



**EDUCACIÓN**

SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO  
NACIONAL DE MÉXICO

# Tecnológico Nacional de México

Centro Nacional de Investigación  
y Desarrollo Tecnológico

## Tesis de Maestría

Sistema de predicción de tiempos de arribo utilizando el sistema  
GTFS

presentada por  
**Ing. Andres Ocampo Ortiz**

como requisito para la obtención del grado de  
**Maestra en Ciencias de la Computación**

Director de tesis  
**Dra. Alicia Martínez Rebollar**

Cuernavaca, Morelos, México. Julio de 2021.



**EDUCACIÓN**  
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO  
NACIONAL DE MÉXICO

Centro Nacional de Investigación y Desarrollo Tecnológico  
Departamento de Ciencias Computacionales

Cuernavaca, Mor., **25/junio/2021**

OFICIO No. DCC/158/2021  
Asunto: Aceptación de documento de tesis  
CENIDET-AC-004-M14-OFICIO

**DR. CARLOS MANUEL ASTORGA ZARAGOZA**  
SUBDIRECTOR ACADÉMICO  
PRESENTE

Por este conducto, los integrantes de Comité Tutorial del C. **ANDRÉS OCAMPO ORTIZ**, con número de control M17CE099, de la Maestría en Ciencias de la Computación, le informamos que hemos revisado el trabajo de tesis de grado titulado "**SISTEMA DE PREDICCIÓN DE TIEMPOS DE AJRIBO UTILIZANDO EL SISTEMA GTFB**" y hemos encontrado que se han atendido todas las observaciones que se le indicaron, por lo que hemos acordado aceptar el documento de tesis y le solicitamos la autorización de impresión definitiva.

Dra. Alicia Martínez Rebolgar  
Directora de tesis

Dr. Hugo Estrada Esquivel  
Codirector de Tesis

Dr. Márquez-Espinoza Sánchez  
Revisor 1

Dr. Joaquín Pérez Ortega  
Revisor 2

C.c.p. Depto. Servicios Escolares.  
Expediente / Estudiante  
JSGG/iba

**cenidet**  
Centro Nacional de Investigación y Desarrollo Tecnológico



Interior Internado Palmira S/N, Col. Palmira

C.P. 62490, Cuernavaca, Morelos

Tel. (01) 777 3 62 77 70, ext. 3201.

e-mail: [dcc@cenidet.tecnm.mx](mailto:dcc@cenidet.tecnm.mx)

[www.tecnm.mx](http://www.tecnm.mx) | [www.cenidet.tecnm.mx](http://www.cenidet.tecnm.mx)





**EDUCACIÓN**  
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO  
NACIONAL DE MÉXICO

Centro Nacional de Investigación y Desarrollo Tecnológico

Subdirección Académica

Cuernavaca, Mor.,  
No. De Oficio:  
Asunto:

09/Julio/2021  
SAC/091/2021  
Autorización de  
impresión de tesis

**ANDRÉS OCAMPO ORTIZ**  
CANDIDATO AL GRADO DE MAESTRO EN CIENCIAS  
DE LA COMPUTACIÓN  
PRESENTE

Por este conducto, tengo el agrado de comunicarle que el Comité Tutorial asignado a su trabajo de tesis titulado "SISTEMA DE PREDICCIÓN DE TIEMPOS DE ARRIBO UTILIZANDO EL SISTEMA GTFPS", ha informado a esta Subdirección Académica, que están de acuerdo con el trabajo presentado. Por lo anterior, se le autoriza a que proceda con la impresión definitiva de su trabajo de tesis.

Esperando que el logro del mismo sea acorde con sus aspiraciones profesionales, reciba un cordial saludo.

**ATENTAMENTE**

Excelencia en Educación Tecnológica®  
"Educación Tecnológica al Servicio de México"

  
DR. CARLOS MANUEL ASTORCA ZARAGOZA  
SUBDIRECTOR ACADÉMICO



C.p. M.E. Guadalupe Garrido Rivera, Jefa del Departamento de Servicios Escolares.  
Expediente.

CMAZ/CHG

**cenidet**  
Centro Nacional de Investigación  
y Desarrollo Tecnológico



Interior Internado Palmira S/N, Col. Palmira, C. P. 62490,  
Cuernavaca, Morelos Tel. (01) 777 3 62 77 70, ext. 3201,  
e-mail: acad\_cenidet@tecnm.mx  
www.tecnm.mx | www.cenidet.tecnm.mx



## **Agradecimientos**

Primeramente, gracias a toda mi familia y amigos por el apoyo incondicional que me brindaron para poder lograr este objetivo. En especial a mi madre mi principal motor y a mi tía Angela que me dieron los argumentos y la motivación necesaria para concluir esta etapa incluso cuando la motivación parecía perdida, a mis abuelas que se encuentran descansando y no pudieron ver la conclusión de esta etapa y siempre me apoyaron en vida, a mi abuelo y mi padre por su gran apoyo moral, a mis hermanos que siempre están para mí cuando los necesito y me ayudan a superarme cada día más, a mi sobrina Helena que su llegada al mundo me motivo aún más para concluir esta etapa y a mi mejor amigo Josué que me ayudó en la realización de las pruebas de este trabajo.

También, deseo agradecer al Consejo Nacional de Ciencia y Tecnología (CONACYT) por este gran apoyo y posibilidad de crecimiento que me brindó a través de la manutención de su beca para poder realizar el desarrollo de este trabajo de investigación.

Al Tecnológico Nacional de México (TecNM) y al campus Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET) por brindarme hospitalidad en sus instalaciones y prestarme el equipo necesario para poder desarrollar este trabajo de investigación.

A mi directora de tesis la doctora Alicia Martínez Rebollar a quien nunca olvidare por el absoluto apoyo, consejos, ideas y motivación brindada para concluir este trabajo de investigación.

Al doctor Hugo Estrada Esquivel, al doctor Máximo López Sánchez y al doctor Joaquín Pérez Ortega por el apoyo, consejos e ideas aportadas para mejorar este trabajo de investigación.

## **Dedicatoria**

A toda mi familia y amigos por el apoyo incondicional que me brindaron para poder lograr este objetivo. En especial a mi madre mi principal motor y a mi tía Angela que me dieron los argumentos y la motivación necesaria para concluir esta etapa incluso cuando la motivación parecía perdida, a mis abuelas que se encuentran descansando y no pudieron ver la conclusión de esta etapa y siempre me apoyaron en vida, a mi abuelo y mi padre por su gran apoyo moral, a mis hermanos que siempre están para mí cuando los necesito y me ayudan a superarme cada día más, a mi sobrina Helena que su llegada al mundo me motivó aún más para concluir esta etapa.

## **Resumen**

Según el INEGI (2016) el transporte público es un servicio ofrecido diariamente hacia la comunidad, el cual tiene como objetivo ayudar a las personas a trasladarse de un lugar a otro con la finalidad de ahorrarles tiempo y esfuerzo. Miles de personas lo utilizan diariamente para llegar a sus trabajos, escuelas, lugares de entretenimiento etc.

Las ciudades inteligentes es un término que ha ido ganando mucha popularidad en la actualidad, ya que ha captado la atención de los medios de comunicación, las redes sociales y los foros políticos de forma intensiva y recurrente. También se ha vuelto un tema de interés para las empresas y para los ámbitos académicos, ya que es un área la cual se pueden desarrollar grandes proyectos los cuales podrían contribuir a la evolución de las ciudades.

Día con día surgen nuevas necesidades las cuales pueden ser resueltas con la tecnología, un ejemplo de esto son los Sistemas de Transporte Inteligentes que tienen como objetivo mejorar la seguridad y eficiencia en el transporte terrestre.

Este tipo de sistemas facilitan la vida diaria de las personas, ya que otorgan información con la cual se puede administrar de mejor manera el tiempo y tener una mejor planificación del día o inclusive evitar algunos eventos como el no conocer donde se encuentran las paradas de los autobuses, estar expuesto a un asalto, entre otros más.

En este trabajo de investigación se presenta el desarrollo de un sistema de predicción de tiempos de arribo del transporte público utilizando el sistema GTFS como base.

Sin más preámbulos se presenta el documento.

## Contenido

Capítulo 1 Introducción.....	1
1.1.    Introducción. ....	2
1.2.    Planteamiento del problema .....	2
1.3.    Objetivo general .....	3
1.4.    Objetivos específicos.....	3
1.5. Estructura de la tesis.....	3
Capítulo 2 Marco Teórico .....	4
2.1    Ciudades Inteligentes .....	5
2.1.1 Transporte público.....	5
2.1.2    Sistemas de Transportes Inteligentes .....	6
2.1.3    Sistemas de información al viajero .....	7
2.2    Geo localización.....	7
2.2.1 GPS .....	7
2.2.2 Red de sensores inalámbricos.....	7
2.3    GTFS .....	8
Capítulo 3 Estado del arte .....	9
3.1    Clasificación del estado del arte.....	10
3.2    Aspectos descriptivos de los trabajos del estado del arte de predicción de tiempos de arribo del transporte público .....	10
3.3    Algoritmos heurísticos .....	10
3.3.1    Sistema de ubicación en tiempo real para el servicio de autobuses comunitarios ....	11
3.3.2    Sistema inteligente de toma de decisiones para pasajeros de transportes urbanos....	11
3.3.3    Implementación de un sistema de monitoreo e información de pasajeros del autobús en tiempo real .....	12
3.4    Algoritmos clasificadores.....	12
3.4.1    Predicción del tiempo de viaje del autobús utilizando una red neuronal artificial ....	12
3.4.2    Predicción probabilística del avance del autobús utilizando la regresión de máquina de vectores de relevancia.....	13
3.4.3    Modelo de predicción del tiempo de viaje de la red neuronal artificial para autobuses que utilizan solo datos de GPS .....	14
3.5    Sistemas que consideran GTFS.....	14
3.5.1    Brindando servicios de información en tiempo real sobre el transporte publico .....	14

3.5.2	Visualización de datos de especificaciones de alimentación de tránsito general .....	15
3.5.3	Los múltiples usos de los datos GTFS – abriendo la puerta al tránsito y a las aplicaciones multimodales .....	15
3.6	Sistemas que utilizan Metamodelos .....	16
3.6.1	Incertidumbre en las predicciones del tiempo de arribo del autobús: tratamiento de la heterocedasticidad con un enfoque de metamodelo .....	16
3.6.2	Predicción de la hora de llegada del autobús basada en un modelo mixto.....	17
Capítulo 4	Metodología de solución .....	18
4.1	Descripción general de la metodología de solución .....	19
4.2	Fase 1: Recopilación de información .....	20
4.3	Fase 2: Desarrollo del algoritmo .....	25
4.4	Fase 3: Desarrollo del sistema de software .....	30
4.4.1	Servidor Web.....	31
4.4.2	Desarrollo de la Web API .....	31
4.4.3	Placa LinkIt One.....	32
4.4.4	Aplicación Android Andy_Bus.....	34
4.5	Fase 4: Pruebas del sistema.....	38
Capítulo 5	Pruebas y resultados .....	39
5.1	Descripción de las pruebas .....	40
5.2	Pruebas Fase 1 .....	40
5.3	Pruebas Fase 2.....	40
5.4	Pruebas en Escenario 1: Predicción de tiempos de arribo.....	40
5.5	Pruebas en Escenario 2: Predicción de tiempos de arribo.....	44
5.6	Resumen de resultados y conclusiones .....	49
Capítulo 6	Conclusiones.....	50
6.1	Conclusiones .....	51
6.2	Logros Obtenidos .....	52
6.3	Trabajos futuros .....	54
Referencias	.....	55

## Índice de Figuras

Figura 4. 1	Metodología de solución.....	20
-------------	------------------------------	----



Figura 4. 2 Información GTFS .....	21
Figura 4. 3 Ejemplo de información en tabla stops .....	22
Figura 4. 4 Modelo entidad relación de la base de datos depurada.....	22
Figura 4. 5 Aplicación que obtiene longitud y latitud de las paradas.....	23
Figura 4. 6 Información de las paradas de las rutas .....	24
Figura 4. 7 Información histórica obtenida .....	24
Figura 4. 8 Creación de vectores en el camino.....	26
Figura 4. 9 Paso 1: Desarrollo del algoritmo.....	26
Figura 4. 10 Ejemplo del primer paso del algoritmo.....	26
Figura 4. 11 Paso 2: Desarrollo del algoritmo.....	27
Figura 4. 12 Mapa con la información de las paradas .....	28
Figura 4. 13 Ejemplo de paradas con sus tiempos históricos registrados .....	28
Figura 4. 14 Paso 3: Desarrollo del algoritmo.....	29
Figura 4. 15 Algoritmo completo.....	30
Figura 4. 16 Arquitectura del sistema Andy_Bus .....	31
Figura 4. 17 Datos enviados por la web API.....	32
Figura 4. 18 Servicio rastreador de la placa LinkIt One .....	33
Figura 4. 19 Placa LinkIt One y sus accesorios .....	34
Figura 4. 20 Pantalla de inicio de sesión de Andy_Bus .....	35
Figura 4. 21 Pantalla de registro para nuevos usuarios .....	35
Figura 4. 22 Pantalla con los autobuses registrados .....	36
Figura 4. 23 Pantalla de selección de dirección de autobús .....	36
Figura 4. 24 Pantalla de selección de parada.....	37
Figura 4. 25 Pantalla con predicción de tiempo de arribo.....	37
Figura 4. 26 Pantalla de mapa .....	38
Figura 4. 27 Página de administradores .....	38
Figura 5. 1 Resultados viaje 1 con dirección a Guacamayas .....	42
Figura 5. 2 Resultados viaje 2 con dirección a Guacamayas .....	43
Figura 5. 3 Resultados viaje 3 con dirección a Guacamayas .....	44
Figura 5. 4 Resultados viaje 1 con dirección a Av. Universidad .....	46
Figura 5. 5 Resultados viaje 2 con dirección a Av. Universidad .....	47
Figura 5. 6 Resultados viaje 3 con dirección a Av. Universidad .....	48

## Índice de Tablas

Tabla 4. 1 Información de las bases de datos .....	25
Tabla 5. 1 Descripción de paradas de la ruta 1 .....	41
Tabla 5. 2 Resultados viaje 1 con dirección a Guacamayas.....	42
Tabla 5. 3 Resultados viaje 2 con dirección a Guacamayas.....	43

Tabla 5. 4 Resultados viaje 3 con dirección a Guacamayas.....	44
Tabla 5. 5 Descripción de paradas de la ruta 2.....	45
Tabla 5. 6 Resultados viaje 1 con dirección a Av. Universidad.....	46
Tabla 5. 7 Resultados viaje 2 con dirección a Av. Universidad.....	47
Tabla 5. 8 Resultados viaje 3 con dirección a Av. Universidad.....	48
Tabla 5. 9 Tabla de comparaciones.....	48

# Capítulo 1 Introducción

## **1.1. Introducción.**

El transporte público es un servicio ofrecido diariamente hacia la comunidad que tiene como objetivo ayudar a las personas a trasladarse a un destino con la finalidad de ahorrarles tiempo y esfuerzo. Miles de personas utilizan diariamente el transporte público para llegar a sus trabajos, escuelas, lugares de entretenimiento entre otros más.

Actualmente la tecnología está siendo utilizada como una herramienta para mejorar los servicios del transporte público.

Un claro ejemplo de esto son los Sistemas de Transporte Inteligentes los que tienen como objetivo mejorar la seguridad, eficiencia y el confort en el uso del transporte terrestre. Algunos ejemplos de las aplicaciones desarrolladas con este objetivo son los sistemas de cálculo de tráfico, sistemas de navegación entre destinos, o los sistemas para predicción de tiempos de arribo. Este tipo de sistemas facilitan la vida diaria de las personas, ya que otorgan información con la cual es posible tener una mejor planificación del tiempo que se requiere para alcanzar un destino, conocer la ubicación de las paradas de los autobuses, evitar estar expuestos a asaltos entre otros [11].

Uno de los problemas más recurrentes en los actuales sistemas de transporte público es la determinación del tiempo que un usuario necesita esperar para que arribe el siguiente transporte en su ruta. En este trabajo de investigación se desarrolló un sistema de software para la predicción de tiempos de arribo del transporte público, el cual está adaptado a la especificación estándar GTFS, el cual permite tener un mejor control de la ubicación precisa de las diversas paradas de una línea de autobuses y de la distancia entre ellas.

## **1.2. Planteamiento del problema**

La importancia del transporte público en las ciudades de nuestro país es indiscutible. Para el caso de la Ciudad de México se establece que: “cuenta con una población de 8,851,080 habitantes (la Zona Metropolitana del Valle de México tiene 20,137,152 habitantes), según el censo oficial 2010. La mayoría de las personas utilizan a diario la red de transporte público para viajar desde sus hogares a los centros de trabajo, educativos o de entretenimiento. Dicha red incluye distintos medios como el tren ligero, el transporte colectivo concesionado, el Metro y el Metrobús, entre otros. Tan sólo el Metro, según cifras oficiales, transportó en el año 2010 a más de mil cuatrocientos diez millones de usuarios” [9].

