

**HARITI.
PROYECTO PARA EL ENRUTAMIENTO DEL TRANSPORTE ESCOLAR.**

ROBERT ANTONIO BARRAZA GAMERO

**UNIVERSIDAD DE LOS ANDES
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN
MAESTRÍA EN INGENIERÍA DE SISTEMAS Y COMPUTACIÓN
BOGOTÁ D.C.
2006**

CONTENIDO

	pág.
INTRODUCCIÓN	6
1. OBJETIVOS	8
1.1. OBJETIVO GENERAL	8
1.2. OBJETIVOS ESPECÍFICOS	8
2. JUSTIFICACIÓN	9
3. ALCANCE	10
4. MARCO TEÓRICO	12
4.1. TRANSPORTE URBANO	12
4.1.1. Transporte escolar.	12
4.1.2. Entidades reguladores y normatividad.	14
4.1.2.1. Ministerio de Transporte.	14
4.1.2.1.1. Resolución 1122 de 2005.	14
4.1.2.1.2. Decreto Número 3366 de 2003.	15
4.1.2.2. Secretarías de Tránsito y Transporte.	16
4.2. ALGORITMOS Y COMPLEJIDAD	16
4.2.1. Algoritmos deterministas.	16
4.2.2. Algoritmos no deterministas.	17
4.2.3. Máquina de Turing.	18
4.2.4. Complejidad L	18
4.2.5. Complejidad NL.	18
4.2.6. Complejidad P.	19
4.2.7. Complejidad NP.	19
4.2.8. Complejidad NP-Completo.	19
4.2.9. Complejidad NP-Complejo o NP-Hard	19
4.3. ANÁLISIS DEL PROBLEMA VRP	20
4.3.1. Problema del agente viajero.	20
4.3.2. Problema de la mochila (Bin Packing Problem).	22
4.3.3. Definición de VRP.	22
4.3.3.1. Definición formal de VRP.	23
4.3.3.2. Parámetros de VRP.	24
4.4. PROBLEMA DE ENRUTAMIENTO DE VEHÍCULOS ESCOLARES COMO VRP	25
4.5. HEURÍSTICA	25
4.6. METAHEURÍSTICA	25

4.6.1. Búsqueda Tabú	26
4.7. DESCRIPCIÓN DEL PROBLEMA	27
4.7.1. Definición de términos.	27
4.7.2. Problema de enrutamiento de transporte escolar.	28
4.7.3. Definición matemática del problema.	28
4.7.1. Función objetivo.	31
5. SOLUCIÓN PROPUESTA	34
5.1. LEVANTAMIENTO DE REQUERIMIENTOS	34
5.1.1. Identificación de actores.	34
5.2. DISEÑO	35
5.2.1.1. Módulo de lectura de entradas.	35
5.2.1.2. Módulo de fragmentación o segmentación.	35
5.2.1.2.1 Descripción del algoritmo.	36
5.2.1.2.1. Diagrama de clase.	39
5.2.1.3. Módulo geocodificación.	39
5.2.1.4. Módulo metaheurística.	41
5.2.1.5. Módulo de post-optimación o implosión.	42
6. RESULTADOS	45
6.1. RESULTADOS SEGMENTACIÓN	45
6.2. COMPARACIÓN TABU SEARCH – SIMULATED ANNEALING	47

LISTA DE TABLAS

	pág.
Tabla 1. Proyecciones de población en edad escolar. 2002-2013	12
Tabla 2. Estructura dirección Norteamérica.	39
Tabla 3. Ejemplo direcciones colombianas.	39
Tabla 4. Tabla de la malla vial modificada.	40
Tabla 5. Antes y después de las direcciones.	40
Tabla 6. Resultados validación segmentación.	46
Tabla 7. Resultados para Simulated Annealing.	47
Tabla 8. Resultados para Tabu Search.	47

