models further refine pedestrian spawn patterns by simulating individual decision-making processes.
- Data Collection Frequency Reduction:
  - Initially, the simulation recorded data at each simulation step, causing the simulation to run extremely slowly.
  - This was adjusted to record data once per minute, significantly reducing storage size while still capturing essential trends.
- Occupancy Data Integration with PostgreSQL:
  - Originally, pedestrians were manually hardcoded in SUMO's XML files, requiring extensive file modifications for each scenario.
  - This was replaced with a PostgreSQL database, allowing pedestrian generation to be dynamically managed based on retrieval of building occupancy schedules.
  - The database-driven approach improved flexibility and scalability, reducing the need for manual XML file adjustments.
- SUMO Network File Corrections:
  - The initial SUMO network file contained numerous unconnected edges and junctions, making pedestrian routing unreliable and causing simulation failures.
  - Significant manual corrections were required to connect broken edges, adjust network topology, and ensure proper pedestrian pathways.
  - This process was time-consuming but essential for ensuring that pedestrians could move realistically through the campus network without encountering erroneous dead ends or invalid routes.
- Optimizing Pedestrian Spawning:
  - Originally, pedestrians were spawned one at a time, which was slow and inefficient.
  - This was changed to spawning pedestrians in batches, significantly reducing processing overhead and improving simulation performance.
  - The batch spawning approach allowed larger groups of pedestrians to be generated at once, making high-density scenarios run much faster.
- Enhanced Pedestrian Routing and Pathfinding:
  - Adjustments were made to SUMO's pedestrian routing logic to better reflect realistic movement patterns across campus walkways.
  - Walkway congestion data was used to refine route assignment, ensuring that pedestrian movements more closely resembled real-world conditions.

7.2 Tools and Libraries

New tools and libraries were integrated to improve data processing and simulation performance. The following tools were used:

- SUMO (Simulation of Urban Mobility) 1.15+ – Used to simulate pedestrian movement across the campus.
- PostgreSQL 12+ – Used to manage building occupancy data dynamically.
- TraCI (SUMO API, built-in) – Allowed real-time control of pedestrian entities within the simulation.
- psycopg2 (PostgreSQL Connector, latest version) – Used to query occupancy data from the database and feed it into the simulation.

These tools allowed for more efficient data collection, flexible pedestrian generation, and improved integration of real-time building occupancy data into the simulation.

7.3 Simulation Plan Table Updates

Added scenario-based testing (MWF, TTh, Peak, etc.) to provide deeper insights into how pedestrian traffic fluctuates based on different class schedules and activity levels.

8 Reflection and Next Steps

8.1 Challenges Faced

During the development and execution of the pedestrian traffic simulation, several challenges were encountered, requiring iterative problem-solving and model adjustments:

- SUMO Network File Issues
  - The original SUMO network contained numerous unconnected edges and junctions, preventing smooth pedestrian movement.
  - Fixing these connectivity issues was time-consuming, requiring extensive manual corrections to ensure valid paths for pedestrian travel.
- Computational Efficiency and Performance Bottlenecks
  - Initially, pedestrians were spawned one at a time, significantly slowing down simulation execution due to repeated calls for routing calculations.
  - Data collection was originally performed at every simulation step, leading to excessive data output and increased processing time.
- Complexity of Pedestrian Data Management
  - Managing pedestrian movement through hardcoded XML files was inefficient and required frequent manual updates.
  - Migrating to a PostgreSQL-based approach introduced database integration challenges, requiring adjustments to pedestrian query and spawn mechanisms.

8.2 Solutions Implemented

To overcome these challenges, several key solutions were introduced:

- Fixing the SUMO Network
  - Manually corrected broken edges and junctions, ensuring pedestrians could navigate the campus without encountering routing failures.
  - Verified network integrity through test runs before executing full-scale simulations.

- Optimizing Pedestrian Spawning
  - Shifted from single pedestrian spawning to batch spawning, significantly improving simulation speed and reducing processing overhead.
- Enhancing Data Collection Efficiency
  - Reduced data collection frequency from every simulation step to once per minute, leading to more manageable datasets while preserving meaningful trends.
- Transitioning to a Dynamic Database-Driven Model
  - Migrated occupancy data to PostgreSQL, allowing real-time adjustments to pedestrian spawns based on estimated building occupancy.
  - Eliminated the need for static, hardcoded SUMO XML files, making scenario modifications easier.

8.3 Future Plans

While significant progress has been made in optimizing the pedestrian traffic simulation, additional improvements could further enhance accuracy and usability:

- Refining Pedestrian Routing and Pathfinding
  - Current pedestrian movement relies on standard SUMO routing, which does not fully account for real-world navigation behaviors.
  - Future work could implement improved congestion-based routing, allowing pedestrians to dynamically adjust paths based on walkway density.
- Validating the Simulation with Real-World Data
  - The model currently relies on estimated occupancy schedules to predict pedestrian traffic. • Future iterations should incorporate real-world pedestrian count data (e.g., from campus surveillance, sensors, or manual counts) to validate and refine the model.
- Exploring Additional Environmental Factors
  - Future work could include weather conditions, lighting effects, and temporary path closures to assess their impact on pedestrian movement.
- Expanding Scenario Testing
  - Additional simulations could explore emergency evacuation modeling, accessibility-focused routing, or policy-driven pedestrian flow adjustments (e.g., restricted areas).

8.3.1 Final Thoughts

This milestone significantly improved the efficiency and accuracy of pedestrian traffic modeling for the KSU Marietta campus. Despite the challenges, the implemented solutions have reduced processing time, improved data management, and allowed for flexible scenario adjustments. Looking forward, future iterations should focus on validating results, refining pedestrian movement behaviors, and expanding simulation scenarios to further enhance the model's applicability in campus infrastructure planning.

9 References

9.1 Sources and Citations

The following sources were referenced for methodologies related to building occupancy estimation:

- Building Occupancy Estimation
  - **Standard Practices for Occupancy Modeling**: Gaussian distribution models applied to campus buildings to simulate occupancy diversity across different functional spaces (Sun et al., 2020).

– **Agent-Based Modeling**: Reduced-order graphical models of occupant behavior for structured estimation of room-level occupancy (Barooah & Langarita, 2010).

These sources provide a foundation for occupancy estimation methodologies, aligning with established practices in situations where direct measurement tools are unavailable.

9.2 Tools and Libraries

The following tools and libraries were used for simulation execution, data collection, and database integration:

- SUMO (Simulation of Urban Mobility) 1.15+ – Used for pedestrian movement simulation (https://www.eclipse.dev/s
- PostgreSQL 12+ – Database management for building occupancy data (https://www.postgresql.org/).
- TraCI (SUMO API, built-in) – Interface for real-time control and data extraction from the SUMO simulation (https://sumo.dlr.de/docs/TraCI.html).
- Python – Primary programming language used for scripting and data processing (https://www.python.org/).
- psycopg2 – PostgreSQL adapter for Python, used to query occupancy data and integrate it into the simulation (https://www.psycopg.org/).

9.3 Additional Resources

In addition to primary sources and tools, the following resources were consulted during the simulation development:

- KSU Marietta Campus Map – Used as a reference for walkway structures.
- SUMO User Documentation – Official SUMO documentation for pedestrian modeling and simulation setup (https://sumo.dlr.de/docs/).
- Eclipse SUMO Wiki – Community-driven resources for troubleshooting and advanced simulation modeling.