En el caso de la Ciudad de México a pesar de contar con un sistema de transporte público enorme, se cuenta con muy poca tecnología que permite lograr el confort de sus usuarios. En el caso de países europeos y algunos países en América como Estados Unidos cuenta con sistemas de predicción de tiempos de arribo de autobuses públicos, lo cual facilita a los ciudadanos la planeación del tiempo requerido para desplazarse en sus ciudades. Los sistemas de predicción de tiempos de arribo que existen en México cuentan con características y funcionalidades similares a los de los otros países, por ejemplo, tenemos el caso del sistema Deja-bus [8]. Sin embargo, estos sistemas tienen algunos inconvenientes, uno de ellos es que algunos de estos sistemas utilizan el sistema GPS que se encuentra instalado en los autobuses como es el caso por ejemplo del Sistema de Metrobuses de la Ciudad de México. Sin embargo, estos sistemas propietarios no ofrecen funcionalidades para tener acceso a la localización, lo cual dificulta su uso para fines de experimentación. Otro de los

inconvenientes de los sistemas de llegadas de autobuses es que no utilizan formatos estandarizados para la representación de las rutas y las paradas correspondientes a esa ruta. Los sistemas de estandarización de tránsito permiten la definición de los horarios de llegada de cada uno de los autobuses, así como los horarios de inicio y finalización del recorrido de autobuses de cada ruta. Una de las ventajas del uso de formatos estandarizados como es GTFS es que los usuarios, al tener conocimiento de los horarios, pueden planificar mejor sus alternativas de viaje. Actualmente, por ejemplo, los usuarios no suelen conocer los horarios nocturnos de los autobuses, por lo que podrían esperar por autobuses en rutas que ya no se encuentran laborando.

En este trabajo de tesis se propone un sistema de predicción de tiempos de arribo del transporte público similar a los trabajos mencionados anteriormente, con la diferencia de que se propone un dispositivo específico para obtener la posición GPS de los autobuses y que, además, se utiliza la especificación estándar GTFS para conocer las paradas de los autobuses. El sistema propuesto integra un algoritmo para calcular los tiempos de arribo de los autobuses. De esta forma es posible mostrar al usuario la hora en la que arribarán los autobuses a sus paradas, indicando también si alguno de estos autobuses se encuentra con algún retraso. Esto tiene como objetivo que los usuarios de transporte público puedan planificar sus viajes de mejor manera.

### 1.3. Objetivo general

El objetivo de esta tesis es el desarrollo de un sistema de software que permita la predicción de los tiempos de arribo de transporte público utilizando información del estándar GTFS para la descripción de rutas y paradas y el uso de un dispositivo de rastreo GPS para obtener la posición actual de los autobuses.

### 1.4. Objetivos específicos

- Desarrollar un algoritmo para predecir los tiempos de arribo de los autobuses.
- Desarrollar un sistema de software que implementa el algoritmo desarrollado.
- Diseñar y desarrollar un dispositivo para el rastreo de autobuses utilizando tecnología GPS.
- Desarrollo de un sistema para recolectar la información de los autobuses tomando como base el sistema GTFS.
- Conseguir e implementar la información necesaria del estándar GTFS.

### 1.5. Estructura de la tesis

El contenido de la tesis se encuentra organizado en capítulos de la siguiente manera:

1. **Capítulo 2. Marco teórico:** En este capítulo se definen los conceptos principales que se utilizan en esta tesis.
2. **Capítulo 3. Estado del arte:** En este capítulo se describen los trabajos de investigación relacionados con los sistemas de transporte inteligente.
3. **Capítulo 4. Metodología de solución:** En este capítulo se describe la metodología de solución propuesta que se utilizó para el desarrollo del sistema de transporte inteligente para la predicción de tiempos de arribo de los autobuses del transporte público.
4. **Capítulo 5. Pruebas y Resultados:** En este capítulo se presentan las pruebas realizadas utilizando el sistema desarrollado y validar la efectividad del predictor de tiempos de arribo.

5. **Capítulo 6. Conclusiones:** En este capítulo se presentan las conclusiones obtenidas al realizar este proyecto.

## Capítulo 2 Marco Teórico

## Descripción

En este capítulo se describen algunos de los conceptos que requieren ser definidos con claridad para lograr un claro entendimiento del proyecto desarrollado.

### 2.1 Ciudades Inteligentes

Es un término que ha ido ganando mucha popularidad en la actualidad, ya que ha captado la atención de los medios de comunicación, las redes sociales y los foros políticos de forma intensiva y recurrente. También se ha vuelto un tema de interés para las empresas y para los ámbitos académicos, ya que es un área en la cual se pueden desarrollar grandes proyectos los cuales podrían contribuir a la evolución de las ciudades.

El término Ciudades Inteligentes apareció en los años 90 como un modelo urbano basado en tecnología con la finalidad de abordar los grandes retos y preocupaciones que se iban presentando en las ciudades actuales tales como mejorar la eficiencia energética, disminuir las emisiones contaminantes y reconducir el cambio climático. El modelo de las ciudades inteligentes está basado en la conjunción de innovaciones tecnológicas en las áreas de energía, transporte y tecnologías de información y comunicación.

Una de las definiciones del concepto de ciudad inteligente es la proporcionada en [6]. “Una Ciudad Inteligente debe ser aquella donde el progreso se cataliza mediante la conjunción íntima entre innovaciones en las áreas de energía, transporte y tecnologías de información y comunicación” [6].

Otra de las definiciones propuestas para ciudad inteligente establece que “Una Ciudad Inteligente es aquella que aplica las tecnologías de información y comunicación para la mejora de la calidad de vida y la accesibilidad de sus habitantes y asegura un desarrollo sostenible económico, social y ambiental en mejora permanente. Una ciudad inteligente permite a los ciudadanos interactuar con ella de forma multidisciplinaria y se adapta en tiempo real a sus necesidades” [6].

#### 2.1.1 Transporte público

En la vida cotidiana las personas tienen la necesidad de trasladarse a diferentes lugares ya sea a su trabajo, escuela, lugares de entretenimiento, entre otros. Uno de los métodos que utilizan para llegar rápidamente y sin esfuerzo es el transporte público, el cual se describe a continuación.

El transporte público comprende los medios de transporte en los cuales los pasajeros no son los propietarios de los mismos, siendo servidos por terceros (empresas públicas o privadas). El transporte público ayuda al desplazamiento de las personas de un punto a otro en un área de una ciudad, pagando cada persona una tarifa establecida dependiendo de su recorrido [1].

El transporte público tiene gran importancia en la vida cotidiana ya que ayuda a las personas que no pueden costearse un automóvil propio a desplazarse por las ciudades, y también ayuda para reducir la contaminación, ya que reduce el número de automóviles particulares en circulación.

### **2.1.2 Sistemas de Transportes Inteligentes**

En la actualidad la tecnología se ha ido implementando en diferentes áreas de conocimiento, una de las nuevas disciplinas en donde está teniendo gran aportación es en el transporte, ya que ha desarrollado diferentes sistemas para que este pueda ser beneficiado, un ejemplo es el GPS con el cual las personas pueden obtener su posición geográfica, este ha ayudado a desarrollar diferentes aplicaciones para el transporte, un ejemplo es el fluido de tráfico en un determinado punto, con estos datos se puede obtener los caminos más fáciles a seguir para llegar a un destino. Otro claro ejemplo es el de los Sistemas de Transporte Inteligente.

Se conoce como Sistemas Inteligentes de Transporte al conjunto de aplicaciones informáticas y sistemas tecnológicos creados con el objetivo de mejorar la seguridad y eficiencia en el transporte terrestre (carreteras y ferrocarriles), facilitando la labor de control, gestión y seguimiento por parte de los responsables [2].

Estos sistemas obtienen la información de los diferentes elementos de interés de las carreteras, que una vez procesada y analizada, se utiliza para mejorar la seguridad de los conductores, mejorando el tráfico y la comodidad en los desplazamientos [2].

Uno de los componentes importantes del sistema de transporte inteligente es la infraestructura de tráfico. La solución más conveniente que se podría tomar para aliviar la congestión del tráfico es ampliar la capacidad de la red de transporte a través de la construcción de nuevas infraestructuras de tráfico, pero, lamentablemente, está restringida por su extensa duración de recorrido, el alto costo y porque el espacio geográfico se encuentra limitado. Para poder enfrentar dichos problemas, usar completamente las infraestructuras de transporte existentes se vuelve cada vez más importante. En los últimos años, una nueva técnica se ha utilizado la cual es la reserva de carril, como una técnica flexible para la gestión del tráfico, se ha aplicado ampliamente en la vida real.

El objetivo de esta técnica es reservar carriles en una red de transporte existente para que los usuarios especiales cumplan con los requisitos especiales de transporte [4].

Una aplicación o un framework que muestre el tiempo de llegada del autobús en las paradas de autobús puede disminuir la ansiedad de los pasajeros que esperan el autobús. Distribuir información acerca de los tiempos de arribo de los autobuses a través de algunas interfaces, como teléfonos inteligentes, tabletas u otros dispositivos más podría ocasionar como una consecuencia que el sistema de transporte público sea más sencillo de utilizar, por lo cual, esto puede aumentar su competitividad entre los diversos medios de transporte. El aumento del tiempo de espera y la incertidumbre en la llegada del autobús hacen que el sistema de transporte público no sea atractivo para los pasajeros.

Un Sistema de Información de Pasajeros en Tiempo Real utiliza diferentes tecnologías para el rastreo de las ubicaciones de los autobuses en tiempo real y utiliza dicha información para realizar predicciones de los arribos de los autobuses en sus diferentes paradas a lo largo de su recorrido. Cuando esta información se otorga a los pasajeros por medios cableados o inalámbricos, pueden utilizar su tiempo de manera eficiente y podrían llegar a la parada del autobús antes de que este arribe, o darían la opción de poder utilizar un medio de transporte alternativo si el autobús llegara a sufrir



algún retraso. Incluso pueden realizar planes de sus viajes mucho antes de que realmente lo realicen. Esto hará que el sistema de transporte público sea competitivo y amigable para los pasajeros [5].

### **2.1.3 Sistemas de información al viajero**

Los esfuerzos para aumentar el uso del transporte público en áreas urbanas, es parte de una estrategia para disminuir la congestión del tráfico y los problemas asociados, como la contaminación y la mala calidad del aire, han llevado a las autoridades de transporte de diversos tamaños a invertir en sistemas avanzados de información al viajero.

El objetivo de este tipo de sistemas es ir más allá de la información de horarios estáticos y dar a los pasajeros acceso a información en tiempo real acerca del estado de los servicios de autobuses para que puedan planificar mejor sus viajes. Las autoridades de transporte ven los beneficios obtenidos al implementar sistemas de información de llegada de autobuses en tiempo real, algunos de estos son: el mejor servicio ofrecido al cliente, mayor satisfacción y conveniencia del cliente y mayor visibilidad del tránsito en la comunidad.

Una de las percepciones entre los clientes es que los servicios de autobús mejoran y que las personas que viajan tarde en la noche ahora tienen la seguridad de que el próximo autobús no está muy lejos aumentando así la seguridad del viajero [11].

## **2.2 Geo localización**

La geolocalización es una de las herramientas utilizadas por los geógrafos para ubicar o situar a las personas o algunos otros objetos en el espacio mediante sus coordenadas geográficas (longitud, latitud y altitud) y que ha ocupado una nueva dimensión a partir de la aparición de Internet y de los dispositivos móviles [7]. Gracias a la geo localización se han desarrollado diferentes sistemas que han facilitado de manera significativa el rastreo de personas u objetos.

### **2.2.1 GPS**

El Sistema de Posicionamiento Global (GPS) es un servicio que proporciona a él o los usuarios información acerca del posicionamiento, navegación y cronometría. Los 24 satélites del GPS proporcionan servicios a los usuarios, este servicio es gratuito y está disponible para todas las personas de manera permanente y global.

Con la aparición del GPS en la red celular, el monitoreo de vehículos en tiempo real para la administración del transporte se ha hecho posible. Este tipo de tecnologías es posible aplicarlas a sistemas de transporte público, especialmente autobuses, que no pueden cumplir horarios predefinidos ya que se pueden presentar diferentes tipos de sucesos inesperados como son los atascos, averías, accidentes, etc [5].

### **2.2.2 Red de sensores inalámbricos**

Es una red inalámbrica que consta de dispositivos distribuidos espaciados autónomos que utilizan sensores para realizar monitoreo de condiciones físicas o ambientales. Una red de sensores inalámbricos son las cuales pueden lograr comunicación bidireccional y no requiere costo de comunicación. Una red de sensores inalámbricos consta de varios nodos de sensores pequeños que se instalan en cada autobús y en un área de servicio de autobuses, con este sistema la información de la

ubicación y operación de los autobuses puede ser transmitida sin costo alguno mediante la retransmisión de múltiples saltos en la red de sensores inalámbricos [6].

## **2.3 GTFS**

La especificación general de feeds de transporte público (GTFS) define un formato común para los horarios de transporte público y la información geográfica asociada a los datos de la ruta, como son las calles por las cuales circula una ruta y sus correspondientes paradas en la ruta. Los "feeds" GTFS permiten que las empresas de transporte público publiquen sus datos de transporte y que los programadores escriban aplicaciones que consuman esos datos de manera interoperable.

Los feeds GTFS se componen de una serie de archivos de texto. Cada archivo modela una característica en particular de información de tránsito: paradas, rutas, viajes y otros tipos de datos relacionados con los horarios. Las especificaciones de cada uno de los archivos se definen en la referencia de GTFS. Una empresa de transporte puede producir un feed GTFS para compartir su información de transporte público con programadores, quienes podrían desarrollar las herramientas que utilizan los feeds GTFS para añadir la información de transporte público en sus aplicaciones. La especificación GTFS puede ser utilizada para ayudar a los planificadores de viajes, editores de horarios y diversas aplicaciones, etc., a que usen la información de transporte público para hacer mejora de la movilidad en las ciudades.

Muchas aplicaciones son compatibles con los datos en formato de GTFS. El método más fácil para hacer que un feed sea público es almacenarlo en un servidor web y hacer la publicación un anuncio para que esté disponible para su uso. Si eres representante de una empresa de transporte público que controla el transporte público de tu ciudad, puedes usar la especificación GTFS para proporcionar horarios e información geográfica a Google Maps y a otras aplicaciones de Google para mostrar información de transporte público [3].

# Capítulo 3 Estado del arte

## Descripción

En esta sección se presenta el estado del arte de trabajos que realizan la predicción de los tiempos de arribo de autobuses, presentando diferentes tipos de algoritmos con los cuales realizan la predicción. Las categorías de clasificación de los diferentes tipos de algoritmos se presentan en la siguiente sección.

### 3.1 Clasificación del estado del arte

Las categorías utilizadas para la clasificación de los algoritmos de predicción de tiempos de arribo son las siguientes:

- **Algoritmos Heurísticos:** Son algoritmos que proponen uno o más objetivos, normalmente encuentran buenas soluciones, aunque no hay pruebas de que la solución no pueda ser arbitrariamente errónea en algunos casos; o se ejecuta razonablemente rápido, aunque no existe tampoco prueba de que siempre será así.
- **Algoritmos clasificadores:** Son algoritmos que, a partir de un conjunto de ejemplos clasificados (conjunto de entrenamiento) intentan asignar una clasificación a un segundo conjunto de ejemplos. Por otra parte, los sistemas de clasificación no supervisados son aquellos en los que no se dispone de una batería de ejemplos previamente clasificados, sino que realizan la predicción a partir de las propiedades de los ejemplos.
- **Sistemas que consideran GTFS:** La Especificación general de feeds de transporte público (GTFS) define un formato común para los horarios de transporte público y la información geográfica asociada a ellos. Los “feeds” GTFS permiten que las empresas de transporte público publiquen sus datos de transporte y que los programadores escriban aplicaciones que consuman esos datos de manera interoperable.
- **Sistemas que utilizan Meta modelos:** Los meta modelos son la fusión de algoritmos que ya existen para poder crear un nuevo algoritmo.

### 3.2 Aspectos descriptivos de los trabajos del estado del arte de predicción de tiempos de arribo del transporte público

Los aspectos utilizados para describir y detallar cada algoritmo analizado son los siguientes:

**Descripción general del trabajo:** Describe el trabajo de investigación tomando en cuenta algunos aspectos importantes como: características principales y técnicas utilizadas.

**Objetivo:** Describe los principales motivos que dieron iniciativa a desarrollar o proponer el trabajo mencionado en los trabajos de investigación.

**Conclusiones:** Describe las conclusiones y resultados obtenidos al momento de concluir con el trabajo de investigación.

