

Proyecto de Simulación e Inteligencia Artificial para Optimización del Transporte en La Habana: Maximizando la Eficiencia de los Ómnibus y Gazelles frente a Desafíos Financieros y Operativos

Sección 1

El transporte público en La Habana enfrenta desafíos críticos que afectan su eficiencia y capacidad para satisfacer las necesidades de la población y la economía en general. Estos desafíos incluyen la escasez de recursos financieros, la disminución de la capacidad de transporte y la inestabilidad en la oferta de servicios. En este contexto, se propone un proyecto de simulación integrado con inteligencia artificial para optimizar la distribución de los vehículos existentes, específicamente los ómnibus y gazelle, en las distintas rutas de la ciudad.

La inteligencia artificial (IA) y la simulación están revolucionando la optimización del transporte público. La IA posibilita el análisis en tiempo real de datos, permitiendo a las empresas monitorear la demanda y ajustar operaciones para mejorar la eficiencia y satisfacción de los pasajeros. Combinada con la simulación, se optimizan rutas existentes y se evalúan escenarios hipotéticos, minimizando riesgos y maximizando la planificación a largo plazo. Además, la IA ayuda a prever retrasos y gestionar eventos inesperados, mejorando la experiencia del usuario.

Sin embargo, a pesar de los avances significativos, existen lagunas y desafíos que deben abordarse para aprovechar plenamente el potencial de la IA y la simulación en la optimización del transporte público. En particular, hay una falta de estudios específicos que aborden las necesidades y desafíos únicos del transporte público en La Habana. Además, los desafíos financieros y operativos pueden limitar la aplicabilidad de las soluciones existentes, lo que destaca la necesidad de considerar adecuadamente estos factores en el desarrollo e implementación de nuevas soluciones.

Objetivos del Proyecto

1. Desarrollar un modelo de simulación del transporte en La Habana que integre paradigmas de inteligencia artificial para analizar la distribución actual de guaguas y gazelle en las distintas rutas.
2. Identificar áreas de mejora en la distribución de vehículos para maximizar la eficiencia y rendimiento del sistema de transporte público.
3. Proporcionar recomendaciones específicas sobre cómo asignar de manera óptima los vehículos disponibles para satisfacer las necesidades de transporte de la población.

Limitaciones

El proyecto depende de la calidad y disponibilidad de datos sobre el sistema de transporte. Hay retos relacionados con la recopilación de datos precisos y completos.

Las soluciones propuestas deben ser factibles con los recursos financieros y operativos actuales en La Habana. Esto puede restringir las opciones de optimización.

Una de las limitaciones importantes de esta investigación radica en los recursos limitados, especialmente en términos de la habilidad y experiencia de los investigadores. Aunque se ha trabajado para desarrollar las metodologías propuestas, las capacidades técnicas y de análisis de los investigadores pueden afectar la profundidad y amplitud del estudio.

Estructura

En la primera sección se presentó una breve introducción al problema y los objetivos del proyecto. Además, se plantearon el alcance y las limitaciones del proyecto para establecer el contexto en el que se desarrollará la investigación.

En la segunda sección, se define el caso de estudio y el modelo seguido para realizar la simulación. Se describen las herramientas y técnicas utilizadas, así como los datos recopilados para el análisis.

En la tercera sección, se explican los paradigmas de inteligencia artificial utilizados y cómo se aplican para optimizar el transporte público. Se detallan los algoritmos y enfoques empleados para analizar los datos y simular escenarios de mejora.

Finalmente, se presentan las conclusiones y recomendaciones basadas en los resultados de la simulación.

Sección 2

Caso de estudio

Nuestro caso de estudio es la ciudad de la Habana en Cuba. Esta ciudad posee una población de alrededor de 2 137 847 habitantes y una densidad poblacional de aproximadamente 2924.2 habitantes por km^2 .

División por zonas

Se utilizará la división de zonas por municipios como método de análisis, fundamentándose en la necesidad de emplear áreas administrativas oficiales. Esta aproximación facilitará la asociación de cada zona con datos estadísticos disponibles, tales como población, empleo, entre otros.

Red básica

Se empleará un grafo dirigido en el que cada nodo representa una cuadra, definida como un segmento de calle entre dos intersecciones. Durante la construcción del grafo, se determina según las direcciones de las calles y las restricciones de giro las aristas del grafo. Para cualquier par de nodos (cuadras) x y y , existe una arista desde x hasta y si hay una intersección que los conecta y es posible avanzar desde x hasta y , ya sea doblando a la derecha, a la izquierda, siguiendo recto o realizando un giro en U.

Cada nodo también contiene información sobre la longitud de la cuadra y elementos que pueden afectar la velocidad de desplazamiento, como cruces peatonales, semáforos o señales de Pare y Ceda el paso. Además, proporciona información sobre las rutas que tienen puntos de parada en esta cuadra y posibles puntos de interés.

Dimensiones temporales

El análisis se desarrollará a lo largo de un período temporal de al menos un año, dividido en trimestres para facilitar la comprensión de las variaciones estacionales en la demanda de transporte. Se presupondrá que a largo plazo, la población de La Habana y sus patrones de movilidad experimentarán cambios graduales debido al crecimiento demográfico, desarrollos urbanos y modificaciones en la economía y el estilo de vida de la población.

Se tomará el año 2019 como referencia para establecer los patrones de demanda de transporte, debido a la disponibilidad de datos detallados para dicho año. Se considerará que los parámetros del sistema, como la capacidad de los vehículos de transporte público, la disponibilidad de servicios y la infraestructura vial, pueden experimentar ligeras variaciones dentro de cada período de referencia, pero se mantendrán dentro de rangos aceptables para la simulación.