LISTA DE FIGURAS

	pág.
Figura 1. Alumnos transportados en rutas escolares, 2003.	13
Figura 2. Inscritos nuevos y su asignación en instituciones educativas distritales – 2003 (miles de alumnos)	13
Figura 3. Algoritmo Búsqueda Tabú.	27
Figura 4. Segmentación con KMeans (1000 estudiantes – 30 segmentos).	37
Figura 5. Centroides después del proceso de KMeans.	37
Figura 6. Segmentación con KMeans con capacidad.	38
Figura 7. Puntos alrededor de un estudiante, 4 segmentos, 200 metros.	43
Figura 8. Obtención puntos alrededor de un estudiante.	44
Figura 9. Modelo segmentación.	45
Figura 10. Comportamiento del modelo.	46
Figura 11. Resultados en costo y tiempo de las dos metaheurísticas.	48

INTRODUCCIÓN

El problema de la planificación del transporte en las ciudades colombianas ha tomado mucha importancia en los últimos años. Las diferentes ciudades han afrontado el problema del transporte masivo de distintas formas: el Metro de Medellín, Transmilenio en Bogotá, MIO en Cali, Transmetro en Barranquilla, Transcaribe en Cartagena. Estas ciudades han invertido gran parte de su presupuesto en la aplicación de estas soluciones [1].

Pero el problema del transporte masivo no es el único con que se enfrentan las grandes ciudades, también existe el de los autos particulares, el de las motocicletas y específicamente el del transporte escolar.

Como muchas veces sucede, el problema, en este caso el del transporte escolar, es más complejo de lo que a simple vista se ve, al igual que su relevancia es mayor de lo que podría parecer.

La optimización de rutas de transporte escolar se considera un tema prioritario por parte de las administraciones públicas por la obligatoriedad de su servicio y la trascendencia social y económica que supone. Las rutas de transporte escolar deben ser adaptadas de forma continua a las necesidades de los alumnos de cada centro para cada momento. [2]

Es pues, de suma importancia, la implementación de medios que optimicen los recorridos de los distintos estudiantes a su respectivo plantel educativo y viceversa. La optimización debe incluir entre otras cosas:

- La disminución de la distancia que se desplaza cada vehículo.
- Un seguimiento más cercano de la velocidad de los vehículos para impedir el sobrepaso de los límites establecidos.

El problema del transporte escolar se palpa de manera significativa en ciudades como Bogotá en donde, según cifras de la Secretaría de Educación Distrital, en el 2003 se transportaron 23.343 estudiantes sólo en los estratos 1 y 2; entre los grados 0 a 9 [3], los jóvenes deben desplazarse grandes distancias para llegar a su sitio de estudio, deben tomar varios buses para realizar el recorrido hasta su centro educativo.

Los colegios que cuentan con rutas escolares no escapan a esta problemática, basan sus recorridos en el conocimiento empírico y aproximaciones, muchas veces subjetivas, lo que ocasiona que dichos recorridos demoren más de lo necesario y gasten más recursos. En ambos contextos, tanto para el usuario de buses urbanos, como el de rutas escolares, dicho aumento de costos se transmite a los acudientes de los estudiantes, lo que incrementa aún más el costo de la educación. Esto sin mencionar problemas de cansancio en los estudiantes y baja en su desempeño, los cuales escapan a la intención de este trabajo.

El presente trabajo busca solucionar parte de la problemática del transporte escolar. Se centra en la optimización de los recorridos de los diferentes buses de rutas escolares, de tal manera que sirva como herramienta para ayudar a las optimizaciones antes mencionadas.

Para la resolución de este problema se utilizarán las técnicas por VRP (Vehículo Routing Problem), debido a que el presente trabajo es compartido con el Ing. César Niño Méndez, se decide dividir su realización en dos, según la metaheurística implementada. Este proyecto se encarga de la implementación de la metaheurística Tabú Search para resolver el problema y la del Ing. César Niño Méndez la metaheurística Simulated Annealing.

1. OBJETIVOS

1.1. OBJETIVO GENERAL

Diseñar e implementar un aplicativo que determine un conjunto de rutas, que minimicen la distancia recorrida de los vehículos de una flota de buses escolares en su trayecto para recoger y repartir a los estudiantes de un colegio.