### 3.3 Algoritmos heurísticos

Las heurísticas generalmente son usadas cuando no existe una solución óptima bajo las restricciones dadas (tiempo, espacio, etc.).

### **3.3.1 Sistema de ubicación en tiempo real para el servicio de autobuses comunitarios**

Este trabajo de investigación [10] presenta el desarrollo de un sistema que tiene como objetivo predecir el tiempo de arribo del transporte público. Hace mención sobre la metodología que fue utilizada para el desarrollo de la herramienta y sobre como el sistema será implementado en una aplicación Mobile para el Sistema Operativo Android.

El sistema desarrollado para calcular el tiempo de arribo de las rutas cuenta con dos fases (actualizador de enlaces y tiempo estimado de llegada) para realizar obtener el tiempo estimado de arribo. Las dos fases mencionadas describen cómo se utilizan y cómo es que se realiza la predicción haciendo uso de ellas, ya que permite registrar las paradas que tiene el autobús proporcionando su longitud y latitud, las cuáles serán almacenadas en una base de datos, calcula la distancia que existe entre paradas, estos son parámetros utilizados para que el algoritmo pueda trabajar.

El sistema calcula el tiempo de arribo gracias a que el autobús se encuentra enviando su ubicación GPS actual en tiempo real. Con esto se permite obtener un cálculo del tiempo estimado de arribo mediante la posición de la ruta y con la posición a la que se encuentra su siguiente parada. Estos factores se promedian con el tiempo estimado de espera (el tiempo que tarda la ruta anterior en recorrer esa misma distancia), esta variable se va actualizando gracias a que se obtiene el tiempo que tarda cada una de las rutas, esto para poder obtener un estimado aproximado de cómo está el tráfico o si ocurrió algún otro suceso que retrase los tiempos de arribo.

Este trabajo de investigación tiene como conclusión que el utilizar GTFS provee una ventaja para desarrollar este tipo de sistemas, en los cuales se quiera obtener estimaciones del tiempo de arribo de los autobuses del transporte público, por lo cual seguirán trabajando en las siguientes fases de su trabajo de investigación.

### **3.3.2 Sistema inteligente de toma de decisiones para pasajeros de transportes urbanos**

Este trabajo de investigación [11] presenta la implementación de un sistema de predicción de tiempos de arribo del transporte público (autobuses) y el desarrollo de un contador de pasajeros del transporte público. Del cual las personas podrán obtener estos datos del transporte mediante un sistema de software desarrollado para el Sistema Operativo Android.

Para desarrollar el algoritmo se registran las paradas que tiene el camino que recorre el autobús otorgando la ubicación (longitud y latitud), esta información es almacenada en una base de datos que se encuentra ubicada en un servidor web, una vez registradas las paradas se utilizara el GPS del Smartphone del conductor para proporcionar su ubicación en tiempo real. El sistema calcula la hora de arribo con el lapso de tiempo que tarda un autobús en pasar de una parada a otra, el cual se almacena en el servidor y es el tiempo estimado que tarda el siguiente autobús en recorrer esa misma distancia. El sistema también calcula el tiempo que tarda en recorrer el camino los siguientes autobuses de una parada a otra, la cual se actualiza en el servidor y es el tiempo estimado del autobús que sigue.

El contador de pasajeros está implementado con dos sensores ultrasónicos y una placa de Arduino. El método utilizado tiene una lógica interesante, ya que funciona de la siguiente manera: si la persona pasa por el sensor uno y después por el sensor dos, esto significa que se encuentra abordando la ruta, por lo contrario, si pasa por el sensor dos primero y después por el sensor uno significará que la persona está descendiendo de ella. Estos datos son almacenados en un servidor y se van actualizando cada que los sensores perciben una nueva acción realizada por las personas.

Las conclusiones obtenidas por este trabajo de investigación son muy informativas, ya que proporciona información acerca de cómo funciona el sistema, que aspectos se toman en cuenta para el desarrollo del algoritmo y su funcionamiento, muestra resultados acerca del margen de error con el que se cuenta la herramienta.

### **3.3.3 Implementación de un sistema de monitoreo e información de pasajeros del autobús en tiempo real**

Este trabajo de investigación [5] presenta el desarrollo de un sistema que predice el tiempo de arribo del transporte público. La metodología utilizada y sobre como el sistema será desarrollado creando una aplicación Mobile para el Sistema Operativo Android.

El sistema desarrollado cuenta con dos fases (actualizador de enlaces y tiempo estimado de llegada) para realizar obtener el tiempo estimado de arribo. Fácil de utilizar y realizar la predicción ya que permite registrar las paradas que tiene el autobús proporcionando su longitud y latitud, las cuáles serán almacenadas en una base de datos, calcula la distancia que existe entre paradas, estos son parámetros utilizados para que el algoritmo pueda trabajar. El sistema calcula el tiempo de arribo gracias a que el autobús se encuentra enviando su ubicación GPS en tiempo real. Con esto se permite obtener un cálculo del tiempo estimado de llegada mediante la posición de la ruta y con la posición a la que se encuentra su siguiente parada. Estos factores se promedian con el tiempo estimado de espera (el tiempo que tardo la ruta anterior en recorrer esa misma distancia), esta variable se va actualizando gracias a cada una de las rutas y cada que una ruta recorre el camino de una parada a la siguiente, esto para poder obtener un estimado aproximado de cómo está el tráfico o si ocurrió algún otro suceso que retrase los tiempos de arribo.

Este trabajo de investigación propone una solución viable a una de las necesidades que los pasajeros, conductores y administradores de vehículos, utilizando sistema GPS de dispositivos móviles y la red celular para obtener la ubicación de los autobuses.

## **3.4 Algoritmos clasificadores**

Estas técnicas son utilizadas normalmente para el reconocimiento de patrones. Mediante estas técnicas, es posible entrenar un clasificador para que aprenda autónomamente la forma de realizar una buena clasificación de los objetos.

### **3.4.1 Predicción del tiempo de viaje del autobús utilizando una red neuronal artificial**

Este trabajo de investigación [12] muestra cómo utilizar una red neuronal artificial para poder realizar predicciones del tiempo de arribo de los autobuses. Aborda una pequeña descripción de métodos que han sido utilizados para realizar esta tarea y los compara con el propuesto. Su enfoque principal es la

comparación con el método de regresión, ya que los autores creen que ese es el método más preciso para realizar predicciones de tiempos de arribo.

El método de Red Neuronal Artificial trabaja en dos fases: aprendizaje y recordar. En la primera etapa es en donde se le darán los parámetros con los cuales se trabajará la red neuronal artificial, y en la segunda etapa es en donde se realiza todo el proceso de cálculo con los parámetros dados en la primera etapa. Los parámetros que toma en cuenta el método son: distancia entre paradas, tiempo de retrasos, tiempos de espera del autobús en una parada y tiempo de llegada a la parada, ya que son los parámetros que más influyen al momento de realizar la predicción con este método.

En este trabajo de investigación se realizaron diferentes pruebas, cada una de ellas se entrenó con diferentes conjuntos de datos y se pudo comprobar que el método trabaja mejor que el método de regresión, por lo cual el objetivo se cumplió. Este trabajo tiene buenas conclusiones una de ellas es que permite identificar qué factores son los que más influyen al momento de tratar de realizar una predicción de tiempo de arribo de los autobuses. De igual manera se sabe qué factores no influyen y algunos métodos que se pueden ocupar (Modelo Basado en Historial, Modelo de Regresión, Modelos de filtración de Kalman).

### **3.4.2 Predicción probabilística del avance del autobús utilizando la regresión de máquina de vectores de relevancia**

En este trabajo de investigación [13] se propone un algoritmo basado en máquinas de vectores de relevancia para predecir los avances del autobús al incorporar las series de tiempo de los avances del autobús, el tiempo de viaje y la demanda que tuvo de los pasajeros en las paradas anteriores. Este método fue utilizado por que puede generar el resultado de predicción probabilística.

El trabajo de investigación menciona algunos factores que afectan la predicción de los tiempos de arribo, tales como, el avance y la velocidad del autobús, la demanda de los pasajeros, el comportamiento del conductor del autobús, condiciones climáticas entre otros más. Este método propuesto asigna una probabilidad a cada avance de autobús con las condiciones que pueden afectar. La innovación de este trabajo de investigación recae en tres puntos principales, el primero de ellos es que puede predecir probabilísticamente el avance del autobús con el intervalo de confianza, el segundo es que este trabajo utilizó datos de tarjetas inteligentes para extraer la demanda de los pasajeros en las paradas para la predicción del avance del autobús, y el tercero es que se puede capturar la fluctuación del avance en el nivel de parada utilizando el método propuesto para la mitigación de la acumulación de autobuses y la estimación de tiempos de espera de los pasajeros.

Este algoritmo fue comparado con cinco algoritmos y técnicas que han sido utilizadas para desarrollar algoritmos como este las cuales son: máquinas de vector soporte, algoritmo genérico de máquina de vector soporte, KNN (vecino más cercano) y redes neuronales artificiales. El resultado obtenido en este trabajo de investigación, indica que el algoritmo de máquinas de vector de relevancia supera en términos de precisión e intervalos de confianza a los cinco algoritmos con los que fue comparado. Ya que cuando se establece el nivel de confianza en 95% más del 95% de los autobuses se encuentran dentro de los términos de la predicción realizada.

### **3.4.3 Modelo de predicción del tiempo de viaje de la red neuronal artificial para autobuses que utilizan solo datos de GPS**

Este trabajo de investigación [14] propone un modelo basado en las redes neuronales artificiales para predecir de forma precisa el tiempo de llegada de los autobuses, usando solo la información del sistema de posicionamiento global (GPS) para proporcionar información en tiempo real en una parada determinada.

Para el desarrollo de este algoritmo los investigadores hicieron uso de un sensor GPS para obtener la ubicación actual del autobús, y así poder considerar explícitamente la información de la hora de llegada y salida de los autobuses en sus diferentes paradas. El desarrollo de este modelo fue basado en los patrones históricos de hora de llegada y salida y en la información de hora de llegada y salida en tiempo real. Una de las ventajas que tiene este modelo es que fue desarrollado con una red neuronal artificial, que en comparación con los modelos que más se utilizan para hacer predicción este tiene una ventaja importante, es que puede ser entrenada tomando en cuenta otros factores que los demás modelos no toman en cuenta, como, por ejemplo, la velocidad, el clima, la distancia, entre otras más como datos de entradas. Para realizar este modelo se tomaron como base, otros cuatro modelos que son los más utilizados al momento de realizar predicción de tiempos de arribo, los cuales son los algoritmos de filtrado de Kalman, modelos de regresión, modelos de máquinas de aprendizaje y los modelos históricos promedio.

Este modelo fue comparado con un modelo promedio histórico según dos criterios: precisión de predicción y robustez, y en el estudio realizado se demostró que la red neuronal artificial superó el enfoque promedio en ambos aspectos, por lo cual obtuvo mejores resultados que el modelo promedio histórico.

## **3.5 Sistemas que consideran GTFS**

Este tipo de formato es utilizado para que las empresas de transporte público puedan establecer un sistema de estandarización de horario y para poder compartir su información con los usuarios del transporte público.

### **3.5.1 Brindando servicios de información en tiempo real sobre el transporte público**

Este trabajo de investigación [15] presenta el desarrollo de un Framework nombrado MOBANA. Este implementa el sistema GTFS y GTFS-RealTime entre otros sistemas y técnicas para poder calcular el tiempo de arribo del transporte público. MOBANA permite visualizar un mapa de la ciudad, paradas y rutas, posición de los vehículos en tiempo real, la posición programada de los vehículos en tiempo real, eventos de tráfico y feeds.

MOBANA utiliza el sistema GTFS para realizar casi todos los cálculos, pero agrega un sistema de mensajería distribuida, con el cual mediante los tweets publicados en diferentes zonas se logra calcular y observar el flujo de tráfico que existe en los diferentes caminos, mediante la posición GPS o mediante Hashtags. Desarrollaron un algoritmo para poder procesar esta información y así utilizar esta metodología. Crearon otro método para poder predecir los retardos que los autobuses presenten ya sea que se encuentre en medio de un accidente o algún otro suceso que retrase su llegada, y el cálculo del tiempo que tardara está en llegar a su destino para procesar esta información con el



sistema GTFS y poder mostrarla al usuario mediante un Framework que fue montado en una página web.

Este trabajo de investigación muestra cómo se puede implementar el sistema GTFS y GTFS-RealTime. Así como los diferentes usos que se le puede dar y que información podemos obtener de este. Menciona una comparación de diferentes sistemas que ofrecen este tipo de servicios la cual describe sus ventajas y desventajas. Presenta como se puede implementar una fusión de los sistemas para obtener datos y poder realizar el objetivo de manera más eficiente. La manera en la que se puede implementar el servidor para almacenar y mostrar los datos obtenidos.

### **3.5.2 Visualización de datos de especificaciones de alimentación de tránsito general**

Este trabajo de investigación [16] muestra cómo utilizar el sistema GTFS y como se adapta a un sistema. En este trabajo de investigación fue adaptado a una página web en la cual podrán ingresar todas las personas y se muestra algunos ejemplos de cómo la visualizan los usuarios. Nos muestran algunas ideas para combatir con algunas de las limitaciones con las que cuenta el sistema GTFS, como es el caso de número máximo de paradas que se permite registrar y entre otros.

El sistema web está basado en el sistema GTFS, pero para el desarrollo de este utilizaron principalmente cuatro tablas de la gran variedad que tiene el sistema. Las cuales son: Route, Stop, Trip y StopTime. En donde están registrados los datos de la ruta, de la parada y del viaje, ya que la tabla StopTime contiene llaves foráneas de las otras 3 tablas.

Uno de los problemas que se comenta es que existen errores al obtener la información del sistema, ya que al obtener los datos se deben analizar, procesar y corregir para posteriormente almacenarlo en la base de datos. Otro error abordado es la limitante de captura de paradas, puesto que el sistema solo permite registrar 8 paradas por ruta más su inicio y su fin dando como resultado 10 registros por ruta en total. Este trabajo aborda la solución que le dieron a estos problemas. Nos muestra una visualización de su sistema y una pequeña explicación de cómo funciona.

Este trabajo de investigación obtuvo conclusiones que ayudan a los investigadores, ya que muestra cómo se puede adaptar este sistema al software en desarrollo, da una muestra de cómo obtener los datos y solucionar los errores y limitaciones que tiene la versión gratuita del GTFS.

### **3.5.3 Los múltiples usos de los datos GTFS – abriendo la puerta al tránsito y a las aplicaciones multimodales**

Este trabajo de investigación [17] describe todos los beneficios y usos que se pueden dar a sistemas que utilicen GTFS, ya que las agencias de transporte público pueden proporcionar muchos nuevos tipos de servicios de información al público o para las operaciones internas de la empresa, con poco o ningún costo para la agencia.

En este trabajo de investigación se mencionan algunos aspectos que dieron pie a que se creara GTFS, también menciona algunas de las agencias que se encuentran trabajando con este sistema y como les ha ayudado en su crecimiento y como es que han sido aprovechados todos los recursos que ofrece. También se mencionan algunos riesgos con los que se pueden enfrentar al momento de utilizar el GTFS y cómo enfrentarse a ellos. Se mencionan también algunos trabajos que han sido

desarrollados utilizando el GTFS y cómo es que han funcionado. Otro aspecto mencionado es como se puede dar de alta la información en el GTFS, ya que se necesitan llenar una serie de archivos .txt proporcionando información acerca del transporte público, así como información de las paradas de los autobuses, información acerca del autobús, de la ruta que siguen los autobuses entre otras más.

La conclusión obtenida por los autores es que GTFS se ha convertido en un estándar para la representación de datos de tránsito programados, gracias a esto está siendo utilizado por muchos tipos de diferentes aplicaciones de las agencias de transporte público. Con este estándar cualquier departamento de transporte puede aprovechar la información que está siendo publicada para ayudar a crear verdaderos sistemas de información de viajeros multimodales.

### **3.6 Sistemas que utilizan Metamodelos**

Esta técnica se utiliza para poder mejorar la solución a un problema aumentando la eficiencia o crear una nueva solución para un problema.

#### **3.6.1 Incertidumbre en las predicciones del tiempo de arribo del autobús: tratamiento de la heterocedasticidad con un enfoque de metamodelo**

En este trabajo de investigación [18] se describe un metamodelo desarrollado con la fusión de algoritmos de predicción de tiempos de arribo que ya existen. El propósito principal de este metamodelo es el poder utilizar las diferentes ventajas que tienen los algoritmos ya existentes, y capturar la incertidumbre asociada al realizar predicciones ya que según los autores esta área no se analiza adecuadamente al momento de realizar estas predicciones.