Los promedios de pasajeros transportados en cada trimestre del año son los siguientes, en orden cronológico: 1 029 100, 1 069 100, 932 500 y 901 400. Se destaca que los meses de septiembre y octubre, marcados por la situación conocida en Cuba como "la conyuntura", experimentaron una severa afectación en el transporte del país debido a la situación del petróleo. Estos datos se considerarán relevantes, dado que este tipo de situaciones no son poco comunes en la realidad actual del país.

En las urbes de gran tamaño, las horas pico suelen concentrarse durante las horas de la mañana y la tarde, coincidiendo con los horarios de entrada y salida del trabajo y la escuela. Dado que no se dispone de información específica sobre las horas pico en la capital del país, se hará una suposición general considerando las horas pico como aquellas comprendidas entre las 7:00 a.m. y las 10:00 a.m., y entre las 5:00 p.m. y las 8:00 p.m.

Estimación de la demanda

Para estimar la demanda de transporte público en La Habana, tomaremos en consideración los desplazamientos relacionados con el trabajo y la educación. Aunque no contamos con datos específicos sobre la población económicamente activa (PEA) por municipios, haremos una aproximación basada en los siguientes datos:

- La población total de La Habana con 15 años o más es de 1,827,259 personas.
- De estas, la población económicamente activa (PEA) es de 854,061 personas.
 - 836,564 de ellas están empleadas, mientras que 17,497 están desempleadas.

Esto significa que aproximadamente el 47% de la población de 15 años o más es económicamente activa, y el 98% de la PEA está empleada. Basándonos en estas estadísticas, la distribución de la PEA por municipio y su ocupación se resume en la siguiente tabla:

| Municipio | 15+y | PEA | Ocupados | Desocupados |
|---------------|---------|--------|----------|-------------|
| Playa | 156.937 | 73.352 | 71.848 | 1.504 |
| Plaza | 125.200 | 58.518 | 57.318 | 1.200 |
| Centro Habana | 115.228 | 53.858 | 52.754 | 1.104 |

| Municipio | 15+y | PEA | Ocupados | Desocupados |
|-----------------------|-----------|---------|----------|-------------|
| Habana Vieja | 68.911 | 32.209 | 31.549 | 660 |
| Regla | 37.345 | 17.455 | 17.097 | 358 |
| Habana del Este | 149.688 | 69.964 | 68.530 | 1.434 |
| Guanabacoa | 107.579 | 50.282 | 49.251 | 1.031 |
| San Miguel del Padron | 133.905 | 62.587 | 61.304 | 1.283 |
| Diez de Octubre | 172.873 | 80.801 | 79.145 | 1.656 |
| Cerro | 107.548 | 50.268 | 49.238 | 1.030 |
| Marianao | 114.859 | 53.685 | 52.584 | 1.101 |
| La Lisa | 124.020 | 57.967 | 56.779 | 1.188 |
| Boyeros | 171.666 | 80.237 | 78.592 | 1.645 |
| Arroyo Naranjo | 172.213 | 80.492 | 78.842 | 1.650 |
| Cotorro | 69.287 | 32.385 | 31.721 | 664 |
| Total | 1.827.259 | 854.060 | 836.552 | 17.508 |

Además, para comprender la distribución de la fuerza laboral, consideraremos la relación entre los municipios de residencia de los empleados y los municipios donde se encuentran las entidades empleadoras. La siguiente tabla muestra la distribución porcentual de los empleados por municipio:

| Municipio | Playa | Plaza | Centro Habana | Habana Vieja | Regla | Habana del Este | Guanabacoa | San Miguel del Padron | Diez de Octubre | Cerro | Marianao | La Lisa | Boyeros | Arroyo Naranjo | Cotorro | Total |
|-----------------------|-------|-------|---------------|--------------|-------|-----------------|------------|-----------------------|-----------------|-------|----------|---------|---------|----------------|---------|-------|
| Playa | 53.8 | 19.1 | 3.1 | 5.1 | 0.5 | 0.8 | 0.2 | 0.3 | 0.7 | 3.3 | 4.7 | 2.4 | 4.6 | 1.1 | 0.4 | 46.2 |
| Plaza | 20.2 | 52.6 | 4.8 | 6.6 | 0.9 | 0.6 | 0.3 | 0.6 | 0.8 | 4.2 | 2.1 | 0.8 | 4.4 | 0.9 | 0.1 | 47.4 |
| Centro Habana | 17.2 | 26.9 | 16.3 | 15.1 | 2.1 | 1.2 | 0.6 | 0.5 | 0.9 | 6 | 1.9 | 1.0 | 6.9 | 2.5 | 0.7 | 83.7 |
| Habana Vieja | 12.9 | 18.1 | 5.4 | 43.2 | 3.3 | 1.3 | 0.5 | 0.6 | 1.2 | 4.5 | 1.9 | 0.7 | 4.6 | 1.2 | 0.7 | 56.8 |
| Regla | 9.4 | 11.8 | 3.9 | 15.6 | 36.2 | 4.7 | 4 | 1.2 | 1.1 | 4.3 | 2.3 | 0.3 | 2.0 | 2.0 | 1.1 | 63.8 |
| Habana del Este | 18 | 18.9 | 4.2 | 14.9 | 6.3 | 18.2 | 3.1 | 1.4 | 1.4 | 4.9 | 1.9 | 0.6 | 3.5 | 2.4 | 0.4 | 81.8 |
| Guanabacoa | 11.8 | 11.9 | 3.5 | 11.5 | 9.8 | 10.1 | 18.6 | 2.9 | 2.0 | 6.9 | 1.9 | 0.5 | 4.6 | 2.7 | 1.5 | 81.4 |
| San Miguel del Padron | 16.1 | 15.0 | 3.5 | 10.9 | 4.4 | 3.9 | 3.3 | 14.0 | 3.9 | 8.2 | 2.5 | 0.8 | 5.6 | 4.1 | 3.9 | 86.0 |
| Diez de Octubre | 19.8 | 19.6 | 4.1 | 11.4 | 2.1 | 1.4 | 0.7 | 1.5 | 17.1 | 8.4 | 2.5 | 0.8 | 7.2 | 2.4 | 0.8 | 82.9 |
| Cerro | 19.2 | 22.7 | 4.9 | 11.0 | 1.3 | 1.0 | 1.1 | 0.5 | 2.9 | 22.8 | 2.9 | 1.3 | 6.2 | 1.6 | 0.5 | 77.2 |
| Marianao | 30.1 | 16.3 | 3.2 | 4.4 | 1.0 | 0.8 | 0.3 | 0.4 | 1.2 | 5.7 | 21.2 | 3.6 | 9.1 | 2.0 | 0.7 | 78.8 |
| La Lisa | 36.3 | 17.0 | 3.0 | 3.5 | 0.6 | 0.7 | 0.7 | 0.2 | 1.0 | 3.5 | 6.2 | 18.7 | 5.9 | 2.3 | 0.5 | 81.3 |
| Boyeros | 16.8 | 15.5 | 2.0 | 3.5 | 0.4 | 0.6 | 0.2 | 0.3 | 1.3 | 4.5 | 3.1 | 2.6 | 44.8 | 3.5 | 0.6 | 55.2 |
| Arroyo Naranjo | 16 | 15.8 | 3.4 | 7.4 | 1.7 | 1.6 | 0.6 | 1.2 | 2.4 | 7.0 | 3.9 | 1.0 | 15.4 | 21.2 | 1.4 | 78.8 |
| Cotorro | 10.5 | 13.2 | 2.2 | 5.6 | 1.5 | 4.2 | 1.5 | 1.6 | 2.5 | 6.6 | 1.2 | 1.7 | 7.4 | 5.0 | 35.2 | 64.8 |

