

# Proyecto de IA-Simulación

## *Tema: Movilidad Humana en La Habana*

Yoan René Ramos Corrales C-412      David Cabrera García C-411

14 de octubre de 2024

### **Abstract**

This project proposal for the Artificial Intelligence and Simulation courses aims to develop a mobility trajectory generator for Havana. The generator will simulate the movement of the city's inhabitants, focusing primarily on the bus networks. The primary goal is to estimate the number of residents present in each municipality of the capital over the course of a 24-hour day.

### **Resumen**

Esta propuesta de proyecto para las asignaturas de Inteligencia Artificial y Simulación tiene como objetivo desarrollar un generador de trayectorias de movilidad para La Habana. El generador simulará el movimiento de los habitantes de la capital, considerando principalmente las redes de ómnibus. El objetivo principal es estimar la cantidad de habitantes que se encuentran en cada municipio de la capital durante las 24 horas de un día.

# Índice

<b>1. Introducción</b>	<b>3</b>
<b>2. Metodología</b>	<b>3</b>
2.1. Preliminares . . . . .	3
2.2. Definición del Medio Ambiente . . . . .	3
2.2.1. Características del Medio Ambiente . . . . .	4
2.3. Definición del Agente . . . . .	5
2.3.1. Creencias (Beliefs) . . . . .	5
2.3.2. Deseos (Desires) . . . . .	5
2.3.3. Intenciones (Intentions) . . . . .	5
2.3.4. Preferencias . . . . .	6
2.3.5. Decisiones del Agente . . . . .	6
2.4. Simulación por Eventos Discretos . . . . .	7
2.4.1. Manejo de Eventos . . . . .	7
2.4.2. Interacción de los Agentes con el Entorno . . . . .	7
2.4.3. Eventos Clave en la Simulación . . . . .	7
2.5. Uso de Modelos de Lenguaje para la Configuración y Ejecución de la Simulación . . . . .	7
2.5.1. Generación y Ejecución de la Simulación . . . . .	8
2.5.2. Análisis de los Resultados . . . . .	8
2.5.3. Ventajas del Uso de LLMs en la Simulación . . . . .	8
<b>3. Resultados y Discusión</b>	<b>9</b>
<b>4. Conclusiones</b>	<b>9</b>

# 1 Introducción

Los estudios sobre la movilidad humana se enfocan en describir cómo las personas se desplazan dentro de un sistema o red. El análisis y la predicción de estos movimientos tienen aplicaciones en diversas áreas, como la propagación de enfermedades, la planificación urbana, la ingeniería del tráfico y los mercados financieros, entre otras. En el caso de enfermedades infecciosas como la COVID-19, se sabe que la movilidad humana es un factor clave para su propagación.

Actualmente, se recopila una enorme cantidad de datos sobre trayectorias de movilidad a través de sensores y diversas aplicaciones. Sin embargo, su uso directo presenta desafíos debido a preocupaciones sobre la privacidad, consideraciones comerciales, datos faltantes y altos costos de implementación. Por este motivo, la generación de datos sintéticos de trayectorias de movilidad ha surgido como una tendencia emergente para mitigar las dificultades asociadas al uso de datos reales. La investigación en la generación de trayectorias sintéticas está ganando gran atención, como lo demuestra el creciente número de publicaciones en este campo interdisciplinario.

Atendiendo a lo anterior se propone entonces la creación de un modelo capaz de simular el movimiento de los habitantes de la La Habana considerando principalmente las redes de ómnibus de la capital.

## 2 Metodología

### 2.1 Preliminares

Nuestra simulación se basa en un **modelo de eventos discretos**, que captura los cambios de estado en momentos específicos del tiempo [6]. El sistema está diseñado bajo un enfoque de **sistemas basados en agentes** [8], donde el **medio ambiente** es la red de transporte público de La Habana, y los **agentes** son pasajeros que interactúan con dicho medio.