1.2. OBJETIVOS ESPECÍFICOS

- Construir una herramienta que facilite el ingreso de la información necesaria para la determinación de las diferentes rutas. Información de estudiantes, vehículos, colegios, depósitos, etc.
- Diseñar un aplicativo que sirva para unir los diferentes módulos y combinarlos.
- Diseñar un algoritmo para fragmentar los diferentes grupos de estudiantes.
- Diseñar el algoritmo para determinar la ruta que optimice los recorridos escolares, utilizando la metaheurística de búsqueda Tabú.
- Comparar con otras metaheurísticas para establecer la mejor.
- Realizar un aplicativo para mostrarle los resultados obtenidos al usuario. Desplegando las diferentes rutas obtenidas.

2. JUSTIFICACIÓN

En la actualidad existen varios ejemplos de ciudades que han afrontado el problema de optimización de las rutas de transporte escolar como uno de primer orden y de suma importancia. Para citar dos, se tienen.

- En las islas Baleares debido a las distancias que recorrían los diferentes estudiantes para llegar a su lugar de estudio, ocasionado a su vez a la variabilidad poblacional, se hizo necesaria la elaboración de un sistema que fuera capaz de optimizar las diferentes rutas escolares [4]. El proyecto lleva por nombre SIGTEBAL “Sistema de Información Geográfica para el diseño, gestión, análisis y planificación de rutas de transporte escolar en las Baleares”.

El objetivo del proyecto SIGTEBAL consiste en mejorar el sistema de gestión y planificación de rutas de transporte escolar en Baleares. Para ello, en primer lugar se diagnostica el sistema de rutas de transporte utilizado actualmente por la administración autonómica de Baleares en los Institutos de Educación Secundaria (IES) de la isla de Mallorca. Posteriormente se proponen rutas de transporte escolar en función de distintos criterios y finalmente, se desarrolla una aplicación informática, con funcionalidad SIG, para gestionar y planificar las rutas adecuándolas a las demandas educativas existentes. [5]

- En la localidad londinense de Wandsworth se ha venido realizando un estudio acerca de la problemática del transporte escolar [6], implementando políticas dentro de las cuales resalta la que dice, que cada estudiante no puede permanecer más de una hora en cualquier de los dos recorridos que hace diariamente, de la casa al colegio o viceversa. Este proyecto hace parte de un proyecto en todo Reino Unido conocido como “*Safe Routes to Schools*” [7] el cual busca mejorar la experiencia de movilización de los escolares, propone medios de transportes alternativos, así como mejoras en las vías.

3. ALCANCE

El problema del transporte escolar es realmente vasto. Incluye, entre otras cosas, un conjunto de normatividades para controlar el desplazamiento de los vehículos encargados de la tarea de transportar estudiantes (velocidad máxima, característica mínima de los vehículos, distintivos a utilizar, etc.), mejorar los sistemas de verificación del cumplimiento de dichas normas, controlar la idoneidad de los tripulantes de los vehículos, la señalización de la ciudad, etc.

Dentro del gran conjunto de problemas que incluye el transporte escolar, se encuentra el de mejorar la obtención de las diferentes rutas que deben tomar los diferentes vehículos para llevar a los estudiantes desde sus casas hasta el colegio y viceversa. Es éste precisamente, el objetivo de este trabajo.

Desarrollar una herramienta en la que, teniendo: la ubicación de los diferentes estudiantes, la ubicación del colegio, el lugar del garaje de los vehículos, las características de los vehículos y la malla vial sobre la cual se mueven estos vehículos; determine recorridos que minimicen el desplazamiento de los diferentes estudiantes desde y hacia sus casas.

La ubicación de los estudiantes viene dada por un identificador de cada estudiante y la dirección de su residencia. La ubicación del colegio y del garaje, incluye la dirección del colegio. Del vehículo se necesita saber ante todo, un identificador y su capacidad. Todas estas ubicaciones tienen que caer dentro de la malla vial que se tiene.

No es el propósito de este trabajo realizar una herramienta de geocodificación, por lo que se obvia el paso de tener la ubicación de los estudiantes en forma de sus direcciones, a su representación en coordenadas que se puedan ubicar en la malla vial. Esto no quiere decir que no se realice dicho paso, sino que éste no se efectúa de manera automática.