Este trabajo de investigación fue creado para que los usuarios del transporte público puedan obtener información del transporte público en tiempo real, y con esto poder hacer un mejor plan de viaje, con la herramienta las personas se sentirán más seguras principalmente las que viajan de noche. Otra de las causas de estos sistemas es que beneficia a las autoridades del transporte, ya que ellos son los que obtienen mayor beneficio de esto puesto que incrementa la satisfacción de los viajeros, proveen mejor servicio a la comunidad y otros beneficios más.

Los autores mencionan que el problema al momento de realizar predicciones es complicado, ya que los tiempos de viaje de los autobuses son resultados de interacciones no lineales y complejas de varios factores que llegan a influir en la demanda o capacidad del autobús como pueden ser el tráfico, la demanda de los pasajeros, condiciones climáticas entre otros más. Esto se menciona como un gran problema ya que los autores mencionan que la mayoría de las veces estos problemas son ignorados o tomados como constantes y esto es un grave error ya que provoca que los algoritmos no funcionen de manera correcta. El enfoque alternativo utilizado en este trabajo de investigación es el uso de intervalos de predicción, adaptando un enfoque diferente para construirlos, usando regresión por cuantiles, que toman en cuenta la incertidumbre en la estructura del modelo como el ruido en los datos de entrada, utilizando un modelo más sofisticado y aplicándolo a un servicio en tiempo real.

Este trabajo de investigación obtuvo varios resultados, uno de ellos es el desarrollo un enfoque para manejar adecuadamente la heterocedasticidad en las predicciones mediante el uso de la regresión por cuantiles. Utilizando este método se proporciona la cobertura deseada. Utilizando los intervalos de predicción en cortas distancias y en largas distancias fueron obtenidos buenos resultados en comparación con otros trabajos.

### **3.6.2 Predicción de la hora de llegada del autobús basada en un modelo mixto**

En este trabajo de investigación [19] se propone un método para la predicción de tiempos de llegada de los autobuses. El propósito principal de este método es lograr una mejor predicción que los métodos ya existentes, para lograr este propósito, el método propuesto es una fusión de varios métodos que ya existen, cada uno de estos métodos es utilizado para llevar a cabo diferentes tareas dentro del método propuesto. Los métodos utilizados para crear el nuevo método son: KNN (vecino más cercano), K-means, Modelos de filtración de Kalman y el modelo de transferencia histórica de Markov.

El método creado consiste de tres pasos. El primer paso es el entrenamiento de patrones, en este paso es utilizado el método de KNN y K-means para obtener los patrones del flujo de tráfico. La segunda etapa es una predicción de un solo paso, el cual está basado en el filtro de Kalman el cual fue ajustado en tiempo real con la modificación del histórico modelo de patrones de fluctuación de tráfico. En la tercera etapa utilizan la influencia de la ley histórica en la predicción a larga distancia, y combinan la predicción de un sólo paso con el modelo de transferencia histórica de Markov, esto lo hacen para llevar a cabo la predicción de varios pasos.

Los resultados que se obtuvieron muestran que la predicción de un solo paso funciona mejor en precisión y eficacia que la predicción de flujo de tráfico a corto plazo y que el filtro de Kalman. La predicción de varios pasos proporciona mayor veracidad y confiabilidad al realizar pronósticos de tiempos de viaje comparado con los modelos de flujo de tráfico a corto plazo y con los de patrón de tráfico histórico. Otro de los resultados obtenidos es que la propuesta de este trabajo de investigación puede adaptarse o enfrentarse rápidamente a condiciones anormales que ocurran en el viaje como un accidente de tráfico y obtener una predicción de tiempo de arribo del autobús más precisa. Uno de los problemas que se presenta con el método propuesto es el uso de la memoria, ya que los métodos de KNN y el modelo de transferencia histórica de Markov requieren una gran cantidad de memoria.

# Capítulo 4 Metodología de solución

En este capítulo se describe la metodología de solución propuesta para la creación del sistema de software con el cual se realiza la predicción de los tiempos de arribo de autobuses de transporte público. Como primer punto se describe de manera general la metodología de solución y posteriormente se describe de manera detallada cada una de las fases que componen la metodología.

## **4.1 Descripción general de la metodología de solución**

En este proyecto de investigación se desarrolló un sistema de software que tiene la capacidad de predecir los tiempos de arribo de los autobuses a las correspondientes paradas de una ruta de transporte público. Se desarrolló también un prototipo de un sistema de rastreo. El sistema utiliza una placa LinkIt One y un chip de telefonía móvil de la compañía Telcel como apoyo para poder recopilar la información de contexto del transporte público.

La solución propuesta pretende mejorar la eficiencia del transporte público, como componentes de hardware se utilizaron recursos tecnológicos no tan sofisticados con el fin de facilitar su implementación en una ciudad. La metodología propuesta consta de 4 fases (Figura 4.1).

### **Fase 1: Recopilación de información sobre la ruta, paradas y horarios y su representación en formato GTFS**

El objetivo de esta fase fue obtener información sobre la ruta de transporte público, la localización de cada una de las paradas que componen la ruta y los horarios en los que cada autobús debe llegar a cada una de las paradas. Esta información se representa siguiendo el estándar de la especificación GTFS, que ha demostrado ser una excelente herramienta para obtener una mejor predicción de tiempos de arribo de los autobuses.

Otra tarea que se realizó en esta fase fue crear la base de datos en la cual se almacena la información recopilada, también se depuro la base de datos GTFS y se creó una nueva base de datos solo con la información que se consideró necesaria para la realización de este trabajo de investigación.

### **Fase 2: Desarrollo del algoritmo**

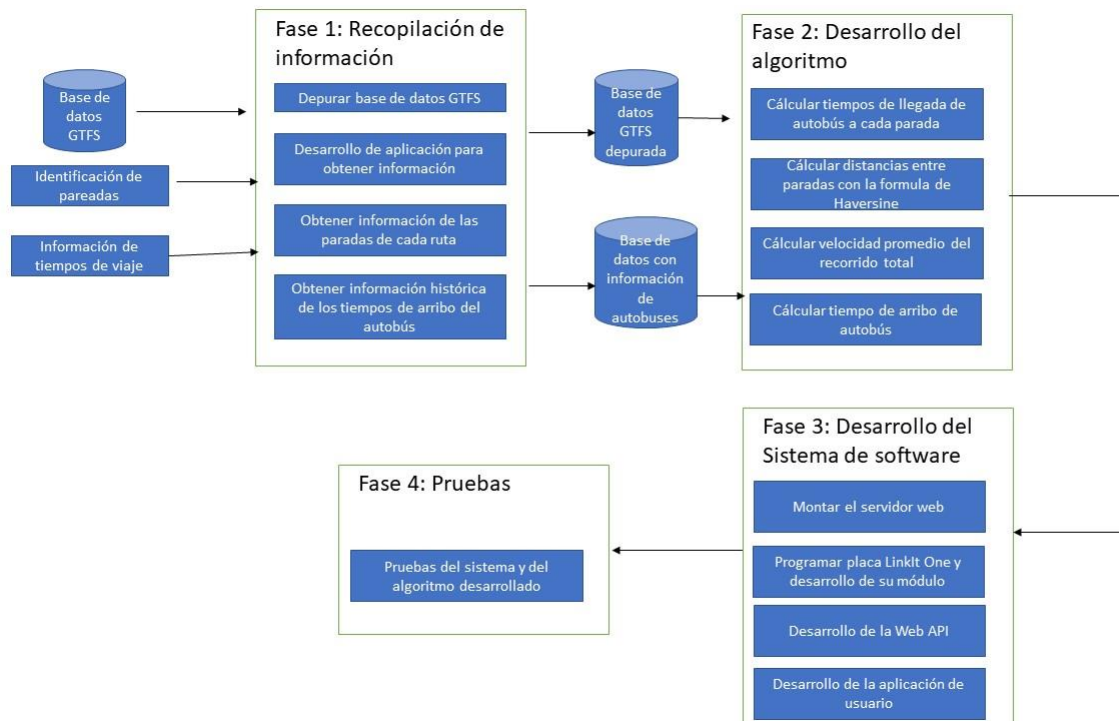
En esta fase se desarrolló el algoritmo encargado de procesar la información de contexto y realizar la predicción de los tiempos de arribo de los autobuses de transporte público.

### **Fase 3: Desarrollo de sistema de software**

En esta fase se realizó el desarrollo del sistema de software que implementa los algoritmos propuestos en la tesis.

### **Fase 4: Pruebas del sistema**

En esta fase se realizan experimentación de la solución propuesta, realizando viajes y registrando los tiempos que fueron predecidos por el sistema.



**Figura 4. 1 Metodología de solución.**

## 4.2 Fase 1: Recopilación de información

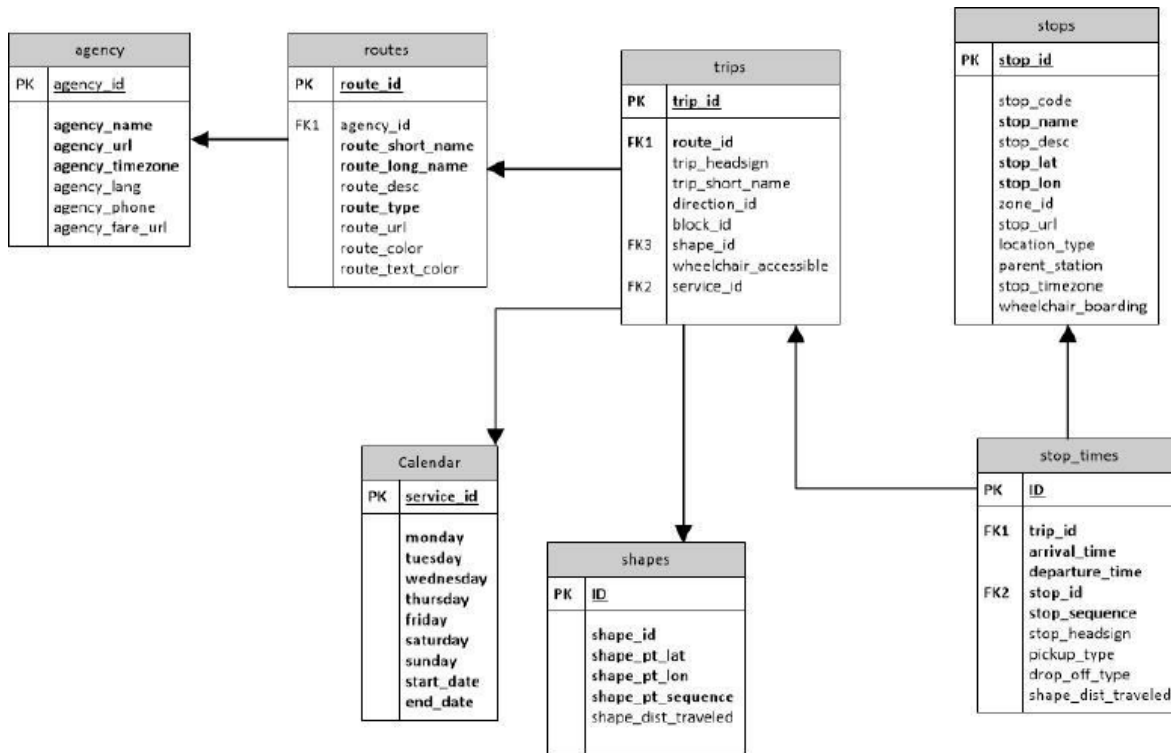
Esta fase de la metodología se divide en cuatro pasos que se describen a continuación.

El primer paso de la metodología propuesta fue la representación de una ruta y sus componentes utilizando la especificación GTFS. Es importante hacer notar que en el caso de ciudades grandes de nuestro país las rutas de transporte público ya son representadas utilizando GTFS. Sin embargo, en ciudades pequeñas no se cuenta con esta información, por lo que en esta tesis se parte del hecho que no existe una representación de rutas en formato GTFS, por lo que la primera fase de nuestro método será la creación de las rutas y sus componentes en GTFS.

El primer paso es obtener la base de datos que contiene la representación de la especificación GTFS. Esta base de datos se obtuvo de la página oficial de Google <https://developers.google.com/transit/gtfs/examples/gtfs-feed>. La base de datos del estándar GTFS está compuesta de 11 tablas las cuales son: agency, calendar, calendar\_dates, fare\_attributes, fare\_rules, frequencies, routes, shapes, stop\_times, stops y trips. Como puede observarse, la base de datos GTFS está compuesta principalmente por los datos de las agencias de transporte público, los datos de las diferentes rutas que circulan por una ciudad, los diferentes considerados en una ruta, las paradas correspondientes a una ruta, los horarios en los cuales los autobuses debe llegar a cada parada.

La figura 4.2 se puede observar la información GTFS obtenida de la página oficial de Google <https://developers.google.com/transit/gtfs/examples/gtfs-feed>.

La base de datos GTFS fue analizada y depurada para identificar solo aquellos archivos y datos que pueden adaptarse al desarrollo del sistema de predicción propuesto.



**Figura 4. 2 Información GTFS**

Una de las tablas de la base de datos GTFS representa las paradas (stops). La figura 4.3 muestra un ejemplo de la tabla stops, la cual está formada por los campos stop\_id que representa el identificador único de la parada, stop\_code que representa el código de la parada, stop\_name que representa el nombre de la parada, stop\_lat que representa la latitud de la parada y stop\_lon que es la longitud de la parada.

stop_id	stop_code	stop_name	stop_lat	stop_lon
86	86	50th St @ Palm River Rd	27.943531	-82.401742
403	403	Nebraska Av @ Twiggs St	27.951797	-82.450892
404	404	Neb @ Cass Mr N-S Sta 3	27.953998	-82.451237
405	405	Nebraska Av @ Maryland Av	27.95455	-82.451146
407	407	Neb @ Estelle Mr N-S Sta 5	27.958445	-82.451237
409	409	Nebraska Av @ 7th Av	27.960157	-82.451116
413	413	Nebraska Av @ 15th Av	27.965799	-82.451097
414	414	Nebraska Av @ 19th Av	27.968736	-82.451178
418	418	Nebraska Av @ 26th Av	27.974132	-82.45115
420	420	Nebraska Av @ Lake Av	27.97769	-82.451129
422	422	Nebraska @ MLK (@ 33Rd Av)	27.98077	-82.450987
424	424	Nebraska Av @ Chelsea St	27.984984	-82.451094
427	427	Nebraska Av @ New Orleans Av	27.990083	-82.451097
429	429	Nebraska Av @ McBerry St	27.992712	-82.451102
430	430	Nebraska Av @ Frierson Av	27.994217	-82.451098
431	431	Nebraska Av @ Hillsborough Av	27.995641	-82.451104
433	433	Nebraska Av @ Powhattan Av	27.998681	-82.451168
435	435	Nebraska Av @ Henry Av	27.999928	-82.451193
436	436	Nebraska Av @ Idelwild Av	28.001809	-82.451163
438	438	Nebraska Av @ Fern St	28.004934	-82.451267
439	439	Nebraska Av @ Lambright St	28.006811	-82.451266
440	440	Nebraska Av @ Knollwood St	28.008152	-82.451263

Figura 4. 3 Ejemplo de información en tabla stops

La depuración de la base de datos se realizó por que en este trabajo de investigación no es necesario utilizar toda la información del sistema GTFS. La depuración de la base de datos se realizó con las tablas agency, routes, stops, trips y stop\_times quedando la base de datos como se muestra en la figura 4.4.

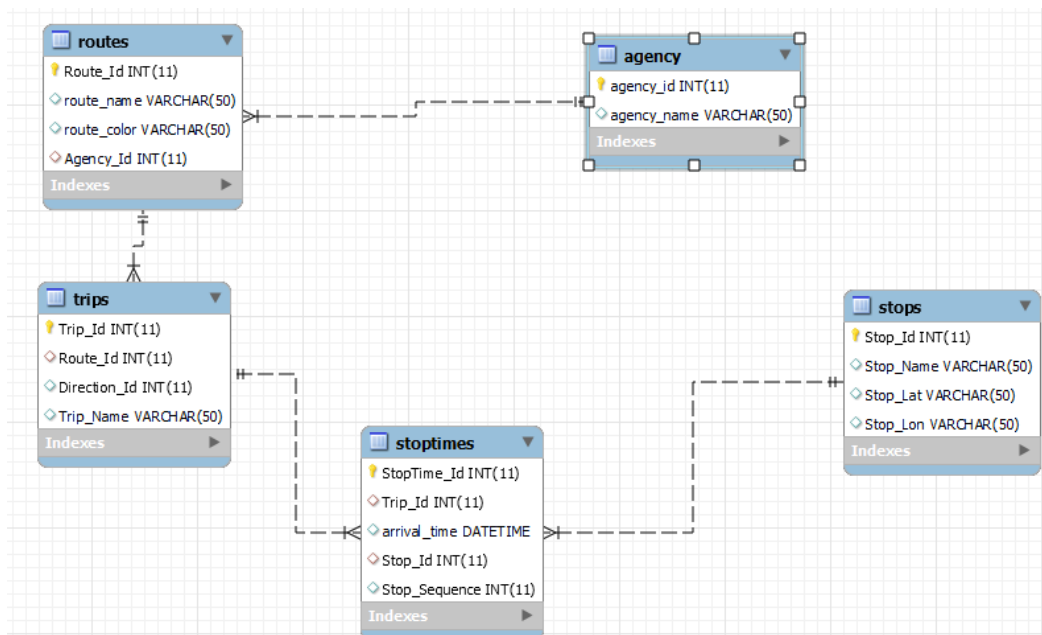
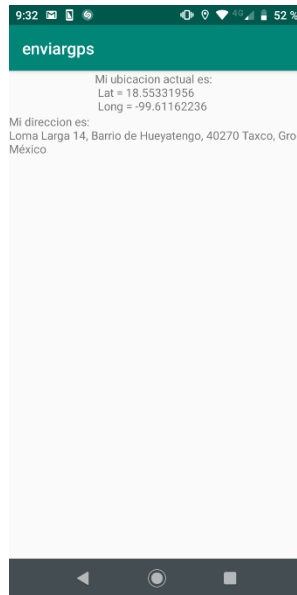


Figura 4. 4 Modelo entidad relación de la base de datos depurada

El segundo paso en esta fase consistió en desarrollar una aplicación móvil que permita identificar las paradas de las rutas consideradas en este trabajo de investigación. La aplicación móvil fue



desarrollada para el sistema operativo Android. En la figura 4.5 se puede observar la interfaz de la aplicación desarrollada.