### 2.2 Definición del Medio Ambiente

El medio ambiente en el que operan los agentes está sobre un **multigrafo**  $G = (V, E, \rho, w)$  que modela la red de transporte público de La Habana [4]. El multigrafo se define de la siguiente manera:

- $V = \{v_1, v_2, \dots, v_n\}$  es el conjunto de vértices que representan las **paradas de transporte público**.
- $E = \{e_1, e_2, \dots, e_m\}$  es el conjunto de aristas, donde cada arista  $e \in E$  representa una **conexión directa** entre dos paradas a través de una ruta específica.
- $\rho : E \rightarrow R$  es una función que asigna a cada arista una ruta específica  $\rho_i \in R$ , donde  $R$  es el conjunto de rutas disponibles en la red de transporte.
- $w : E \rightarrow \mathbb{R}^+$  es una función de peso que asigna una distancia a cada arista, representando la **distancia** entre las paradas por una ruta específica.

### 2.2.1. Características del Medio Ambiente

- **Desplazamiento de las guaguas:** Las guaguas se desplazan a lo largo del multigrafo  $G$ , siguiendo rutas predefinidas. Cada guagua sigue una ruta asignada  $\rho_i \in R$ , que corresponde a un conjunto de aristas en el multigrafo que conectan varias paradas, pero no actúan como agentes autónomos en la simulación.

$$\text{Ruta de la guagua } i : \rho_i = (v_{i1}, v_{i2}, \dots, v_{ik})$$

Donde  $v_{i1}, v_{i2}, \dots, v_{ik} \in V$  son las paradas consecutivas en la ruta  $\rho_i$ .

- **Paradas y municipios:** Cada parada  $v \in V$  pertenece a un municipio específico de La Habana. Formalmente, existe una función  $\mu : V \rightarrow M$ , donde  $M$  es el conjunto de municipios, y  $\mu(v)$  asigna un municipio a cada parada  $v$ . Esto permite modelar la distribución geográfica de las paradas dentro de la ciudad y analizar las rutas dentro de cada municipio o entre diferentes municipios.
- **Movilidad y tiempos de desplazamiento:** Cada guagua tiene un tiempo de desplazamiento asociado a cada arista  $e_{\ell}^{ij} \in E_{ij}$  en su ruta, que está determinado por la función de peso  $w(e_{\ell}^{ij})$ , la cual representa la **distancia** entre las paradas  $v_i$  y  $v_j$  en una ruta específica.
- **Frecuencia de las guaguas:** Cada ruta  $\rho_i$  tiene un intervalo de tiempo  $\tau_i$  que representa la frecuencia con la que las guaguas parten de la primera parada  $v_{i1}$ . Este parámetro es esencial para simular la disponibilidad de transporte en diferentes rutas.
- **Accesibilidad entre paradas:** Para cada agente, es posible llegar desde cualquier parada  $v_i \in V$  a cualquier otra parada  $v_j \in V$  mediante una o más rutas, ya sea de forma directa o a través de transbordos. Esta propiedad garantiza que la red de transporte cubre toda la ciudad y que los pasajeros pueden llegar a sus destinos, incluso si deben cambiar de guagua en varias ocasiones.
- **Agentes pasajeros:** Los pasajeros son los agentes del sistema y toman decisiones sobre qué ruta tomar para llegar a su destino. Estas decisiones dependen del estado del multigrafo (disponibilidad de guaguas, tiempos de espera, distancias entre paradas, etc.).
- **Interacción dinámica:** Los pasajeros interactúan dinámicamente con el multigrafo  $G$ . Deciden abordar una guagua en una parada  $v_i$  dependiendo de la ruta que ésta sigue, la distancia a su destino y el tiempo de espera. A medida que las guaguas se desplazan por las aristas, cambian el estado del entorno, lo cual afecta las decisiones de los pasajeros.
- **Actualización del estado:** El estado del multigrafo  $G$  y de los agentes pasajeros se actualiza continuamente en función de la posición de las guaguas en las rutas, los tiempos de desplazamiento y la disponibilidad de conexiones en cada parada.