Una vez teniendo la ubicación de los estudiantes dentro de la malla vial (en coordenadas dentro de dicha malla), se procede a la determinación de los diferentes segmentos en los que se dividen los estudiantes. Cada segmento generado, debe estar relacionado con un vehículo específico, por lo que no puede superar la capacidad de dicho vehículo. Al final de este paso se debe tener que, cada estudiante pertenece a un segmento, cada segmento tiene asignado un vehículo, el número de estudiantes dentro de este segmento no supera la capacidad del vehículo asignado a él.

Posteriormente se procede a la determinación de rutas que optimicen el recorrido del vehículo dentro de su segmento correspondiente.

Para finalizar se procede a dibujar las diferentes rutas de todos los vehículos dentro de la malla vial.

El aplicativo puede ser utilizado por los directivos de colegios para ayudar al diseño de las rutas escolares necesarias para el transporte de sus estudiantes.

4. MARCO TEÓRICO

4.1. TRANSPORTE URBANO

4.1.1. Transporte escolar. Como era de esperarse, los estudiantes han ido en aumento en los últimos años, en Bogotá por ejemplo se tiene proyectado que el crecimiento anual de población en edad escolar sea del 2.3% para el 2013 (Ver Tabla 1), con lo que alcanzaría el 1.992.307 para ese año.

De igual manera el número de estudiantes con necesidad de movilizarse habrá aumentado, como lo ha estado haciendo en los últimos años. Existe un gran grupo de estudiantes que deben desplazarse hacia otras localidades diferentes a las que viven, por lo que la necesidad de transporte para estos estudiantes es imperativa (Ver. Figura 1, para estratos 1 y 2). Existe a su vez estudiantes que pese a movilizarse dentro de su localidad tienen que recorrer más de dos kilómetros para llegar a su centro educativo.

Tabla 1. Proyecciones de población en edad escolar. 2002-2013 [8].

Año	PEE Total	Crecimiento anual	PEE Estratos 1 y 2	Crecimiento anual	PEE Estrato 3	Crecimiento anual
2002	1.540.798	2,0%	710.615	2,2%	637.511	1,8%
2003	1.573.246	2,1%	727.412	2,4%	648.907	1,8%
2004	1.607.601	2,2%	745.526	2,5%	660.734	1,8%
2005	1.644.015	2,3%	765.080	2,6%	673.013	1,9%
2006	1.682.651	2,4%	786.209	2,8%	685.766	1,9%
2007	1.723.695	2,4%	809.066	2,9%	699.018	1,9%
2008	1.767.355	2,5%	833.825	3,1%	712.795	2,0%
2009	1.813.864	2,6%	860.682	3,2%	727.124	2,0%
2010	1.863.482	2,7%	889.859	3,4%	742.039	2,1%
2011	1.905.010	2,2%	913.089	2,6%	755.254	1,8%
2012	1.947.933	2,3%	937.174	2,6%	768.853	1,8%
2013	1.992.307	2,3%	962.152	2,7%	782.848	1,8%

Figura 1. Alumnos transportados en rutas escolares, 2003 [9].