**Figura 4. 5 Aplicación que obtiene longitud y latitud de las paradas**

Es este paso se obtienen las coordenadas de cada una de las paradas de las rutas consideradas para su análisis en esta tesis. Para llevar a cabo este paso se tuvo que hacer el recorrido por cada una de las paradas de una ruta obteniendo sus coordenadas geográficas (latitud, longitud). La figura 4.6 muestra la tabla con la información de las paradas de las rutas.

```
MariaDB [andybus]> select * from stops;
```

Stop_Id	Stop_Name	Stop_Lat	Stop_Lon
1	Avenida Universidad 153-905, Universidad, Cuernava	18.978498	-99.245266
2	Avenida Universidad 430-7, Santa María Ahuacatitl?	18.973893	-99.246213
3	Don Miguel Hidalgo 15, Lienzo del Charro	18.969117	-99.244813
4	Avenida Universidad 62130, Lienzo del Charro	18.966776	-99.244186
5	México 95 600, Bellavista	18.965786	-99.247064
6	Av. Emiliano Zapata, Bellavista	18.961289	-99.246290
7	El Globo Av.Emiliano Zapata, Bellavista	18.956933	-99.245710
8	Banorte Av. Emiliano Zapata, Bellavista	18.954390	-99.245341
9	La Pradera	18.939488	-99.241364
10	Alvaro Obregon 807-801	18.937544	-99.240883
11	Guacamayas	18.875329	-99.219914
12	Av. Estado de Puebla s/n	18.876969	-99.222654
13	Av. Estado de Puebla 173	18.873510	-99.223358
14	Av. Estado de Puebla 116	18.876284	-99.226658
15	Av. Estado de Puebla 95	18.881705	-99.228063
16	Av. Estado de Puebla 9	18.886716	-99.228580
17	Av. Jose Maria Morelos y Pavon Nte. 87-81	18.894904	-99.228945
18	Av. Jose Maria Morelos	18.896639	-99.229815
19	Av. Jose Maria Morelos y Pavon Nte. 80-86	18.902485	-99.231508
20	Las Palmas	18.905924	-99.231900
21	Jose Ma. Morelos y Pavon 810	18.908106	-99.232837

```
21 rows in set (0.004 sec)
```

Figura 4. 6 Información de las paradas de las rutas

El tercer paso consiste en obtener la información histórica de los tiempos de arribo de los autobuses seleccionados, para generar la base de datos históricas de los recorridos fue necesario realizar 1 viaje por ruta, una de las rutas contaba con 11 paradas y la otra con 10. Una vez obtenida esta información obtenemos como salida nuestra base de datos con los tiempos históricos como se muestra en la figura 4.7.

```
MariaDB [andybus]> select * from stoptimes;
```

StopTime_Id	Trip_Id	arrival_time	Stop_Id	Stop_Sequence
1	1	2021-03-15 12:00:00	1	1
2	1	2021-03-15 12:02:23	2	2
3	1	2021-03-15 12:04:11	3	3
4	1	2021-03-15 12:06:26	4	4
5	1	2021-03-15 12:09:50	5	5
6	1	2021-03-15 12:10:49	6	6
7	1	2021-03-15 12:12:54	7	7
8	1	2021-03-15 12:14:05	8	8
9	1	2021-03-15 12:18:03	9	9
10	1	2021-03-15 12:18:43	10	10
11	2	2021-05-29 12:00:00	11	1
12	2	2021-05-29 12:01:18	12	2
13	2	2021-05-29 12:02:40	13	3
14	2	2021-05-29 12:05:06	14	4
15	2	2021-05-29 12:07:06	15	5
16	2	2021-05-29 12:09:10	16	6
17	2	2021-05-29 12:11:05	17	7
18	2	2021-05-29 12:13:00	18	8
19	2	2021-05-29 12:14:21	19	9
20	2	2021-05-29 12:15:08	20	10
21	2	2021-05-29 12:16:19	21	11

```
21 rows in set (0.024 sec)
```

Figura 4. 7 Información histórica obtenida

Una vez concluida esta primera fase se cuenta con una ruta de transporte público representada con el estándar de representación GTFS. Es importante comentar que esta representación de las rutas en GTFS podrá ser utilizada en otras tesis que necesiten utilizar rutas de transporte.

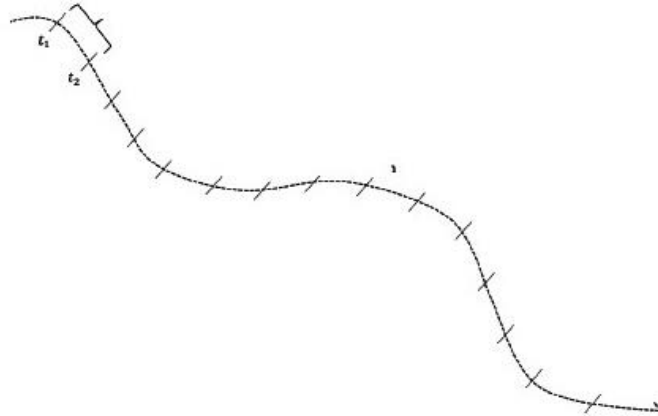
### 4.3 Fase 2: Desarrollo del algoritmo

La segunda fase de nuestra metodología consistió en desarrollar el algoritmo de predicción de tiempos de arribo. El algoritmo se desarrolló tomando como entrada la información de la base de datos GTFS depurada, de igual manera también se toma como dato de entrada la información de nuestra base de datos histórica. La información es procesada de la siguiente manera:

A partir de la información de las paradas (latitud, longitud) y los registros de los tiempos de arribo de los autobuses que fueron recolectados (tabla 4.1), se crean los vectores o también conocidos como aristas correspondientes a el recorrido que realiza el autobús, un vector es el recorrido que realiza el autobús de una parada a otra, es decir, de la parada 1 a la parada 2 es un vector, de la parada 2 a la parada 3 es otro vector y así sucesivamente. Para la creación de los vectores existe la condición de que la ruta del autobús debe de tener mínimo dos paradas ya que con una parada no se puede formar un vector, otra de las reglas para obtener el total de los vectores es que el número de vectores que existirán a lo largo de una ruta es igual al número de paradas -1. En la figura 4.8 se muestra un ejemplo de cómo se forman los vectores a lo largo del recorrido.

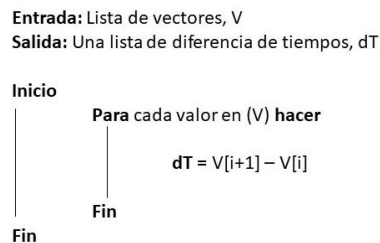
Paradas Ruta 1 De Av. Universidad rumbo a Guacamayas				
No. de Parada	Nombre de parada	Latitud	Longitud	Tiempo historico
1	Avenida Universidad 153-905, Universidad	18.978498	-99.245266	00:00
2	Avenida Universidad 430-7, Santa Maria Ahuacatitlan	18.973893	-99.246213	02:23
3	Don Miguel Hidalgo 15, Lienzo del Charro	18.969117	-99.244813	04:11
4	Avenida Universidad 62130, Lienzo del Charro	18.966776	-99.244186	06:26
5	Mexico 95 600, Bellavista	18.965786	-99.247064	09:50
6	Av. Emiliano Zapata, Bellavista	18.961289	-99.246290	10:49
7	El Globo, Av.Emiliano Zapata, Bellavista	18.956933	-99.245710	12:54
8	Banorte, Av. Emiliano Zapata, Bellavista	18.954390	-99.245341	14:05
9	La Pradera	18.939488	-99.241364	18:03
10	Alvaro Obregon 807-801	18.937544	-99.240883	18:43

**Tabla 4. 1 Información de las bases de datos**



**Figura 4. 8 Creación de vectores en el camino**

Una vez creados los vectores se realiza el cálculo de tiempos de recorrido del autobús, tomando los tiempos de llegada de los autobuses a sus paradas, ya con esta información se realiza una sustracción de tiempos para saber cuánto tarda en recorrer cada uno de los vectores formados.



**Figura 4. 9 Paso 1: Desarrollo del algoritmo**

Sustituyendo los datos en la figura 4.9  $V[i]$  representa el tiempo de llegada del autobús a la parada A y  $V[i+1]$  representa el tiempo de llegada a la parada B, por lo que  $dT$  es la diferencia de tiempo, lo que corresponde al tiempo que tarda en recorrer el vector formado por la parada A y la parada B. En la figura 4.10 se muestra un ejemplo con datos reales de la forma en la cual se obtienen los tiempos de los vectores en donde 04:11 es el tiempo de llegada a la siguiente parada y 02:23 es el tiempo de llegada a la parada actual.

$$dT=04:11-02:23$$

**Figura 4. 10 Ejemplo del primer paso del algoritmo**

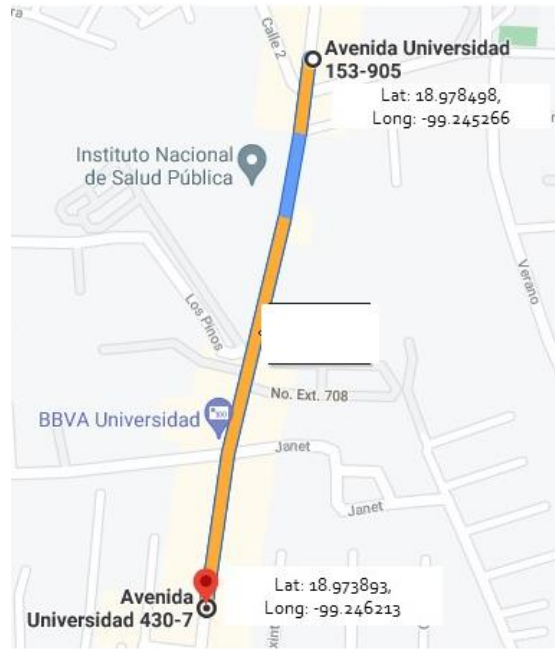
Se obtiene la distancia que hay en cada uno de los vectores creados, para este paso se hace uso de la longitud y latitud de las paradas, ya que para obtener dicha información se aplica la fórmula de Haversine [21].

La fórmula de Haversine se utiliza para calcular la distancia entre dos puntos en la superficie de la Tierra especificando su longitud y latitud. Es una fórmula que calcula distancias en línea recta por lo cual, al calcular la distancia de un trayecto que cuenta con algunos desvíos puede generar pérdidas grandes de la distancia entre esos dos puntos. Con la finalidad de utilizar este algoritmo sin tener pérdidas significantes se siguió la siguiente estrategia: incluir puntos adicionales en el trayecto de tal forma que se reduzcan al mínimo las líneas curvas. En nuestro caso es posible recorrer todo el trayecto de una cierta ruta e incluir estos puntos intermedios en las zonas con curvas. El mecanismo que se siguió para verificar la exactitud de algoritmo de Haversine obtener a distancia real entre dos puntos obtenida con un odómetro y verificarla contra la distancia obtenida con el algoritmo. En nuestro caso, solo se contaba con la información de la localización de las paradas, por lo se utilizó el algoritmo de Haversine para estimar la distancia entre esos dos puntos. Los resultados de la experimentación indicaron que existía una variación mínima entre las distancias calculadas entre el algoritmo y las distancias reales. Una de las ventajas de utilizar el algoritmo de Haversine es que resulta más sencilla la automatización del cálculo de la distancia entre dos puntos. De esta forma, cuando se desea incorporar una nueva ruta solo de debe obtener la localización entre dos puntos y el algoritmo nos proporcionará, en forma automática, un estimado de la distancia entre esos dos puntos.

En la figura 4.11 se muestra un ejemplo de cómo se obtiene la distancia de un vector. En la formula R equivale al radio de la tierra. En la figura 4.12 se muestra el mapa con las paradas y su información de latitud y longitud.

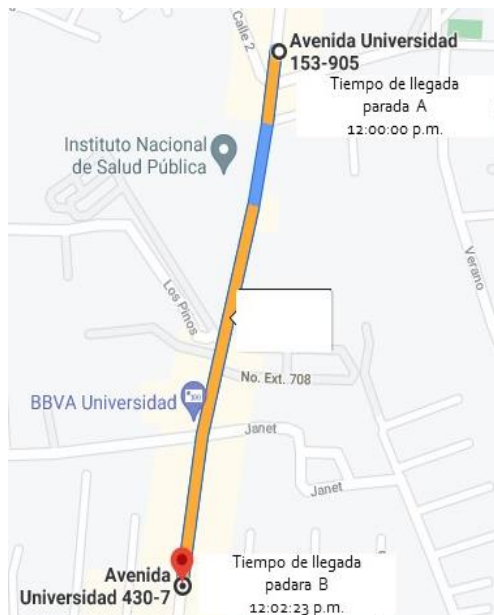
$$\begin{aligned}
 d &= R * c \\
 R &= 6372.8 \text{ km} \\
 dLat &= 3.1416 * (18.973893 - 18.978498) / 180 \\
 dLon &= 3.1416 * (-99.246213 - -99.245266) / 180 \\
 lat1 &= 3.1416 * 18.978498 / 180 \\
 lat2 &= 3.1416 * 18.973893 / 180 \\
 a &= \sin(dLat / 2) * \sin(dLat / 2) + \sin(dLon / 2) * \sin(dLon / 2) * \cos(lat1) * \cos(lat2); \\
 c &= 2 * \text{Atanz}(\sqrt{a})
 \end{aligned}$$

**Figura 4. 11 Paso 2: Desarrollo del algoritmo.**



**Figura 4. 12 Mapa con la información de las paradas**

Se calcula la velocidad promedio del recorrido total de los autobuses a través de todo su trayecto con la distancia y el tiempo obtenido de los vectores. En la figura 4.13 se muestra un ejemplo de las paradas con sus horas de llegada del autobús y en la figura 4.14 se muestra el algoritmo para calcular la velocidad promedio.