## 2.3 Definición del Agente

Cada agente en la simulación está basado en la arquitectura **BDI** (***B**elief-**D**esire-**I**ntention*) [7], donde un agente tiene la capacidad de planificar rutas, tomar decisiones y ajustar su comportamiento según las circunstancias del entorno. Formalmente, el agente  $A$  se define por los siguientes componentes:

### 2.3.1. Creencias (Beliefs)

El agente mantiene una serie de creencias  $\mathcal{B}$  que definen su percepción del entorno y de sí mismo. Estas creencias se modelan como un conjunto de variables de estado:

- $\mathcal{B}.\text{parada\_origen}$ : La parada de transporte público en la que el agente comenzó su viaje.
- $\mathcal{B}.\text{parada\_actual}$ : La parada donde el agente se encuentra actualmente.
- $\mathcal{B}.\text{destino}$ : La parada de destino del agente.
- $\mathcal{B}.\text{paradas\_next}$ : Lista de paradas donde el agente debe cambiar de vehículo en su ruta planificada.
- $\mathcal{B}.\text{current\_time}$ : El tiempo actual en la simulación.
- $\mathcal{B}.\text{ruta\_planificada}$ : La ruta que el agente ha planificado para llegar a su destino.
- $\mathcal{B}.\text{llego\_trabajo}$ : Booleano que indica si el agente ha llegado a su lugar de trabajo.
- $\mathcal{B}.\text{regreso\_casa}$ : Booleano que indica si el agente ha regresado a su casa.
- $\mathcal{B}.\text{cogio\_carro}$ : Booleano que indica si el agente ha decidido utilizar un automóvil privado en lugar del transporte público.
- $\mathcal{B}.\text{caminando}$ : Booleano que indica si el agente ha decidido caminar hacia su destino.

### 2.3.2. Deseos (Desires)

Los deseos  $\mathcal{D}$  del agente son los objetivos que quiere lograr, y pueden ser dinámicos a lo largo del tiempo. En este modelo, los deseos más relevantes son:

- $\mathcal{D}.\text{llegar\_destino}$ : El agente desea llegar a su destino.
- $\mathcal{D}.\text{regresar\_casa}$ : El agente desea regresar a su casa.

### 2.3.3. Intenciones (Intentions)

Las intenciones  $\mathcal{I}$  del agente son los planes concretos que ejecuta para cumplir sus deseos, dadas sus creencias actuales. Por ejemplo:

- $\mathcal{I}.\text{salir\_de\_casa}$ : El agente decide salir de su casa para comenzar su viaje.

- $\mathcal{I}$ .esperar\_guagua: El agente decide esperar una guagua (autobús) en la parada actual.
- $\mathcal{I}$ .caminar: El agente decide caminar hacia su destino si no hay una guagua disponible después de un tiempo o si es una opción más conveniente según sus preferencias.
- $\mathcal{I}$ .coger\_carro: El agente decide utilizar un automóvil privado si la espera en la parada es muy larga o si es más beneficioso.

#### 2.3.4. Preferencias

Cada agente tiene un conjunto de preferencias  $\mathcal{P}$  que influyen la toma de decisiones. Estas preferencias varían entre agentes y son modeladas como valores aleatorios entre 0 y 1, ponderando diferentes factores en la toma de decisiones. Las preferencias principales son:

- $\mathcal{P}$ .rapidez: Preferencia por rutas rápidas.
- $\mathcal{P}$ .comodidad: Preferencia por rutas cómodas (menor número de paradas y cambios de guagua).
- $\mathcal{P}$ .ganancias: Preferencia por poder pagar un automóvil privado.
- $\mathcal{P}$ .laboriosidad: Preferencia por llegar al trabajo.
- $\mathcal{P}$ .condicion\_fisica: Nivel de condición física que afecta la velocidad al caminar o el tiempo máximo que puede soportar antes de decidir tomar un vehículo.
- $\mathcal{P}$ .paciencia: Tolerancia a las esperas prolongadas en la parada. Si se impacienta, el agente podría cambiar sus intenciones y optar por otras rutas o modos de transporte.