En este análisis se considerarán únicamente los estudiantes de niveles preuniversitario, educación técnica y profesional, formación de personal pedagógico y educación superior. Se excluyen las etapas de educación infantil, primaria y secundaria, ya que la mayoría de estos estudiantes son asignados a centros educativos cercanos a sus lugares de residencia. Cabe señalar que, en el caso del nivel preuniversitario, la mayoría de los estudiantes se ubican dentro de su mismo municipio.

A continuación se presenta una tabla que muestra la cantidad de instituciones educativas por municipio para los niveles preuniversitario, técnica y profesional, formación de personal pedagógico y educación especial:

| Municipio | Preuniversitario | Técnica y profesional | Formación Personal Pedagógico | Especial |
|-----------------------|-------------------------|------------------------------|--------------------------------------|-----------------|
| Playa | 3 | 5 | - | 2 |
| Plaza | 3 | 4 | - | 4 |
| Centro Habana | 1 | 3 | - | 4 |
| Habana Vieja | 1 | 2 | - | 2 |
| Regla | 1 | 2 | - | 2 |
| Habana del Este | 3 | 4 | - | 4 |
| Guanabacoa | 2 | 2 | - | 4 |
| San Miguel del Padron | 2 | 2 | - | 4 |
| Diez de Octubre | 4 | 4 | - | 4 |
| Cerro | 2 | 3 | 1 | 5 |
| Marianao | 2 | 3 | - | 5 |
| La Lisa | 2 | 3 | - | 4 |
| Boyeros | 2 | 6 | 2 | 5 |
| Arroyo Naranjo | 4 | 4 | - | 4 |
| Cotorro | 2 | 4 | - | 2 |

Asimismo, se muestra la matrícula de estudiantes para cada municipio y nivel educativo:

| Municipio | Preuniversitario | Técnica y profesional | Formación Personal Pedagógico | Especial |
|------------------|-------------------------|------------------------------|--------------------------------------|-----------------|
| Playa | 2 191 | 4660 | - | 352 |
| Plaza | 2 389 | 3 599 | - | 275 |
| Centro Habana | 726 | 1 330 | - | 285 |
| Habana Vieja | 831 | 139 | - | 178 |
| Regla | 563 | 697 | - | 165 |

| Municipio | Preuniversitario | Técnica y profesional | Formación Personal Pedagógico | Especial |
|-----------------------|-------------------------|------------------------------|--------------------------------------|-----------------|
| Habana del Este | 2 023 | 2 068 | - | 374 |
| Guanabacoa | 746 | 1 340 | - | 580 |
| San Miguel del Padron | 1 087 | 2 059 | - | 399 |
| Diez de Octubre | 2 368 | 3 036 | - | 371 |
| Cerro | 1 020 | 1 766 | 1 178 | 506 |
| Marianao | 1 166 | 1 724 | - | 470 |
| La Lisa | 1 221 | 1 341 | - | 868 |
| Boyerros | 2 444 | 5 025 | 1 842 | 737 |
| Arroyo Naranjo | 3 047 | 1 422 | - | 742 |
| Cotorro | 861 | 1 626 | - | 200 |

Para obtener una estimación del número de estudiantes de educación superior, hemos utilizado los siguientes datos:

- La población no económicamente activa en La Habana es de aproximadamente 973,198 personas.
- De esta población, aproximadamente el 16.1% son estudiantes, lo que equivale a unos 156,685 estudiantes.
- A partir de los 15 años, en niveles de educación no superior, se calcula que hay alrededor de 64,194 estudiantes, lo que deja un total estimado de 92,491 estudiantes de educación superior.

Con base en cifras nacionales disponibles, se estima que la distribución de estudiantes de educación superior es la siguiente:

- 28,672 estudiantes de medicina
- 6,474 de deportes
- 465 de arte
- 21,272 de formación pedagógica
- 12,838 de la Universidad de La Habana
- 10,174 de la CUJAE

El resto de los estudiantes de educación superior residentes en la Habana se distribuye entre universidades de diversas provincias del país.