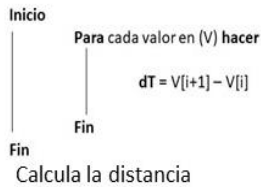
**Figura 4. 13 Ejemplo de paradas con sus tiempos históricos registrados**



Calcula el tiempo de los vectores

**Entrada:** Lista de vectores, V

**Salida:** Una lista de diferencia de tiempos, dT



d=R\*c

R=6372.8 km

dLat = 3.1416 \* (18.973893 - 18.978498) / 180

dLon = 3.1416 \* (-99.246213 - -99.245266) / 180

lat1 = 3.1416 \* 18.978498 / 180

lat2 = 3.1416 \* 18.973893 / 180

a=Sin(dLat / 2) \* Sin(dLat / 2) + Sin(dLon / 2) \* Sin(dLon / 2) \* Cos(lat1) \* Cos(lat2);

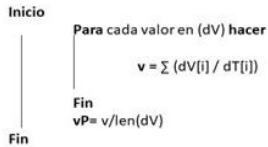
c=2 \* Atan2(sqrt(a))

Calcula la velocidad promedio

**Entrada:** Lista de distancia de vectores, dV[i]

Lista de diferencia de tiempos, dT[i]

**Salida:** Velocidad promedio del recorrido vP.



Calcula tiempo de predicción de arribo del autobús

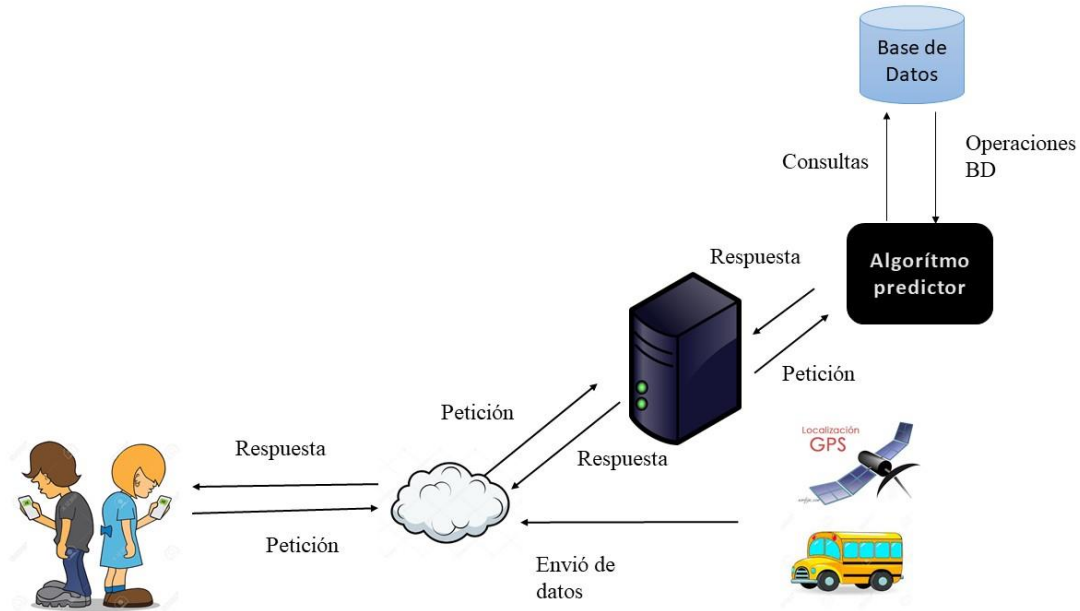
Tiempo=Distancia/Velocidad

**Figura 4. 15 Algoritmo completo**

## 4.4 Fase 3: Desarrollo del sistema de software

En esta fase se describe la arquitectura que fue diseñada para implementar el sistema Andy\_Bus. Este sistema tiene una arquitectura de tipo cliente servidor, lo cual significa que los dispositivos en los cuales se instalaron los sistemas se comunican con un servidor web de manera remota a través de peticiones, el cual obtiene las solicitudes, las analiza y si están bien devuelve una respuesta a los dispositivos (figura 4.16).





**Figura 4. 16 Arquitectura del sistema Andy\_Bus**

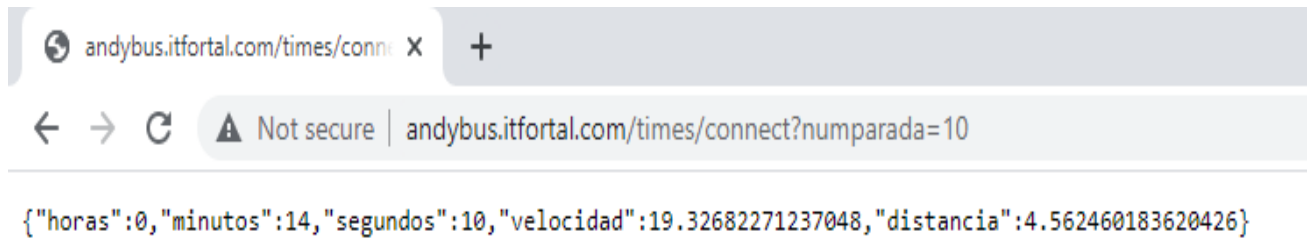
#### **4.4.1 Servidor Web**

La creación del servidor web se consiguió obteniendo un hosting en el cual se pudiera almacenar el algoritmo que realiza las predicciones de tiempos de arribo. En este hosting también se almacenó la base de datos remota, la cual es necesaria para poder acceder a los datos desde cualquier lugar a través de los dispositivos con los sistemas instalados. De igual manera fue necesario obtener un dominio, mediante el cual pudiésemos acceder a estos servicios. El dominio obtenido lleva por nombre <http://andybus.itfortal.com/>.

#### **4.4.2 Desarrollo de la Web API**

La web API desarrollada fue programada en lenguaje C# .net. Este lenguaje fue elegido ya que actualmente es de los más utilizados para realizar sistemas web. Con la finalidad de mantener de manera más ordenada la estructura de la web API se utilizó la arquitectura conocida como MVC (Modelo Vista Controlador).

En la figura 4.17 se muestra un ejemplo de la forma en la cual se envían los datos de la web API al cliente al momento que se realiza una petición.



**Figura 4. 17 Datos enviados por la web API**

Para la creación de la base de datos fue utilizado el manejador de base de datos Mysql (también conocido también como MariaDB). Se decidió utilizar este manejador ya que cuenta con funciones que hacen más amigables la creación y el control de las bases de datos y dicho manejador de código abierto es de los más utilizados actualmente.

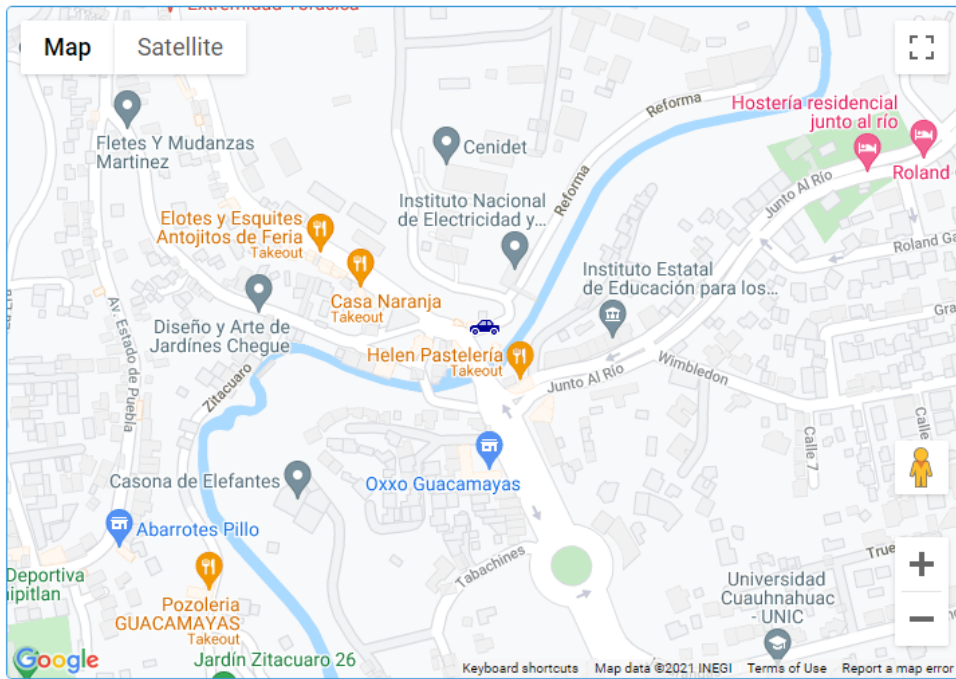
#### **4.4.3 Placa LinkIt One**

Para ubicar los autobuses en su localización actual y poder realizar la predicción de los tiempos de arribo fue necesario utilizar un dispositivo capaz de poder mantener una comunicación con el servidor web y que pudiese enviar, en tiempo real su ubicación, esto con la finalidad de poder mantener ubicado al autobús y tomar los registros de sus tiempos de arribo. El dispositivo utilizado fue una placa LinkIt One, se optó por esta placa ya que cuenta con las funcionalidades y especificaciones requeridas y otra de las razones es que cuenta con una batería propia, por lo cual no necesita de una fuente alterna de energía, esta placa cuenta con ranura en la cual se le puede insertar una tarjeta SIM de telefonía móvil, se le inserto una tarjeta SIM de la compañía Telcel para poder darle acceso a internet a la placa, cabe resaltar que esta placa es utilizada para proyectos relacionas con internet de las cosas, y nuestro sistema está relacionado con ese tema. Para programar la placa se utilizó el IDE de Arduino, ya que es el IDE oficial utilizado por la placa.

Fue necesario programar un servicio web que estuviese en comunicación con la placa LinkIt One ya que este servicio se encargaría únicamente de estar recibiendo la ubicación de la placa, de esta manera al momento de llegar a las paradas de los autobuses fue posible obtener la información de los tiempos de arribo, los cuales se utilizan para poner a funcionar el algoritmo desarrollado.

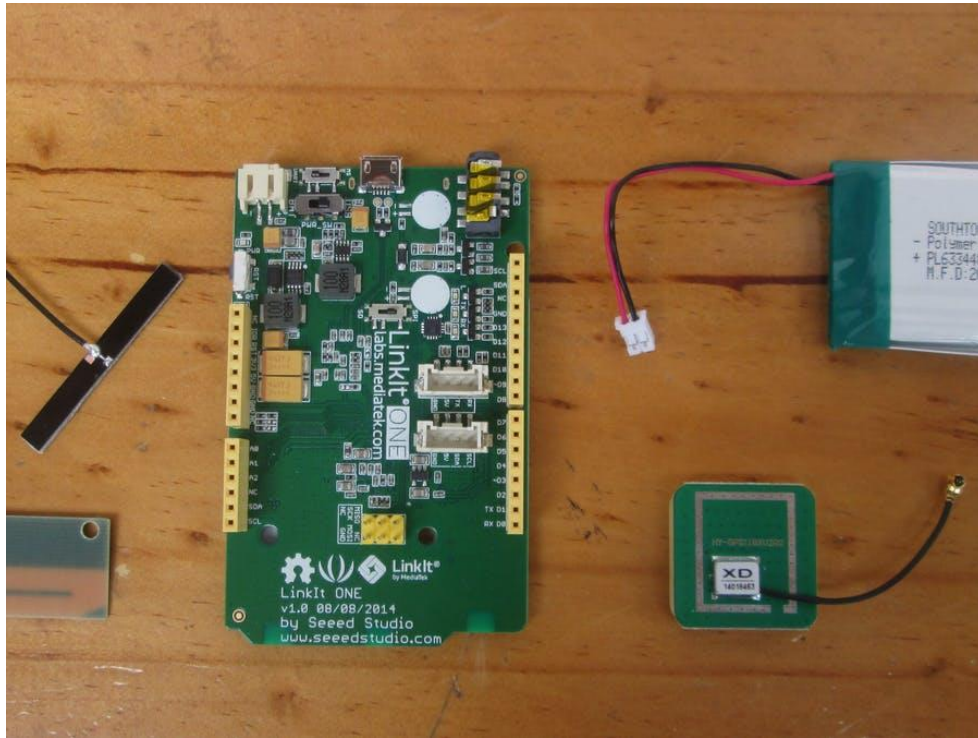
En la figura 4.18 se muestra el servicio desarrollado para rastrear la placa LinkIt One.

rastreador de autobuses



**Figura 4. 18 Servicio rastreador de la placa LinkIt One**

La tarjeta LinkIt One cuenta con una batería propia la cual le da energía para funcionar. La tarjeta envía su posición actual cada 5 segundos mediante el puerto 80 al sistema de rastreo el cual verifica si se encuentra en un rango de 20 metros cerca de la parada y de ser así se registra el tiempo de llegada en la base de datos. La tarjeta fue instalada en uno de la parte de enfrente del autobús justo a lado del chofer. En la figura 4.19 se muestra una imagen de la placa LinkIt One y sus accesorios.

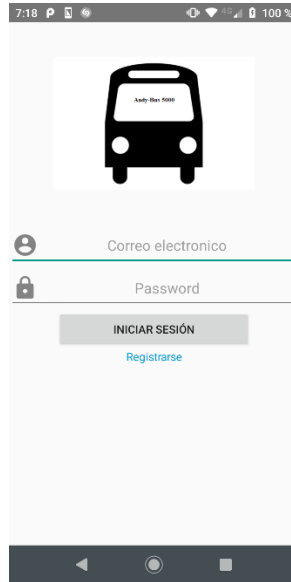


**Figura 4. 19** Placa LinkIt One y sus accesorios

#### **4.4.4 Aplicación Android Andy\_Bus**

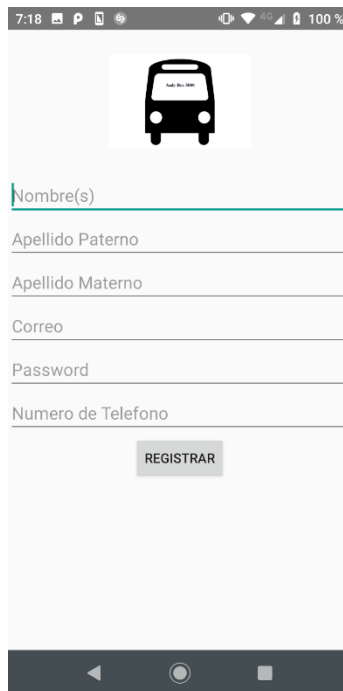
Para poder comunicar al cliente con el servidor y poder realizar las consultas sobre los tiempos de arribo de los autobuses, se desarrolló una aplicación móvil para smartphones con sistema operativo Android, la cual es utilizable para todos los que cuenten con un sistema operativo Android versión 4.4 Kitkat o superior. Gracias a esta aplicación el cliente puede ver los autobuses, los paraderos que tienen y también pueden consultar la parada de la cual necesitan obtener la predicción de tiempos de arribo.

Como primera función de la aplicación se encuentra la pantalla de inicio en la cual el usuario podrá iniciar sesión para posteriormente realizar las consultas. En la figura 4.20 se muestra la primera pantalla de la aplicación.



**Figura 4. 20 Pantalla de inicio de sesión de Andy\_Bus**

En el caso de usuarios no registrados, se cuenta con una interfaz para registrar una nueva cuenta, en la cual podrán realizar su registro para posteriormente realizar el inicio de sesión y acceder a los datos. En la figura 4.21 se muestra la pantalla en la cual el usuario puede realizar su registro.



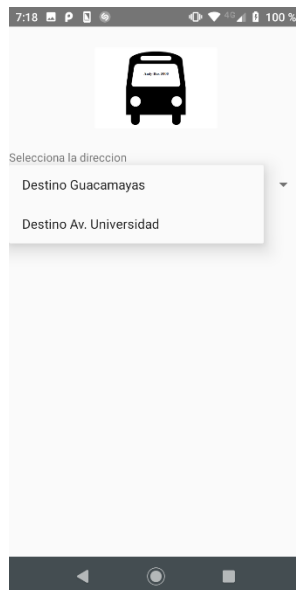
**Figura 4. 21 Pantalla de registro para nuevos usuarios**

Una vez que el usuario inicia sesión se le presentan los autobuses que están registrados en el sistema y de los cuales podrá obtener la información. En esta opción el usuario elegirá el autobús que sea de su interés. En la figura 4.22 se muestra la pantalla que enseña al usuario los autobuses registrados.



**Figura 4. 22 Pantalla con los autobuses registrados**

Cuando se elige el autobús se muestran las direcciones a las que viaja el autobús. En la figura 4.23 se muestra la pantalla de selección de dirección.



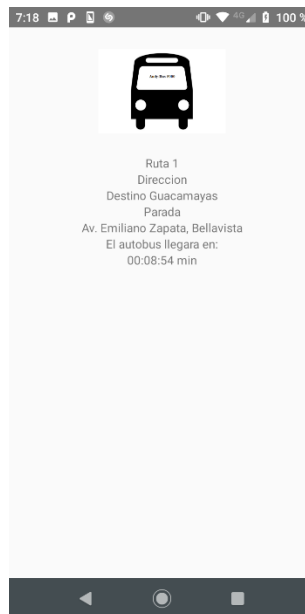
**Figura 4. 23 Pantalla de selección de dirección de autobús**

Una vez que se ha seleccionado la dirección en la que viaja el autobús se muestran sus paradas. El usuario debe seleccionar la parada en la que se encuentre o de la que se requiera obtener la predicción de tiempo de arribo. En la figura 4.24 se muestra la pantalla de selección de parada.



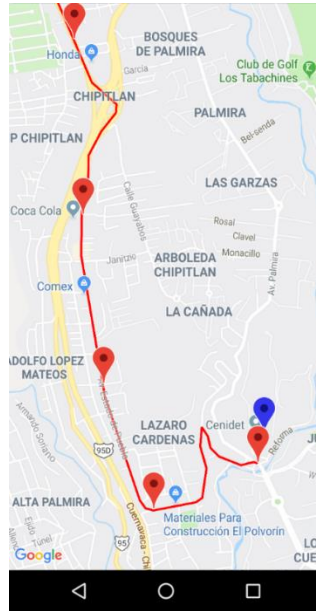
**Figura 4. 24 Pantalla de selecci3n de parada**

Una vez que se selecciona la parada, la aplicaci3n muestra en la siguiente interfaz un resumen con las selecciones que realizo el usuario, as3 como el tiempo de predicci3n de arribo del autob3s a la parada. En la figura 4.25 se muestra la pantalla que contiene el pron3stico de tiempo de arribo del autob3s.