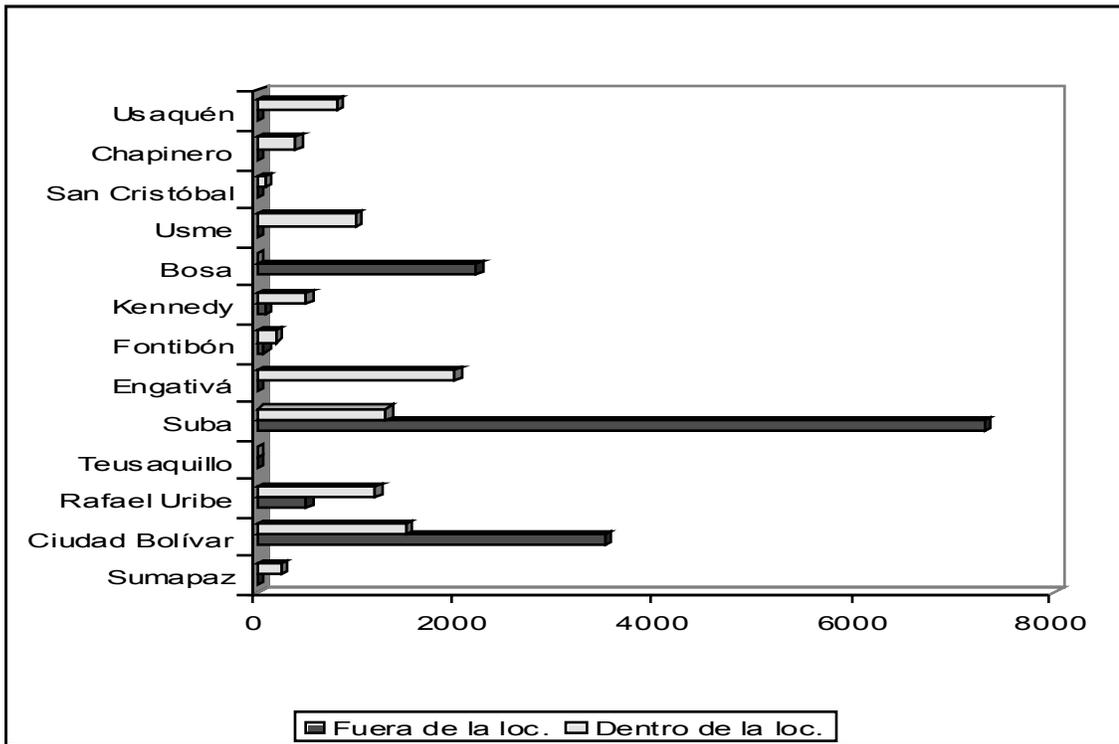
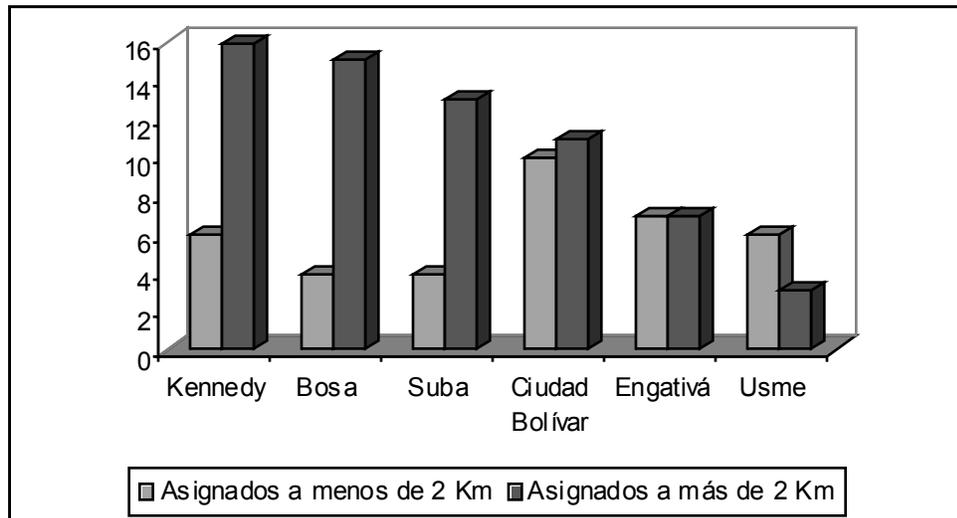


Figura 2. Inscritos nuevos y su asignación en instituciones educativas distritales – 2003 (miles de alumnos) [10].



4.1.2. Entidades reguladoras y normatividad.

4.1.2.1. Ministerio de Transporte. Según las siguientes leyes y reglamentaciones, es el ministerio de transporte el encargado de administrar todo lo relacionado con el transporte, incluido el escolar.

- Artículo 189, numeral 11 de la Constitución Política de Colombia.
- Ley 105 de 1993.
- Ley 336 de 1996.

En virtud de las facultades legales conferidas por las anteriores leyes (y otras), el Ministerio de Transporte ha proferido resoluciones que normalizan el transporte escolar en los diferentes entes territoriales de la nación. Entre los cuales cabe destacar.

4.1.2.1.1. Resolución 1122 de 2005.

ARTÍCULO PRIMERO.