Para estimar la demanda de transporte público en diferentes horarios del día dividiremos el proceso en dos partes.

Primero nos enfocaremos en las horas pico, que suelen ser de 7:00 a.m. a 10:00 a.m. y de 5:00 p.m. a 8:00 p.m. El primer paso es estimar la cantidad de personas que se mueven para trabajar durante estos períodos. Esto se logra utilizando datos de la población económicamente activa (PEA) y la distribución de la fuerza laboral entre municipios, lo que permite calcular cuántas personas se desplazan entre municipios para llegar a sus empleos. Luego, se debe estimar el número de estudiantes que utilizan el transporte público en estos mismos horarios, considerando las cifras de instituciones educativas por municipio y la matrícula estudiantil por nivel educativo. Una vez que se tienen estos dos componentes principales, se suman para obtener la demanda total de transporte público en las horas pico.

Para calcular la demanda de transporte público en horarios no pico, es necesario considerar que los patrones de viaje son menos regulares y pueden ser más diversos. En primer lugar, hay que tener en cuenta a las personas que trabajan en horarios no tradicionales, los desplazamientos relacionados con actividades recreativas, compras y otros asuntos personales, que tienden a ser más dispersos durante el día y el fin de semana. Además, los viajes para citas médicas, trámites gubernamentales y otros servicios también son factores a tener en cuenta.

Para abordar la demanda de transporte público en horarios no pico, proponemos una simplificación del problema mediante la generación de planes aleatorios basados en la población desempleada y la población no activa económicamente. Aunque reconocemos que este enfoque puede no capturar con precisión los patrones reales de viaje, lo consideramos necesario porque la información específica que necesitamos no está disponible y nuestro equipo carece de la experiencia para derivarla de fuentes existentes.

La principal desventaja de este enfoque es la falta de correlación con la realidad. Al usar datos generados aleatoriamente, no se tiene en cuenta la variedad de motivos que impulsan los viajes fuera de las horas pico, como actividades recreativas, turnos laborales irregulares, visitas médicas, compras, entre otros. Esto puede llevar a resultados poco precisos y a un sistema de transporte menos eficiente.

Además, este método tiene un alto grado de incertidumbre, ya que no puede anticipar eventos inesperados o cambios en el comportamiento de los usuarios. Sin embargo, ante la falta de información concreta y recursos limitados, podría ser un punto de partida para obtener una estimación básica de la demanda.

Definición de la simulación

Supuestos y Limitaciones

Durante el desarrollo del proyecto, se realizaron ciertas simplificaciones y se tomaron algunos supuestos para hacer manejable el problema y facilitar el análisis. Estas consideraciones son las siguientes:

- **Estimaciones no Comprobadas al 100%:** Aunque se utilizaron datos demográficos para estimar ciertos parámetros, no se comprobó al 100% la veracidad de los cálculos. Esto podría afectar la precisión del modelo y su capacidad para reflejar con exactitud la realidad de la población.
- **Selección de Solo 2 Vehículos:** Se decidió enfocarse en dos tipos de vehículos, ómnibus y gazelles, para simplificar el modelo. Reconocemos que esta elección no es completamente realista, ya que el sistema de transporte suele incluir una mayor variedad de vehículos, lo que podría limitar la representatividad del modelo.

- **Esquema Monetario Atrasado:** El modelo utiliza un esquema monetario que puede no reflejar adecuadamente la realidad económica del país. Esto puede influir en las decisiones relacionadas con la economía del transporte y afectar la precisión de las simulaciones financieras.
- **Simplificación del Grafo:** Para simplificar el grafo de la ciudad, se consideraron los nodos como cuadras enteras en lugar de ubicaciones precisas. Esta simplificación permite reducir la complejidad del modelo pero puede limitar la precisión al no considerar la posición exacta de los pasajeros dentro de una cuadra.
- **Ubicación Generalizada de Viviendas:** Para simplificar el problema, se asumió que una persona vive en una cuadra, sin especificar una casa o dirección concreta. Esta suposición facilita el modelo pero puede reducir la precisión al no considerar la variabilidad de la ubicación de las personas dentro de una cuadra.

Definición del Medio

- Entorno Físico:
 - Las calles y avenidas de La Habana, incluyendo intersecciones, cruces, pasos peatonales, señales de pare y ceda el paso, semáforos y lugares de interés.
- Vehículos:
 - Guaguas
 - Gazelle
- Agentes:
 - Conductores de guaguas y gazelle con diferentes niveles de habilidad.
 - Pasajeros que esperan en paradas específicas o solicitan gazelle, dependiendo de sus planes y el estado del medio.

Condiciones Iniciales

- Distribución inicial de guaguas y gazelle en la ciudad.
- Distribución de las rutas y puntos de parada.
- Estado inicial de las guaguas y gazelle (nivel de petróleo, estado mecánico).
- Ubicación inicial de obstáculos (accidentes, reparaciones).

Interacciones y reglas

- Accidentes y Tráfico:
 - Los accidentes ocurren con una probabilidad variable, según la hora del día y la densidad del tráfico.
 - Los obstáculos enlentecedores y bloqueadores afectan el tráfico y la elección de rutas.

Recursos y Restricciones

- Recursos:
 - Cantidad de guaguas y gazelle disponibles.
- Restricciones:
 - Limitaciones en el tiempo para completar rutas.
 - Regulaciones de tráfico.

Escalas y Unidades

- Distancia: Kilómetros.
- Tiempo: Minutos.
- Escala Temporal: Simulación discreta con pasos de tiempo específicos.