**Figura 4. 25 Pantalla con predicci3n de tiempo de arribo**

La aplicaci3n tambi3n cuenta con la funcionalidad de desplegar un mapa en el cual se pueden observar las paradas que tiene el trayecto del autob3s y tambi3n puede ver en donde est3 ubicado el autob3s. En la figura 4.26 se muestra la pantalla de la aplicaci3n que despliega el mapa en el cual el marcador color azul representa el autob3s y los marcadores rojos sus paradas.



**Figura 4. 26 Pantalla de mapa**

Como actividad extra se desarrolló una página web en la que los administradores de la herramienta pueden dar de alta nuevas agencias, autobuses, rutas, paradas. En la figura 4.27 se muestra la interfaz de la página web en donde se registran las paradas de un autobús.

### Registrar nueva parada

Latitud

Longitud



**Figura 4. 27 Página de administradores**

## 4.5 Fase 4: Pruebas del sistema

Esta fase de la metodología se enfoca en probar el sistema y algoritmo desarrollado, en el siguiente capítulo se describe a detalle como es que se realizó esta fase.



# Capítulo 5 Pruebas y resultados

En este capítulo se describe el plan de pruebas establecido para llevar a cabo la evaluación del algoritmo y del sistema de software para realizar la predicción de tiempos de arribo de los autobuses.

## **5.1 Descripción de las pruebas**

El objetivo principal de este capítulo es evaluar el algoritmo utilizado en el sistema de software para poder realizar la predicción de tiempos de arribo de los autobuses. Los resultados obtenidos de las pruebas se presentan en tablas, en las cuales se registran las predicciones realizadas y las medidas de evaluación del sistema y el algoritmo.

- En la primera fase de las pruebas se seleccionaron las paradas que fueron tomadas en cuenta de la ruta que sigue el autobús, el autobús que fue seleccionado es el número 1 de la ciudad de Cuernavaca que tiene como paraderos iniciales y finales las calles de Guacamayas y Av. Universidad. También se seleccionó el día y hora en la que se realizaron las pruebas.
- En la segunda fase se realiza la prueba del sistema, ya que se ejecutó para obtener la predicción de los tiempos de arribo del autobús y ver la efectividad que tiene el algoritmo en su predicción. Para esta fase se realizaron algunos viajes en autobús para realizar la comparación de los tiempos de arribo reales con los tiempos predichos por el sistema de software.

## **5.2 Pruebas Fase 1**

En esta fase se analizaron las paradas que se tomaron en cuenta para realizar dicha predicción tomando en cuenta las características de ellas optando por la mejor opción. Otro factor que se tomó en cuenta fue el día y la hora, ya que se optó por realizar las pruebas en un día y horario normal, es decir, un día en el que no esté lloviendo, no haya desfiles, etc, y un horario en el que no haya mucha congestión de tráfico ya que de esta forma se puede realizar una mejor evaluación del sistema y del algoritmo.

## **5.3 Pruebas Fase 2**

En esta fase se realizan los viajes en el autobús obteniendo los tiempos de arribo reales y los tiempos de arribo pronosticados por el sistema de software. Cabe resaltar que se realizaron 3 viajes en el autobús número 1 los cuales iniciaron en la parada de autobús de guacamayas con dirección a Av. Universidad y otros 3 viajes en el mismo autobús que iniciaron en Av. Universidad con dirección a Guacamayas.

## **5.4 Pruebas en Escenario 1: Predicción de tiempos de arribo**

Las pruebas se realizaron en el autobús número 1 donde se realizaron 3 viajes, los cuales iniciaron en Av. Universidad y tomaron rumbo a su base en Guacamayas. Se tomaron en cuenta las primeras 10 paradas que conforman el trayecto mencionado como se muestra en la tabla 5.1. Para realizar las pruebas se fue cronometrando el viaje desde la parada inicial hasta la parada final seleccionada, y a la misma vez se fue consultando los tiempos de predicción de arribo del algoritmo desarrollado para posteriormente realizar una comparación de tiempos y ver el margen de error y la precisión que tiene el algoritmo al momento de realizar la predicción.

Los viajes se realizaron durante el mismo día y la misma hora con el fin de poder darle mayor precisión al algoritmo ya que la congestión del tráfico llega a variar constantemente a lo largo del día. Se eligió un día en el que no se presentaran intermitencias de tráfico como son desfiles, lluvias, entre otras más.

En las tablas 5.2, 5.3 y 5.4 se muestran los tiempos de arribo reales de los autobuses, así como los tiempos de arribo pronosticados por el algoritmo y de igual manera la diferencia de tiempo que se obtuvo entre los dos tiempos. Y en las figuras 5.1, 5.2 y 5.3 se muestran las gráficas de los datos obtenidos y su línea de dispersión.

<b>Paradas Ruta 1</b>			
<b>De Av. Universidad rumbo a Guacamayas</b>			
<b>No. de Parada</b>	<b>Nombre de parada</b>	<b>Latitud</b>	<b>Longitud</b>
1	Avenida Universidad 153-905, Universidad	18.978498	-99.245266
2	Avenida Universidad 430-7, Santa María Ahuacatitlan	18.973893	-99.246213
3	Don Miguel Hidalgo 15, Lienzo del Charro	18.969117	-99.244813
4	Avenida Universidad 62130, Lienzo del Charro	18.966776	-99.244186
5	Mexico 95 600, Bellavista	18.965786	-99.247064
6	Av. Emiliano Zapata, Bellavista	18.961289	-99.246290
7	El Globo, Av.Emiliano Zapata, Bellavista	18.956933	-99.245710
8	Banorte, Av. Emiliano Zapata, Bellavista	18.954390	-99.245341
9	La Pradera	18.939488	-99.241364
10	Alvaro Obregon 807-801	18.937544	-99.240883

**Tabla 5. 1 Descripción de paradas de la ruta 1**

Ruta 1 Viaje 1 De Av. Universidad rumbo a Guacamayas								
No. de Parada	Tiempo de Arribo Real (segundos)	Tiempo de Arribo del Algoritmo con odómetro (segundos)	Tiempo de Arribo del Algoritmo con Haversine (segundos)	Diferencia entre tiempo real y el del algoritmo (segundos)	Distancia calculada con odómetro (km)	Distancia calculada con Haversine (km)	Error Absoluto Porcentual del algoritmo	Precision del Pronostico del algoritmo
1	00:00	00:00	00:00	00:00	0 km	0 km	0%	100%
2	02:07	02:23	02:23	00:16	0.5 km	0.521 km	12.60%	87.40%
3	03:35	04:11	04:12	00:37	1.1 km	1.073 km	17.21%	82.79%
4	05:55	05:37	06:33	00:38	1.5 km	1.341 km	10.70%	89.30%
5	08:52	08:07	08:58	00:06	1.877 km	1.663 km	1.12%	98.88%
6	09:55	09:48	11:17	01:22	2.432 km	2.17 km	13.78%	86.22%
7	11:26	10:41	11:26	00:00	2.93 km	2.658 km	0%	100%
8	12:15	11:43	12:16	00:01	3.129 km	2.944 km	0.13%	99.87%
9	17:07	16:13	16:47	-00:20	4.922 km	4.653 km	1.95%	98.05%
10	17:35	17:37	18:11	00:36	5.13 km	4.875 km	3.41%	96.59%

Tabla 5. 2 Resultados viaje 1 con dirección a Guacamayas

MAPE (Medición de Error de Pronostico) = 6.77%.

Precisión = 93.23%.

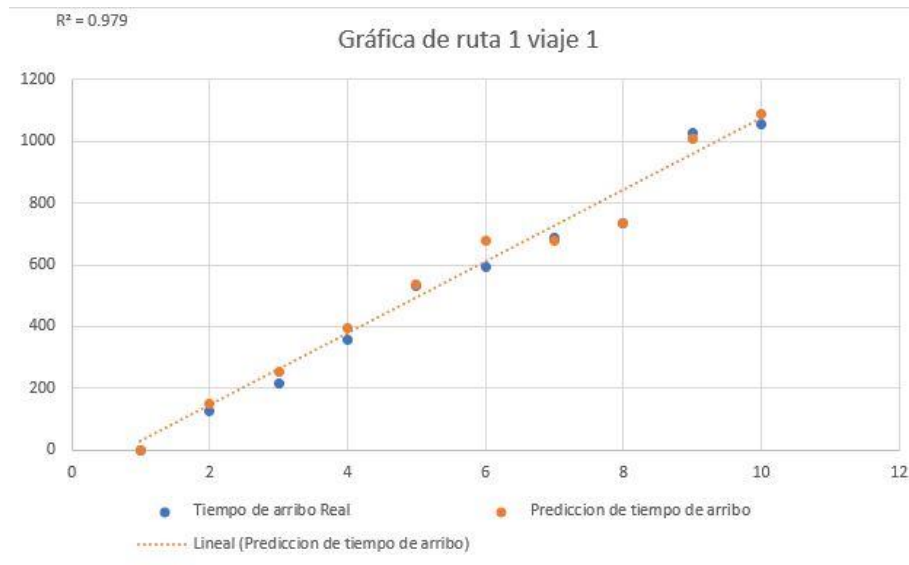


Figura 5. 1 Resultados viaje 1 con dirección a Guacamayas

Ruta 1 Viaje 2 De Av. Universidad rumbo a Guacamayas								
No. de Parada	Tiempo de Arribo Real (segundos)	Tiempo de Arribo del Algoritmo con odómetro (segundos)	Tiempo de Arribo del Algoritmo con Haversine (segundos)	Diferencia entre tiempo real y el del algoritmo (segundos)	Distancia calculada con odómetro (km)	Distancia calculada con Haversine (km)	Error Absoluto Porcentual del algoritmo	Precisión del Pronostico del algoritmo
1	00:00	00:00	00:00	00:00	0 km	0 km	0%	100%
2	01:12	02:07	02:07	00:55	0.5 km	0.521 km	76.38%	23.62%
3	04:35	03:35	02:51	-01:44	1.1 km	1.073 km	37.82%	62.18%
4	06:59	05:39	05:20	-01:39	1.5 km	1.341 km	23.63%	76.37%
5	08:48	08:08	07:47	-01:01	1.877 km	1.663 km	11.56%	88.44%
6	09:56	09:42	09:23	-00:33	2.432 km	2.17 km	5.54%	94.46%
7	11:18	10:28	10:13	-01:05	2.93 km	2.658 km	9.59%	90.41%
8	12:00	11:15	11:01	-00:59	3.129 km	2.944 km	8.20%	91.80%
9	16:00	15:40	15:26	-00:34	4.922 km	4.653 km	3.54%	96.46%
10	16:26	16:49	16:44	00:18	5.13 km	4.875 km	1.82%	98.18%

**Tabla 5. 3 Resultados viaje 2 con dirección a Guacamayas**

MAPE (Medición de Error de Pronóstico) = 19.79%.

Precisión = 80.21%.



**Figura 5. 2 Resultados viaje 2 con dirección a Guacamayas**

Ruta 1 Viaje 3 De Av. Universidad rumbo a Guacamayas								
No. de Parada	Tiempo de Arribo Real (segundos)	Tiempo de Arribo del Algoritmo con odometro (segundos)	Tiempo de Arribo del Algoritmo con Haversine (segundos)	Diferencia entre tiempo real y el del algoritmo (segundos)	Distancia calculada con odómetro (km)	Distancia calculada con Haversine (km)	Error Absoluto Porcentual del algoritmo	Precision del Pronostico del algoritmo
1	00:00	00:00	00:00	00:00	0 km	0 km	0%	100%
2	02:25	01:12	01:12	-01:13	0.5 km	0.521 km	50.35%	49.65%
3	04:31	05:40	05:19	00:48	1.1 km	1.073 km	17.71%	82.29%
4	06:46	07:18	08:15	01:29	1.5 km	1.341 km	21.92%	78.08%
5	08:25	09:33	09:41	01:16	1.877 km	1.663 km	19.79%	80.21%
6	09:28	11:06	11:12	01:44	2.432 km	2.17 km	18.31%	81.69%
7	10:42	11:07	11:11	00:29	2.93 km	2.658 km	4.52%	95.48%
8	11:18	11:33	11:37	00:19	3.129 km	2.944 km	2.80%	97.20%
9	15:36	15:17	15:21	-00:15	4.922 km	4.653 km	1.61%	98.39%
10	16:01	16:32	16:36	00:35	5.13 km	4.875 km	3.64%	96.36%

Tabla 5. 4 Resultados viaje 3 con dirección a Guacamayas

MAPE (Medición de Error de Pronóstico) = 15.63%.

Precisión = 84.37%.



Figura 5. 3 Resultados viaje 3 con dirección a Guacamayas

## 5.5 Pruebas en Escenario 2: Predicción de tiempos de arribo

Las pruebas se realizaron en el autobús numero 1 donde se realizaron 3 viajes, los cuales iniciaron en Guacamayas y tomaron rumbo a su base en Av. Universidad. Se tomaron en cuenta las primeras 10

paradas que conforman el trayecto mencionado como se muestra en la tabla 5.5. Para realizar las pruebas se fue cronometrando el viaje desde la parada inicial hasta la parada final seleccionada y a la misma vez se fue consultando los tiempos de predicción de arribo del algoritmo desarrollado para posteriormente realizar una comparación de tiempos y ver el margen de error y de precisión que tiene el algoritmo al momento de realizar la predicción.

Los viajes se realizaron durante el mismo día y la misma hora con el fin de poder darle mayor precisión al algoritmo ya que la congestión del tráfico llega a variar constantemente a lo largo del día. Se eligió un día en el que no se presentaran intermitencias de tráfico como son desfiles, lluvias, entre otras más.

En las tablas 5.6, 5.7 y 5.8 se muestran los tiempos de arribo reales de los autobuses, así como los tiempos de arribo pronosticados por el algoritmo y de igual manera la diferencia de tiempo que se obtuvo entre los dos tiempos. Y en las figuras 5.4, 5.5 y 5.6 se muestran las gráficas de los datos obtenidos y su línea de dispersión.

<b>Paradas Ruta 2 De Guacamayas rumbo a Av. Universidad</b>			
No. de Parada	Nombre de parada	Latitud	Longitud
1	Guacamayas	18.875329	-99.219914
2	Av. Estado de Puebla s/n, Lazaro Cardenas	18.876969	-99.222654
3	Av. Estado de Puebla 173, Lazaro Cardenas	18.873510	-99.223358
4	Av. Estado de Puebla 116, Lazaro Cardenas	18.876284	-99.226658
5	Av. Estado de Puebla 95, Lazaro Cardenas	18.881705	-99.228063
6	Av. Estado de Puebla 9, Lazaro Cardenas	18.886716	-99.228580
7	Av. José María Morelos y Pavón Nte. 87-81, Chipitlan	18.894904	-99.228945
8	José María Morelos, Chipitlan	18.896639	-99.229815
9	Av. José María Morelos y Pavón Nte. 80-86, Las Palmas	18.902485	-99.231508
10	Las Palmas	18.905924	-99.231900
11	José Ma. Morelos y Pavón 810, Las Palmas	18.908106	-99.232837

**Tabla 5. 5 Descripción de paradas de la ruta 2**

Ruta 2 Viaje 1 De Guacamayas rumbo a Av. Universidad								
No. de Parada	Tiempo de Arribo Real (segundos)	Tiempo de Arribo del Algoritmo con odómetro (segundos)	Tiempo de Arribo del Algoritmo con Haversine (segundos)	Diferencia entre tiempo real y el del algoritmo (segundos)	Distancia calculada con odómetro (km)	Distancia calculada con Haversine (km)	Error Absoluto Porcentual del algoritmo	Precision del Pronostico del algoritmo
1	00:00	00:00	00:00	00:00	0 km	0 km	0%	100%
2	01:16	01:18	01:18	00:02	0.4 km	0.341 km	2.63%	97.37%
3	02:48	02:43	02:40	-00:08	0.85 km	0.733 km	4.76%	95.24%
4	05:25	05:29	05:01	-00:24	1.372 km	1.197 km	7.39%	92.61%
5	07:31	07:43	07:11	-00:20	2.06 km	1.818 km	4.44%	95.56%
6	09:25	10:01	09:22	-00:03	2.572 km	2.378 km	0.53%	99.47%
7	11:35	11:56	11:24	-00:11	3.668 km	3.289 km	1.59%	98.41%
8	12:38	13:55	13:28	01:00	3.889 km	3.503 km	6.59%	93.41%
9	14:42	13:52	15:03	00:19	4.582 km	4.177 km	2.38%	97.62%
10	15:39	14:16	15:28	-00:11	4.977 km	4.562 km	1.17%	98.83%
11	16:20	15:50	16:49	00:29	5.17 km	4.824 km	2.96%	97.04%

**Tabla 5. 6 Resultados viaje 1 con dirección a Av. Universidad**

MAPE (Medición de Error de Pronóstico) = 3.45%.