#### 2.3.5. Decisiones del Agente

El agente toma decisiones basadas en sus creencias, deseos e intenciones. Estas decisiones incluyen:

- **Elección de Ruta:** El agente utiliza un algoritmo de búsqueda  $A^*$  para planificar su ruta en el multigrafo que representa la red de transporte público [5]. Esta ruta se ajusta dinámicamente según sus preferencias y el estado actual del sistema (como la disponibilidad de guaguas y la congestión).
- **Cambio de Ruta:** Si el agente considera que una ruta alternativa es más favorable (por ejemplo, si una guagua diferente llega a la parada), puede recalcular su ruta usando  $A^*$  con diferentes estrategias [5], como minimizar la distancia o reducir el número de paradas.
- **Impaciencia y Abandono de la Parada:** Si el agente lleva demasiado tiempo esperando una guagua, se puede impacientar. En este caso, el agente puede decidir abandonar la parada y caminar o coger un carro hacia su destino. El tiempo de espera antes de abandonar la parada se calcula en función de su preferencia de paciencia.

## 2.4 Simulación por Eventos Discretos

La simulación desarrollada sigue un enfoque de eventos discretos, en el cual las interacciones de los agentes con el entorno se modelan en función de la ocurrencia de eventos programados en el tiempo. A continuación, se detallan los aspectos clave de este enfoque y cómo los agentes interactúan con el sistema de transporte simulado.

### 2.4.1. Manejo de Eventos

La simulación se organiza en torno a una cola de prioridad (*heap*) de eventos [1]. Cada evento contiene una marca temporal y un conjunto de argumentos que describen la acción a realizar. Los tipos principales de eventos incluyen:

- **Llegada de Agente a Parada:** Un agente alcanza una parada y elige su próximo curso de acción, ya sea subirse a una guagua, continuar caminando, o esperar.
- **Inicialización de Guagua:** Una guagua comienza su recorrido por una ruta específica según la frecuencia de su ruta.
- **Movimiento de Guagua:** La guagua avanza de una parada a la siguiente, permitiendo que los pasajeros suban y bajen.

### 2.4.2. Interacción de los Agentes con el Entorno

En cada evento, los agentes actualizan sus creencias e intenciones basándose en la información disponible en el entorno tal como se describe en la sección 2.3.

### 2.4.3. Eventos Clave en la Simulación

El flujo de la simulación depende del desencadenamiento de los eventos y las decisiones de los agentes. Los eventos se procesan en orden cronológico, lo que permite que la simulación avance de manera dinámica. El entorno, representado por un grafo de conexiones de rutas de guaguas, proporciona el contexto para las acciones de los agentes. La interacción de los agentes con este grafo y con las guaguas es fundamental para el éxito de sus trayectorias.

A medida que los agentes completan sus trayectos o regresan a sus puntos de origen, el sistema de eventos sigue generando nuevas acciones, asegurando que la simulación capture la complejidad de la movilidad en un entorno urbano dinámico.

## 2.5 Uso de Modelos de Lenguaje para la Configuración y Ejecución de la Simulación

En el desarrollo de la simulación de movilidad humana para la ciudad de La Habana, se ha implementado un enfoque que integra modelos de lenguaje avanzados (LLMs, por sus siglas en inglés) para automatizar tanto la configuración como la ejecución de la simulación [3]. Este enfoque permite a los usuarios interactuar con el sistema mediante lenguaje natural, facilitando el proceso de configuración y la posterior interpretación de los resultados generados por la simulación.