Definición de los agentes

Agente Pasajero

Los pasajeros serán modelados como agentes con infraestructura BDI individuales que representan a personas en una muestra representativa de la población. Estos agentes se moverán dinámicamente por la ciudad, interactuando con el entorno y otros agentes, como los conductores de autobuses.

Características

- **Perfil Demográfico:** Cada agente pasajero tendrá un perfil demográfico que reflejará características relevantes para la simulación generadas en base a los datos demográficos disponibles descritos con anterioridad.
- **Objetivos de Viaje:** Cada agente pasajero tendrá objetivos de viaje específicos, como llegar a su lugar de trabajo, hogar, o destino de ocio. Estos objetivos guiarán las decisiones de ruta y comportamiento del agente.

Reglas del Agente Pasajero:

Decisión de Abordaje:

- El agente evalúa las condiciones actuales, como la distancia a la parada de autobús más cercana, la disponibilidad de autobuses en esa parada y su destino.
- Si el tiempo de espera es aceptable y hay espacio disponible en el autobús, el agente decide abordar.
- Si el tiempo de espera es largo o el autobús está lleno, el agente puede optar por esperar en la parada o buscar alternativas de transporte.

Selección de Ruta:

- Una vez a bordo, el agente determina la ruta óptima para llegar a su destino. Esto puede implicar realizar transbordos si es necesario.
- El agente considera factores como la duración del viaje, la cantidad de transbordos y la distancia a caminar desde su salida del autobús hasta su destino.

Interacción con el Entorno:

- Durante el viaje, el agente monitorea activamente el entorno, detectando cambios en la ruta, condiciones del tráfico, paradas adicionales, etc.
- El agente ajusta su comportamiento en respuesta a estos cambios, como decidir bajarse del autobús en una parada no planificada si es más conveniente para llegar a su destino.

Propiedades del Agente Pasajero:

- **Reactividad:** El agente muestra reactividad al percibir su entorno y responder de manera oportuna a los cambios que ocurren durante su viaje en autobús. Por ejemplo, decide abordar el autobús en función del tiempo de espera y la disponibilidad de espacio, así como adapta su comportamiento en respuesta a cambios en la ruta o condiciones del tráfico.

durante el viaje.

- Proactividad: El agente pasajero demuestra proactividad al mostrar un comportamiento dirigido a objetivos (principalmente llegar a un destino) y tomar la iniciativa para lograrlos durante su viaje en autobús. Por ejemplo, planifica su ruta de manera anticipada considerando diversos factores como la duración del viaje y la congestión del tráfico.

Generación de poblaciones

El proceso de generación poblacional se fundamenta en la necesidad de crear poblaciones de pasajeros que sean representativas de la realidad demográfica y de movilidad de La Habana. Para lograr esto, comenzamos con datos concretos sobre la distribución de la población en cada municipio de la ciudad. Estos datos se detallan al inicio de esta sección y las distribuciones de probabilidad inferidas en base a ellos se guardan en 'demographic_data.json'. La clase PopulationGenerator utiliza un enfoque basado en la probabilidad acumulativa para asignar valores aleatorios a una variedad de atributos demográficos, como el municipio de residencia, el estado laboral, el lugar de trabajo y el nivel educativo y generar personas con esto.

Los atributos se generan en un orden específico para que las personas sean lo más realistas posibles. Primero generamos su municipio, con el municipio generamos su estatus de empleo usando los datos sobre la PEA presentados en el caso de estudio, en base a esto generamos el municipio en el que trabajan, y así sucesivamente. Una vez generados estos datos en base a sus distribuciones estimadas, usamos lógica difusa para generar una posible cantidad de dinero y un posible tiempo de espera.

Sistemas difusos para la generación de datos

El uso de sistemas difusos en la generación de poblaciones de pasajeros agrega un nivel de realismo y aleatoriedad controlada al proceso. En nuestro caso, empleamos un sistema difuso para modelar ciertos atributos demográficos y de comportamiento de viaje, como el nivel de ingresos y el tiempo máximo de espera en una parada. En ambos sistemas se utilizó la librería skfuzzy para simplificar el trabajo.

El sistema difuso FuzzySystemMoney se encarga de inferir un estimado del dinero con que anda cada individuo en función de su edad y municipio de residencia. Utiliza reglas difusas basadas en datos históricos para asignar un nivel de ingresos que se ajuste a la distribución observada en la población real. Por ejemplo reconocimos una tendencia a más altos ingresos en municipios como Plaza y Playa en comparación con otros municipios como Arroyo Naranjo. Además, las personas más jóvenes o los ancianos tenían más tendencia a tener bajos ingresos que la población adulta. Aunque 'money' representa el dinero con el que anda la persona y no su nivel general de ingresos, para simplificar el problema asumimos una equivalencia. Dado el bajo coste de los vehículos que simulamos, el dinero no representa un factor tan determinante y no será un problema para la mayoría de personas que se simularán, aún así lo incluimos para tener en cuenta los casos posibles pero poco probables, como que una persona no tenga dinero suficiente para una gacela y tenga que optar obligatoriamente por las guaguas.

El sistema difuso FuzzySystemWaitingTime determina el tiempo máximo de espera en una parada para cada individuo, considerando factores como la edad, el dinero con el que andan y el estado laboral. Para definir las reglas difusas se consideró que por ejemplo las personas más jóvenes suelen ser más impacientes. Además un trabajador o un estudiante suele estar más apurado que una persona en un viaje por motivos de ocio. Además las personas con menos dinero serán más impacientes ya que su capacidad para optar por alternativas de transporte menos económica es más limitada.

Agente Conductor

Agente inteligente que simula el comportamiento de un conductor de ómnibus utilizando lógica basada en reglas:

Supongamos que el entorno de la simulación es una ciudad con calles y paradas de ómnibus, y el objetivo del conductor del ómnibus es recoger pasajeros en las paradas y llevarlos a sus destinos. El ómnibus suele seguir una ruta fija dentro de la ciudad.