...

a) Tipos de alarma:

Tipo 1: Dispositivo sonoro que se activa una vez el vehículo alcance los 61 km/h (sesenta y un kilómetros por hora) y se desactiva solamente cuando se alcance una velocidad menor a 58 km/h (cincuenta y ocho kilómetros por hora).

...

e) Tipo de alarma de acuerdo con la modalidad de servicio:

...

- Los vehículos de transporte público especial que movilicen pasajeros únicamente dentro de las zonas urbanas, los vehículos particulares autorizados para la prestación del servicio de transporte escolar y los vehículos de propiedad de los establecimientos educativos que prestan el servicio de transporte escolar deberán instalar la alarma tipo 1.

4.1.2.1.2. Decreto Número 3366 de 2003.

CAPÍTULO X

SANCIONES A LOS PROPIETARIOS DE VEHÍCULOS PARTICULARES DE TRANSPORTE ESCOLAR.

ARTÍCULO 35.- Serán sancionados con amonestación escrita, los propietarios de vehículos particulares de transporte escolar, que incurran en las siguientes infracciones:

- a) No reportar ante la autoridad que le otorgó el servicio, los cambios de domicilio.
- b) No mantener el vehículo en óptimas condiciones de comodidad y aseo.

ARTÍCULO 36.- Serán sancionados con multa de uno (1) a tres (3) salarios mínimos mensuales legales vigentes los propietarios de vehículos particulares de transporte escolar, que incurran en las siguientes infracciones:

- a) Prestar el servicio de transporte escolar sin portar el permiso expedido por la autoridad municipal competente o con este vencido.
- b) Prestar el servicio de transporte escolar sin portar los distintivos exigidos para la operación.
- c) Prestar el servicio sin llevar el adulto acompañante.
- d) Prestar el servicio sin contar con el sistema de comunicaciones en perfecto estado de funcionamiento.
- e) No mantener vigentes los seguros de responsabilidad civil contractual y extra contractual exigidos en las disposiciones vigentes.
- f) No mantener el vehículo en óptimas condiciones de seguridad.

CAPÍTULO XI

SANCIONES A LAS ENTIDADES EDUCATIVAS CON EQUIPOS PROPIOS O EMPRESAS PRIVADAS CON EQUIPOS PROPIOS DEDICADOS AL TRANSPORTE DE SUS ESTUDIANTES O ASALARIADOS

ARTÍCULO 37.- Serán sancionadas con amonestación escrita, las entidades educativas con equipos propios o empresas públicas o privadas con vehículos propios dedicadas al transporte especial, que incurran en las siguientes infracciones:

- a) Carecer de un programa y sistema de mantenimiento preventivo para los vehículos.
- b) Permitir la prestación del servicio en vehículos conducidos por personas no idóneas.

- c) No contar, para la prestación del servicio con la presencia de un adulto acompañante.
- d) No mantener vigentes pólizas de responsabilidad civil contractual y extracontractual que ampara los riesgos inherentes al transporte.

4.1.2.2. Secretarías de Tránsito y Transporte. Las secretarías de transporte de las diferentes entidades territoriales, y por medio de potestad emanada desde el Ministerio de Transporte, son autónomas y están obligadas a publicar resoluciones para la regulación del transporte en sus respectivas jurisdicciones. Como ha sucedido en el pasado y debido a los trágicos acontecimientos del bus escolar del Agustiniense, la Secretaría de Tránsito y Transporte de Bogotá D.C (STT de ahora en adelante). ha tomado la delantera en cuanto a normatividad para transporte escolar específicamente. Publicó la resolución número 50 de 2005 “Por el cual se establecen medidas especiales y se adoptan elementos de prevención y seguridad para el transporte escolar en la Ciudad de Bogotá D.C.” que recoge, entre otras, las normas ya publicadas por el Ministerio de Transporte.