### 2.5.1. Generación y Ejecución de la Simulación

El modelo de lenguaje utilizado, *Gemini-1.5 Flash* [2], ha sido integrado al sistema mediante una API, permitiendo que, a partir de una entrada textual del usuario, se genere un archivo de configuración en formato JSON que contiene los parámetros clave para la simulación. Estos parámetros incluyen:

- Cantidad de ómnibus.
- Tiempo total de simulación.
- Número de agentes.
- Archivos de distribución de población y trabajos por municipio.
- Frecuencias de salida de las rutas de transporte.
- Probabilidades de retorno a los hogares.
- Intervalos de partida de los agentes.

Tras la generación del archivo JSON, el sistema procede a ejecutar automáticamente la simulación utilizando los parámetros establecidos por el modelo de lenguaje. La ejecución se realiza sin intervención adicional del usuario, permitiendo que el proceso sea ágil y eficiente.

### 2.5.2. Análisis de los Resultados

Al finalizar la simulación, el LLM no solo proporciona los resultados brutos generados, sino que además realiza un análisis detallado de los mismos. Los resultados de la simulación, que se guardan en archivos CSV, son procesados y formateados para su presentación de manera comprensible. A través del modelo de lenguaje, el usuario puede recibir un análisis estadístico completo de los datos, acompañado de gráficos generados automáticamente para facilitar la visualización de las dinámicas de movilidad. Estos gráficos incluyen la distribución de personas por municipio a lo largo del día, reflejando el impacto de las diferentes configuraciones de transporte y comportamiento de los agentes.

Un ejemplo típico de interacción con el LLM puede incluir preguntas como: “¿Qué sucede si solo está disponible la ruta P11?” o “¿Cómo afectaría a la población de Habana del Este la falta de guaguas en la ruta P1?”. El LLM ajusta los parámetros de la simulación según la consulta, ejecuta la simulación y presenta un análisis exhaustivo de los resultados, considerando las condiciones impuestas por el usuario.

### 2.5.3. Ventajas del Uso de LLMs en la Simulación

- Automatización integral del proceso de configuración, ejecución y análisis de simulaciones.
- Reducción significativa de la complejidad para los usuarios, permitiendo interactuar mediante lenguaje natural sin necesidad de conocimientos técnicos detallados.



- Capacidad de personalizar dinámicamente la simulación en respuesta a preguntas o hipótesis específicas planteadas por el usuario.
- Generación automática de informes estadísticos y gráficos que facilitan la interpretación de los resultados.

### 3 Resultados y Discusión

### 4 Conclusiones

### Referencias

- [1] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*. MIT Press, Cambridge, MA, 3rd edition, 2009. Contains detailed explanations of heaps and other data structures.
- [2] Google DeepMind. Gemini-1.5 flash language model. <https://developers.generativeai.google>, 2023. Consultado: 2024-10-14.
- [3] A. Brown et al. Large language models: A survey. <https://arxiv.org/abs/2005.14165>, 2020. Consultado: 2024-10-14.
- [4] Daniel Evans and Rebecca Smith. A multigraph approach to modeling public transportation networks. *Journal of Transportation Research*, 48:14–23, 2021. Discusses the use of multigraphs to represent public transportation systems.
- [5] Peter E. Hart, Nils J. Nilsson, and Bertram Raphael. A formal basis for the heuristic determination of minimum cost paths. In *IEEE Transactions on Systems Science and Cybernetics*, volume 4, pages 100–107. IEEE, 1968. Introduces the A\* algorithm for pathfinding.
- [6] Averill M. Law. *Simulation Modeling and Analysis*. McGraw-Hill Education, New York, NY, 5th edition, 2014. Focuses on discrete-event simulation techniques and their applications.
- [7] Anand S. Rao and Michael P. Georgeff. Bdi agents: From theory to practice. *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS)*, pages 312–319, 1995. Describes the BDI (Belief-Desire-Intention) model for agent systems.
- [8] Michael Wooldridge. *An Introduction to MultiAgent Systems*. John Wiley & Sons, Chichester, UK, 2nd edition, 2009. Overview of agent-based systems, including BDI agents.