Reglas del conductor del ómnibus:

Regla 1: Si hay una parada de ómnibus cerca, detenerse en ella.

Regla 2: Si el ómnibus está lleno, no recoger más pasajeros.

Regla 3: Si no hay pasajeros subiendo o bajando del ómnibus, continuar conduciendo.

Regla 4: Si al terminar una ruta el ómnibus se queda sin combustible, dirigirse a la estación de servicio más cercana.

Regla 5: Si está en la estación de servicio y ya tiene combustible, incorporarse a una ruta.

Regla 6: Si hay un desvío, conducir por la calle despejada y buscar cómo incorporarse a la ruta nuevamente.

Regla 7: Si hay un accidente, conducir por la calle despejada y buscar cómo incorporarse a la ruta nuevamente.

Regla 8: Si las señales de tráfico indican que se debe parar, detener el ómnibus.

Regla 9: Si las señales de tráfico indican que se puede continuar, continuar conduciendo.

El agente conductor posee las siguientes propiedades:

Reactivo:

Es esencial que el agente conductor sea reactivo para percibir y responder a cambios en el entorno, por ejemplo, cuando se llena el vehículo, señales de tráfico o si hay un accidente. Se debe utilizar un bucle de control en el que el agente constantemente perciba su entorno, mediante datos proporcionados por el entorno de simulación. Basado en la información percibida, el agente toma decisiones inmediatas para responder a los cambios en su entorno. Por ejemplo, si detecta algo en la vía, detendrá el vehículo inmediatamente.

Proactivo:

El agente conductor debe ser capaz de tomar decisiones proactivas para alcanzar sus objetivos, como elegir la mejor ruta posible para ir a una estación de combustible al ver que pronto lo necesitará. Para lograr esto se deben definir objetivos claros para el agente y utilizar algoritmos de planificación para alcanzarlos. El agente planifica sus acciones futuras basándose en su conocimiento actual del entorno y sus objetivos.

Adaptable:

Dado que el entorno de tráfico urbano puede ser muy dinámico y sujeto a cambios inesperados, el agente conductor debe ser adaptable. Por ejemplo, si surge un accidente o una construcción en la ruta planificada, el conductor debe ser capaz de reorganizar su plan de acción, encontrar rutas alternativas y comunicar la información a los pasajeros de manera eficiente para minimizar las interrupciones en el servicio.

Integrar las características en un agente puede ser beneficioso para garantizar un rendimiento óptimo en una variedad de situaciones. Un agente reactivo puede proporcionar respuestas rápidas a eventos inmediatos, mientras que un agente adaptable puede ajustar su comportamiento en función de cambios en el entorno a lo largo del tiempo. Esta combinación

puede permitir al agente mantener un equilibrio entre la eficiencia en situaciones normales y la capacidad de adaptarse a situaciones cambiantes.

La diferencia principal radica en cómo cada tipo de agente maneja los cambios en el entorno. Un agente reactivo toma decisiones basadas únicamente en la información presente en el momento, sin tener en cuenta la historia pasada o el futuro, en cambio uno adaptable o proactivo sí tiene en cuenta más información que la del presente.

Sección 3

Conocimiento

El desarrollo de agentes inteligentes para simulaciones de transporte, como en el caso del proyecto del agente "BusDriverAgent", se basa en un marco de inteligencia artificial que emplea reglas lógicas para guiar la toma de decisiones. Este enfoque se alinea con paradigmas tradicionales de IA que utilizan conjuntos de reglas, estructuras de autómatas y lógica condicional para modelar comportamientos complejos.

En este contexto, el agente se estructura como un autómata finito, donde el comportamiento se define por estados y transiciones. El estado actual del agente refleja su situación en un momento dado, como "WAITING_AT_STOP", "DRIVING", o "REFUEL". Cada estado puede tener condiciones específicas que determinan la transición a otro estado.

El proceso de toma de decisiones del agente está dirigido por un conjunto de reglas lógicas que evalúan información contextual proporcionada por el entorno. Estas reglas definen las acciones que debe tomar el agente según las condiciones observadas:

Transiciones entre Estados: Las reglas dictan las transiciones entre diferentes estados del agente. Por ejemplo, si el agente está "WAITING_AT_STOP" y se cumplen ciertas condiciones como la disponibilidad de espacio en el autobús o el tiempo de espera, la regla correspondiente llevará al agente a "START_DRIVING".

Lógica Condicional: Cada regla establece condiciones que se deben cumplir para desencadenar una transición. Estas condiciones se basan en datos del entorno, como señales de tráfico, posición actual, o niveles de combustible. La lógica condicional permite al agente tomar decisiones adaptativas.

Ejecutar Acciones Basadas en Reglas: Cuando un agente toma una decisión basada en una regla, ejecuta acciones correspondientes a esa transición. Estas acciones se modelan como eventos o comportamientos específicos, como detenerse para repostar combustible, cambiar de ruta para evitar un obstáculo, o detenerse en una parada de autobús para recoger pasajeros.

El entorno proporciona la información necesaria para evaluar las condiciones, mientras que las reglas definen cómo el agente debe responder a esos estímulos. El resultado es un agente capaz de tomar decisiones complejas basadas en datos contextuales, permitiendo la simulación precisa de procesos como el transporte urbano y la interacción con otros elementos del sistema.

Grafo

El proceso de construcción del grafo de la ciudad para este proyecto se llevó a cabo utilizando información obtenida de OpenStreetMap (OSM). Esta elección se basó en la gran cantidad de datos geospaciales que OSM proporciona, permitiendo la construcción de un modelo detallado de la infraestructura de transporte. Sin embargo, extraer y procesar esta información presentó desafíos significativos.