Precisión = 96.55%.



**Figura 5. 4 Resultados viaje 1 con dirección a Av. Universidad**



Ruta 2 Viaje 2 De Guacamayas rumbo a Av. Universidad								
No. de Parada	Tiempo de Arribo Real (segundos)	Tiempo de Arribo del Algoritmo con odómetro (segundos)	Tiempo de Arribo del Algoritmo con Haversine (segundos)	Diferencia entre tiempo real y el del algoritmo (segundos)	Distancia calculada con odómetro (km)	Distancia calculada con Haversine (km)	Error Absoluto Porcentual del algoritmo	Precision del Pronostico del algoritmo
1	00:00	00:00	00:00	00:00	0 km	0 km	0%	100%
2	00:52	01:16	01:16	00:24	0.4 km	0.341 km	46.15%	53.85%
3	02:18	02:46	02:27	00:09	0.85 km	0.733 km	6.52%	93.48%
4	04:15	05:16	04:24	00:09	1.372 km	1.197 km	3.53%	96.47%
5	06:28	07:50	06:38	00:10	2.06 km	1.818 km	2.57%	97.43%
6	08:52	10:11	08:39	-00:13	2.572 km	2.378 km	2.45%	97.55%
7	11:26	12:50	11:15	-00:11	3.668 km	3.289 km	1.61%	98.39%
8	12:26	12:18	12:34	00:08	3.889 km	3.503 km	1.07%	98.93%
9	14:06	13:59	14:38	00:32	4.582 km	4.177 km	3.78%	96.22%
10	15:10	14:35	15:33	00:23	4.977 km	4.562 km	2.52%	97.48%
11	16:30	15:05	16:12	-00:18	5.17 km	4.824 km	1.82%	98.18%

**Tabla 5. 7 Resultados viaje 2 con dirección a Av. Universidad**

MAPE (Medición de Error de Pronóstico) = 7.2%.

Precisión = 92.8%.



**Figura 5. 5 Resultados viaje 2 con dirección a Av. Universidad**

Ruta 2 Viaje 3 De Guacamayas rumbo a Av. Universidad								
No. de Parada	Tiempo de Arribo Real (segundos)	Tiempo de Arribo del Algoritmo con odómetro (segundos)	Tiempo de Arribo del Algoritmo con Haversine (segundos)	Diferencia entre tiempo real y el del algoritmo (segundos)	Distancia calculada con odómetro (km)	Distancia calculada con Haversine (km)	Error Absoluto Porcentual del algoritmo	Precision del Pronostico del algoritmo
1	00:00	00:00	00:00	00:00	0 km	0 km	0%	100%
2	01:15	00:52	00:52	-00:23	0.4 km	0.341 km	30.67%	69.33%
3	02:29	02:44	02:41	00:12	0.85 km	0.733 km	8.05%	91.95%
4	04:40	04:47	04:27	-00:13	1.372 km	1.197 km	4.65%	95.35%
5	06:30	07:15	06:43	00:13	2.06 km	1.818 km	3.33%	96.67%
6	08:19	09:09	08:51	00:32	2.572 km	2.378 km	6.41%	93.59%
7	10:02	10:33	11:33	01:31	3.668 km	3.289 km	15.11%	84.89%
8	10:42	11:00	11:55	01:13	3.889 km	3.503 km	11.37%	88.63%
9	11:52	11:37	13:29	01:37	4.582 km	4.177 km	13.62%	86.38%
10	12:46	12:35	13:46	01:00	4.977 km	4.562 km	7.83%	92.17%
11	13:44	13:55	14:19	00:35	5.17 km	4.824 km	4.24%	95.76%

**Tabla 5. 8 Resultados viaje 3 con dirección a Av. Universidad**

MAPE (Medición de Error de Pronóstico) = 10.53%.

Precisión = 89.47%.



**Figura 5. 6 Resultados viaje 3 con dirección a Av. Universidad**

En la siguiente tabla se muestra el peor, mejor y mediano escenario que se obtuvo al momento de realizar las pruebas del algoritmo.

Mejor Escenario	Escenario Medio	Peor Escenario
96.55%	92.8% y 89.47%	80.21%

**Tabla 5. 9 Tabla de comparaciones**

## 5.6 Resumen de resultados y conclusiones

Los resultados de las pruebas realizadas se enfocaron al correcto funcionamiento del algoritmo desarrollado para la predicción de tiempos de arribo de los autobuses para comprobar si es viable la utilización del mismo.

Como se puede observar en las pruebas realizadas los viajes realizados en el escenario 2 tienen un rango de error menor que el de las pruebas realizadas en el escenario 1, uno de los factores que se pudieron observar para que se diera esto es que en el escenario 2 el trayecto que toman los autobuses es menos transitado que el trayecto del escenario 1, otro de los factores que llegan a influir en estos márgenes de error es la manera en la que conduce el chofer de los autobuses, ya que hay algunos que conducen más lentos que otros, también llegan a influir algunos factores como la cantidad de gente que suben en cada una de sus paradas ya que esto produce aproximadamente un retraso de 5 segundos por persona que aborda. Otro factor que nos puede mostrar una mejor precisión en el algoritmo es las gráficas mostradas, ya que se puede ver la línea de dispersión, y mediante a ella podemos observar cual de los escenarios nos muestra una mejor dispersión de los datos.

Utilizando un sistema estandarizado como es *GTFS (General Transit Specification Feed.)* se puede tener un mejor control de predicción de tiempos de arribo, ya que con este sistema de estandarización se establecen paradas generales y específicas para todos los autobuses que siguen una misma ruta, ya que de no ser de esta forma los autobuses pueden realizar cuantas paradas deseen en el sitio que quieran sin seguir ningún tipo de norma, lo cual, genera muchos factores para provocar el retraso de los autobuses como es congestión de tránsito, retraso en la hora de arribo, etc. Por lo cual, se concluye que es más factible adaptar los sistemas de predicción de tiempos de arribo del transporte público al sistema *GTFS*.

Dado a los resultados obtenidos se puede concluir que, si es viable utilizar el algoritmo, ya que tiene un margen de error pequeño por lo cual da una buena predicción de los tiempos de arribo.

# Capítulo 6 Conclusiones

En esta sección se presentan las conclusiones que se obtuvieron al realizar y probar este trabajo de investigación

## 6.1 Conclusiones

El transporte público en México es uno de los servicios con mayor demanda según cifras del INEGI, es muy utilizado para trasladar a las personas de un lugar a otro ahorrándole tiempo y esfuerzo. Sin embargo, presenta varios inconvenientes, como es la falta de una buena organización por parte de las agencias que lo administran, falta de reglamentos hacia los conductores, sistema de estandarización de horarios, entre otros más.

El implementar sistemas de transporte inteligente puede causar la reducción de algunos inconvenientes de este tipo. Esto porque este tipo de sistemas proveen información en tiempo actual sobre qué pasa con el transporte público lo cual permite ser utilizada para tomar decisiones por los usuarios de esta información. Esta es una de las motivaciones que se tienen para implementar este tipo de sistemas en lugares en donde no se cuenta con ellos.

En este trabajo de investigación se desarrolló un sistema de transporte inteligente con infraestructura de mediano costo cuyo objetivo es obtener información del transporte público y con esto procesarla y mejorarlo a través de herramientas tecnológicas. La principal funcionalidad es el de proporcionar información a los usuarios del transporte público para la toma de decisiones o planeación de sus viajes. La información que proporciona es la del tiempo de arribo de los autobuses. Para la realización de este proyecto se desarrolló una metodología de solución que consta de 4 fases: Recopilación de información, desarrollo del algoritmo, desarrollo del sistema de software y concluyendo con la fase de pruebas del sistema.

Una de las fases más importantes es la de recopilar información, ya que esta es la base del desarrollo del sistema, puesto que con ella se realizan las predicciones, por lo cual, fue muy importante estructurarla y obtenerla de manera eficiente. Para el cálculo de las distancias entre las paradas de autobús se utilizó el algoritmo de Haversine que nos permite calcular de forma automática la distancia entre dos puntos.

Dada la información recolectada se desarrolló el algoritmo de predicción de tiempos de arribo para posteriormente poder desarrollarlo en una Web API, la cual, se encarga de mandar la información de los autobuses a los usuarios una vez que le llega la petición.

El sistema se compone de un servidor web, una base de datos remota almacenada en el servidor web, una Web API, una placa de LinkIt One y una aplicación para *Smartphones* con sistema operativo Android. Tanto la aplicación como la placa de LinkIt One interactúan con el servidor web para obtener la información pedida por los usuarios.

Se concluye que el algoritmo y sistema desarrollado son factibles para utilizarse, esto por los resultados obtenidos en la fase de pruebas.

## **6.2 Logros Obtenidos**

Aceptación y presentación del artículo titulado: “*Sistema de Transporte Inteligente para Predecir los Tiempos de Arribo del Transporte Publico*” En la segunda jornada de ciencia y tecnología aplicada que se llevó a cabo los días 04 y 05 de abril de 2019 en la ciudad de Cuernavaca, Morelos.



**SEP**  
SECRETARÍA DE  
EDUCACIÓN PÚBLICA



TECNOLÓGICO  
NACIONAL DE MÉXICO

**EL TECNOLÓGICO NACIONAL DE MÉXICO  
A TRAVÉS DEL CENTRO NACIONAL DE INVESTIGACIÓN  
Y DESARROLLO TECNOLÓGICO**

OTORGAN LA PRESENTE

**CONSTANCIA**

A

**Andrés Ocampo Ortiz, Alicia Martínez Rebollar,  
Javier Ortiz**

POR SU PARTICIPACIÓN CON EL ARTÍCULO: **“SISTEMA DE TRANSPORTE INTELIGENTE PARA PREDECIR LOS TIEMPOS DE ARRIBO DEL TRANSPORTE PÚBLICO”**, PRESENTADO EN LA SEGUNDA JORNADA DE CIENCIA Y TECNOLOGÍA APLICADA, QUE SE LLEVO A CABO LOS DÍAS 4 Y 5 DE ABRIL DE 2019, EN LA CIUDAD DE CUERNAVACA, MORELOS, MÉXICO.

CUERNAVACA, MOR., A 5 DE ABRIL DE 2019.



S.E.P.  
CENTRO NACIONAL DE INVESTIGACION  
Y DESARROLLO TECNOLÓGICO

**DR. VÍCTOR HUGO OLIVARES PEREGRINO**  
DIRECTOR



### **6.3 Trabajos futuros**

Los trabajos futuros que se proponen para ampliar y/o mejorar este trabajo de investigación se listan a continuación.

- Desarrollar un método que haga predicción del tiempo de arribo de los autobuses a la ubicación del usuario, aunque no esté en una parada siempre y cuando el usuario se encuentre en una ubicación por la que pase el autobús.
- Desarrollar un método que le diga al usuario que autobuses debe de tomar para llegar a cierto punto.
- Realizar un estudio de cual sensor o dispositivo que cuente con GPS es mejor utilizar para poder rastrear un autobús.
- Implementar sensores en las paradas los cuales registren la hora de llegada del autobús a la parada.
- Mejorar el algoritmo predictor y la interfaz gráfica para mostrar al usuario las predicciones de todos los autobuses que están próximos a llegar a las paradas.
- Investigar, analizar y aplicar otro método que sea más eficiente que el de Haversine para calcular la distancia.



## Referencias

- [1] EDOMÉX, “Transporte público”, smovilidad.edomex.gob.mx, para. 1, 2018. [Online]. Available: [http://smovilidad.edomex.gob.mx/transporte\\_publico](http://smovilidad.edomex.gob.mx/transporte_publico). [Accesed Feb. 20, 2021].
- [2] WeblogsSL, “Todo sobre los ITS, los sistemas de transporte inteligentes”, motorpasion.com, para. 1, Nov. 18, 2013. [Online]. Available: <https://www.motorpasion.com/espaciotoyota/todo-sobre-los-its-los-sistemas-de-transporte-inteligentes>. [Accesed Feb. 28, 2021].
- [3] D, Google, “¿Qué es la GTFS?”, developers.google.com, para. 1, Mar. 31, 2016. [Online]. Available: <https://developers.google.com/transit/gtfs/?hl=es>. [Accesed Feb. 28, 2021].
- [4] P. Wu, A. Che, F. Chu, and Y. Fang, “Exact and heuristic algorithms for rapid and station arrival-time guaranteed bus transportation via lane reservation,” *IEEE Transactions on intelligent transportation systems*, vol. 18, no. 8, Aug., pp.2028-2043, 2017.
- [5] M.S. Chandurkar, S. Mugade, S. Sinha, M. Misal and P. Borekar, “Implementation of real time bus monitoring and passager information system”, *International journal of scientific and research publications*, vol. 3, no. 5, May., pp.2250-3153, 2013.
- [6] M. Hata, H.Adachi, H.Suzuki, K.Kitase, Y.Matsumoto, A.Watanabe, “Proposal for Novel Bus Location System Using Wireless Sensor Network”, *ITS Japan, no.3209*, Aug., 2013.
- [7] G.B.López, “La geolocalización social”, *Poligonos*, no.27, Jul., pp.2444-0272, 2015.
- [8] L. Serrano, “Diseñan sistema que monitorea y predice llegada de autobuses de transporte público”, *conacytprensa.mx*, para. 1, Aug. 15, 2014. [Online]. Available: <http://www.conacytprensa.mx/index.php/tecnologia/tic/109-disenan-sistema-que-monitorea-y-predice-llegada-de-autobuses-de-transporte-publico>. [Accesed Feb. 28, 2021].
- [9] J. Tadeo, “México: Larealidad del transporte público en Ciudad de México”, *globalvoices.org*, para. 1, Apr. 15, 2014. [Online]. Available: <https://es.globalvoices.org/2011/04/15/mexico-la-realidad-del-transporte-publico-en-ciudad-de-mexico/>. [Accesed Feb. 28, 2021].
- [10] R. Iiya, E. Zhou, M. Niibori, M. Kamada, O. Saitou and S. Shibusawa, “A Handy Real-Time Location System for Community Bus Service”, *19th International Conference on Network-Based Information Systems*, no.19, Sep., pp. 506-509, 2016.
- [11] P. Wences, A. Martinez, H. Estrada and M. Gonzalez, “Decision-Making Intelligent System for Passenger of Urban Transports”, *International Conference on Ubiquitous Computing and Ambient Intelligence*, Oct., pp. 128-139, 2017.
- [12] J. Amita, J. Sukhvir and G. Pradeep, “Prediction of bus travel time using artificial neural network”, *International Journal for Traffic and Transport Engineering*, vol. 5, Aug., pp. 410-424, 2015.

- [13] H. Yu, Z. Wu, D. Chen and X. Ma, “Probabilistic Prediction of Bus Headway Using Relevance Vector Machine Regression”, *IEEE Transaction on Intelligent Transportation Systems*, vol. 18, no. 7, Jul., pp. 1772-1781, 2017
- [14] Z. Kebede, T. Nall and Perkins, “Artificial Neural Network Travel Time Prediction Model for Buses Using Only GPS Data”, *Journal of Public Transportation*, vol. 17, no.2, pp. 45-65, 2014.
- [15] T. Ma, G. Motta and K. Liu, “Delivering Real-Time Information Services on Public Transit: A Framework”, *IEEE Transaction on Intelligent Transportation Systems*, vol. 18, no. 10, Oct., pp. 2642-2656, 2017.
- [16] T. Zhang, M. Chen and C. Lawson, “General Transit Feed Specification Data Visualization”, *22nd International Conference on Geoinformatics*, Jun., 2014.
- [17] A. Antrim and S. Barbeau, “The Many Uses of GTFS Data-Opening the Door to Transit and Multimodal Applications”, *ITS American’s 23rd Annual Meeting & Exposition*, Apr., 2013.
- [18] A. Sullivan, F. Pereira, J. Zhao and H. Koutsopoulos, “Uncertainty in Bus Arrival Time Predictions: Treating Heteroscedasticity With a Metamodel Approach”, *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no.11, Nov., pp. 3286-3296, 2016.
- [19] J. Li, J. Gao, Y. Yang and H. Wei, “Bus Arrival Time Prediction Based on Mixed Model”, *China Communications*, vol. 14, issue. 5, May., pp.38-47, 2017.
- [20] Waze, ”waze”, waze.com, para. 1, 2006. [Online]. Available: <https://www.waze.com/?locale=es-419>. [Accesed Feb. 28, 2021].
- [21] Ken Burnside, ”La ecuacion de Haversine en excel”, techlandia.com, para. 1, 2021. [Online]. Available [techlandia.com/ecuacion-haversine-excel-info\\_201797/](http://techlandia.com/ecuacion-haversine-excel-info_201797/). [Accesed Feb. 28, 2021].