Las bibliotecas y herramientas comúnmente utilizadas para trabajar con datos de OSM demostraron tener limitaciones importantes, especialmente cuando se trataba de capturar las restricciones y señalizaciones de tráfico con precisión. Estas deficiencias obligaron a adoptar un enfoque más manual, trabajando directamente con los datos en bruto de OSM para garantizar la exactitud de la información.

En este proceso, surgieron nuevos retos. Se identificaron problemas como rutas inexistentes o mal construidas en la base de datos de OSM, lo que complicó la creación del grafo de la ciudad. Estos errores y omisiones tuvieron que ser corregidos mediante procesos de validación y verificación exhaustivos, asegurando que el grafo reflejara de manera precisa la red de transporte real. En resumen, la construcción del grafo de la ciudad implicó abordar múltiples desafíos técnicos para lograr un modelo confiable y útil para la simulación del transporte.

Enrutamiento

En nuestro proyecto, necesitamos un algoritmo que permita encontrar rutas válidas entre un punto de origen y posibles destinos, sin que siempre sean necesariamente las mejores. La eficacia de la ruta obtenida puede depender de varios factores. Uno de los más relevantes es la habilidad del conductor, representada por una variable entre 0 y 1, la cual influirá significativamente en el camino que el algoritmo elija.

El algoritmo base que utilizamos es el A* debido a su capacidad para priorizar nodos según el costo. Este costo se estima como la suma del costo acumulado del camino hasta el nodo actual y el valor calculado por una heurística, que representa el costo para llegar al objetivo más cercano.

El uso de esta estructura permite que el algoritmo elija el siguiente nodo a explorar con base en la estimación del costo total, minimizando exploraciones innecesarias. Al avanzar en el proceso, sacamos nodos de la cola de prioridad, examinamos sus conexiones y determinamos si deben ser parte del camino final.

La elección de A* se justifica por su equilibrio entre costo acumulado y estimación heurística, lo que facilita la búsqueda eficiente de rutas en entornos complejos. Esto es particularmente útil para nuestro proyecto, donde la flexibilidad y la eficiencia son requisitos clave.

Cotas

Es importante evitar desviaciones excesivas del destino objetivo. Nuestro algoritmo no es genérico; está específicamente diseñado para resolver nuestro problema. Cuando se ejecuta el algoritmo con un único objetivo, indica que hay un obstáculo que bloquea una ruta, requiriendo una reevaluación de la misma. Nuestro enfoque es claro: regresar a la ruta lo antes posible, asegurándonos de no desviarnos demasiado para mantener a los pasajeros cerca de su parada, en caso de que sea necesario omitirla. Si un nodo potencial se aleja demasiado del nodo objetivo, se excluye del camino.

Condiciones de parada

1. Si se llega a uno de los nodos objetivo, el camino se construirá regresando a través de los padres de los nodos visitados, permitiendo generar la ruta resultante.
2. Se han visitado todos los nodos conectados al nodo inicial por algún camino y no se llegó al nodo objetivo a través de ninguno de ellos. Cuando el algoritmo se ejecuta con un único objetivo, esta condición de parada puede alcanzarse más rápidamente. Esto se debe a que el algoritmo descarta las conexiones que se alejan demasiado del nodo objetivo, incluso si podrían llegar a él por un camino más largo. Al hacer esto, se reduce el tiempo y los recursos

utilizados para buscar rutas que probablemente no serán útiles, acelerando la detección de caminos sin salida.

3. Ejecutar el algoritmo con múltiples objetivos nos permite buscar la estación de servicio más cercana. Para determinar cuándo detener la búsqueda, calculamos el promedio de bloques necesarios para alcanzar el objetivo más lejano y lo multiplicamos por un factor preestablecido. Esto nos da un límite en el número de iteraciones que podemos ejecutar antes de parar el algoritmo si no se encuentra una solución. La razón por la que esta condición de parada es importante es porque permite al algoritmo evitar búsquedas excesivas o interminables. Establecer un límite de iteraciones asegura que el algoritmo no gaste recursos innecesariamente si el objetivo es inalcanzable o demasiado lejano, ayudando a mantener la eficiencia general del sistema.

Correctitud del algoritmo

Para ser considerado correcto, el algoritmo debe devolver un camino válido y factible.

Podemos afirmar que el camino resultante de nuestro algoritmo, en caso de existir, siempre será válido, ya que partimos del nodo de origen y seguimos aristas legales hasta llegar al objetivo. Sin embargo, la factibilidad se refiere a que el camino obtenido cumpla con las condiciones del problema. En el caso de un único objetivo, el camino es factible si se mantiene dentro de un rango aceptable. Para múltiples objetivos, el camino es factible solo si puede calcularse en un número limitado de iteraciones.

Como nuestro algoritmo se basa en A* nuestra heurística base debe ser admisible y consistente:

- **Admisibilidad:** Una heurística es admisible si nunca sobreestima el costo para llegar al objetivo. En nuestro caso, la heurística se basa en la distancia entre el nodo actual y el objetivo, utilizando una estimación optimista que supone un camino directo sin obstáculos. Esto garantiza que la heurística no sobreestima, asegurando la admisibilidad.
- **Consistencia:** Nuestra heurística es consistente porque para cualquier nodo su antecesor estaría más lejos y como la velocidad que se toma para el valor de la heurística siempre es la misma, entonces para un nodo más lejano devolvería un valor mayor. Por tanto el valor de la heurística para cualquier nodo n siempre será menor que para cualesquiera de sus antecesores.

Sin embargo, debemos tener en cuenta que al introducir un margen de error en la heurística, esta deja de ser admisible y consistente, lo que puede resultar en rutas que no sean óptimas. A pesar de esto, el margen de error cumple con el objetivo de nuestro problema, que es obtener una ruta acorde con la habilidad del conductor, ya que el error depende solamente de su capacidad.

El uso del algoritmo A* en este contexto se justifica por su capacidad para encontrar rutas óptimas de manera eficiente mediante la combinación de costos acumulados y estimaciones heurísticas. En el contexto del proyecto, donde el grafo de la ciudad puede ser complejo y con múltiples restricciones de tráfico, A* ofrece la adaptabilidad necesaria para encontrar rutas óptimas incluso con desafíos significativos. El algoritmo puede manejar obstáculos, desvíos y múltiples objetivos al ajustar las estimaciones heurísticas para adaptarse a los cambios en el entorno.

Experimentación

Con el proyecto enfrentamos un reto considerable al intentar construir un modelo de simulación lo más realista posible, incorporando todas las calles y conexiones de La Habana, así como las restricciones y señalizaciones de tráfico disponibles. Este enfoque, aunque ambicioso, llevó a una complejidad excesiva en la implementación del grafo, dificultando las pruebas y experimentos necesarios para validar el modelo.

Desafortunadamente, esta complicación resultó en que no pudimos realizar experimentos con la implementación actual. Reconocemos que, al intentar incluir todos los detalles posibles, el modelo se volvió demasiado complejo y desafiante de manejar. Nos disculpamos por este inconveniente y queremos asegurar que estamos preparadas para responder cualquier pregunta al respecto y de los contenidos que debemos manejar de las asignaturas, así como discutir enfoques alternativos.

Que haríamos con más tiempo para remediarlo:

1. Simplificar el grafo de la ciudad para facilitar la simulación. Esto implicaría reducir el número de nodos, conectando solo las rutas más relevantes para el sistema de transporte público. También se enfocaría en reducir el detalle excesivo, como considerar cuadras enteras en lugar de ubicaciones precisas. Este paso sería crucial para mejorar la capacidad de la simulación y permitir experimentos efectivos.
2. Implementación de Nuevas Estrategias de Distribución de Vehículos: Una vez simplificado el grafo, se implementarían diferentes estrategias para la distribución de guaguas y gazelles en las rutas de la ciudad. Los experimentos se centrarían en tres enfoques principales:
 - **Distribuciones Equitativas por Ruta:** Evaluar la eficiencia del sistema cuando los vehículos se distribuyen equitativamente entre las rutas.
 - **Distribuciones Basadas en Densidad Poblacional:** Ajustar la distribución según la densidad de población, permitiendo observar el impacto en la demanda y la eficiencia del transporte.
Distribuciones Aleatorias: Probar distribuciones aleatorias para entender cómo la variabilidad afecta el sistema y si existen patrones inesperados.
3. Con el modelo simplificado y las nuevas estrategias de distribución, se procedería a recopilar datos de simulación. Los principales parámetros de interés incluirían:
 - **Tiempo de Viaje:** Se mediría el tiempo necesario para completar una ruta, permitiendo evaluar la eficiencia de cada estrategia de distribución.
 - **Niveles de Congestión:** Se analizaría el flujo de tráfico y la congestión en diferentes partes del grafo para identificar cuellos de botella y zonas problemáticas.
 - **Tiempo de espera:** Para medir el tiempo de espera, se habría recopilado información sobre la frecuencia de llegada de vehículos a cada parada, los horarios programados, y la variabilidad en la llegada de guaguas y gazelles. Se habría registrado el tiempo exacto que los pasajeros pasan esperando en las paradas para cada ruta.

Conclusiones

El uso de agentes con arquitectura BDI (Beliefs, Desires, Intentions) permitió modelar pasajeros y conductores con comportamientos reactivos y proactivos, aportando un enfoque más dinámico y adaptable a diversos escenarios de transporte. Los agentes mostraron la capacidad de tomar decisiones fundamentadas, adaptándose a las condiciones del entorno, como el tráfico, la disponibilidad de vehículos y señales de tráfico.

Las reglas de inferencia son fundamentales para dotar a los agentes de inteligencia y capacidad de toma de decisiones informadas. Al incorporar estas reglas, los agentes pueden razonar y tomar acciones basadas en datos disponibles y condiciones contextuales, lo que les permite reaccionar de manera inteligente a su entorno.

La implementación del algoritmo A* para el enrutamiento resultó fundamental, destacando por su capacidad para identificar rutas óptimas mediante la combinación de costos acumulados y estimaciones heurísticas. Este algoritmo facilitó la priorización de nodos basándose en el costo total estimado, lo que minimizó las exploraciones innecesarias. A* se mostró eficiente en entornos complejos, adaptándose a obstáculos y condiciones cambiantes, lo que lo hace adecuado para sistemas de transporte urbanos.

Recomendaciones

El uso de reglas lógicas y condiciones contextuales para orientar el comportamiento de los agentes puede plantear desafíos, especialmente cuando se lidia con datos incompletos o ambiguos. Para superar estas dificultades, se recomienda simplificar el modelo de agentes, centrándose en comportamientos esenciales y evitando reglas redundantes o excesivamente complejas.

Una de las limitaciones observadas fue la excesiva complejidad del grafo utilizado para el enrutamiento, lo que dificultó la implementación del algoritmo y la realización de experimentos. Para mitigar este problema, se recomienda simplificar el grafo eliminando nodos y aristas innecesarios. Esto facilitará la experimentación y mejorará el rendimiento del sistema.

Referencias

- Cascetta, E. (2009). *Transportation Systems analysis: Models and Applications*. Springer.
- Carteni, A. (2017). Urban sustainable mobility. Part 2: Simulation models and impacts estimation. *Transport Problems/Problemy Transportu*, 10(1), 5-16. <https://doi.org/10.21307/tp-2015-001>
- Wolfram, S. (2002). *A New Kind of Science*.
- Garcia, L. (20XX) *Temas de Simulación*.