A parte de la mencionada resolución, la STT ha lanzado el programa “Ruta Pila” dentro del que busca: “Propiciar espacios, reflexiones y ejercicios de corresponsabilidad para promover el conocimiento y cumplimiento de las normas de seguridad en el transporte escolar por parte de los empresarios del sector, conductores, directivas de las instituciones, padres de familia y estudiantes.” [11]

4.2. ALGORITMOS Y COMPLEJIDAD

4.2.1. Algoritmos deterministas. Son aquellos en que en cada paso se sabe cual será el paso siguiente.

Otra definición, más formal, sería:

Un problema Π se dice que es un problema de decisión si las únicas posibles soluciones del mismo son «sí» o «no». Entonces llamamos

D_{Π} al conjunto de todas sus posibles entradas y $Y_{\Pi} \subseteq D_{\Pi}$ al conjunto de todas las entradas que ofrecen resultado afirmativo.

Un algoritmo se dice que es un algoritmo determinista si para cada entrada del problema nos proporciona una solución. En un problema de decisión Π , dada una instancia cualquiera $I \in D_{\Pi}$, un algoritmo determinista nos debe decir si $I \in Y_{\Pi}$ o $I \notin Y_{\Pi}$ ” [12]

4.2.2. Algoritmos no deterministas. Son aquellos en los que al finalizar cada paso se cuenta con un conjunto de pasos alternativos, si se quiere obtener la respuesta correcta se hace necesario recorrer todas estas opciones.

Una definición formal sería:

...un algoritmo no determinista está compuesto de dos etapas. Dada una instancia del problema $I \in D_{\Pi}$, la primera etapa (etapa de conjetura) nos proporciona una estructura S . Esta estructura S junto con la instancia I constituyen la entrada de la segunda etapa (etapa de chequeo), en la que se comprueba si con la estructura S la instancia $I \in Y_{\Pi}$. Diremos que un algoritmo [no] determinista “resuelve” el problema Π , si para cualquier instancia I del mismo se tienen las dos propiedades siguientes:

1. Si $I \in Y_{\Pi}$ entonces en la primera etapa se puede generar una estructura S de manera que con ella se garantiza en la segunda etapa que I produce respuesta afirmativa.
2. Si $I \notin Y_{\Pi}$ no se puede generar en la primera etapa ninguna estructura S que nos lleve, en la segunda etapa, a la respuesta afirmativa para I .

Al contrario de lo que ocurre con los algoritmos deterministas, los algoritmos no determinísticos no resuelven el problema, sino que permiten comprobar si cierta estructura asociada a una entrada determinada del problema es realmente una solución de dicha entrada. [13]

4.2.3. Máquina de Turing. Máquina imaginaria ideada por Alan Turing, matemático británico, que debido a ser pionero en la teoría del ordenador es considerado el padre de la Ingeniería Informática.

“Una máquina de Turing es un autómatas que se mueve sobre una secuencia lineal de datos. En un momento determinado la máquina puede leer un único dato de la anterior secuencia y realizar ciertas acciones en base a una tabla que tiene en cuenta su «estado» actual (interno) y el último dato leído.”[14]

Basado en esta definición, Turing demostró que existían problemas que una máquina no podía resolver.

La entrada de una máquina de Turing viene determinada por el estado actual y el símbolo leído, siendo el cambio de estado, la escritura de un nuevo símbolo y el movimiento las acciones a tomar en función de una entrada. En el caso de que para cada par estado y símbolo posible exista a lo sumo una posibilidad de ejecución, diremos que es una máquina de Turing determinista, mientras que en el caso de que exista al menos un par (estado, símbolo) con más de una posible combinación de actuaciones diremos que se trata de una máquina de Turing no determinista. [15]

4.2.4. Complejidad L. “Es el conjunto de los problemas de decisión (los cuales, someramente, son aquellos cuya respuesta es sí o no), que se pueden resolver en un espacio $\log(n)$ (sin contar el tamaño de la entrada), siendo n el tamaño de la entrada, por una máquina de Turing determinista tal que la solución si existe es única.” [16].

4.2.5. Complejidad NL. “Es el conjunto de los problemas de decisión que pueden ser resueltos en espacio $\log(n)$ (sin contar el tamaño de la entrada), donde n es el tamaño de la entrada, por una máquina de Turing no determinista tal que la solución si existe es única” [17]

4.2.6. Complejidad P. “...en teoría de la complejidad, la clase de complejidad de los problemas de decisión que pueden ser resueltos en tiempo polinómico calculado a partir de la entrada por una máquina de Turing determinista es llamada P” [18] Es un superconjunto de la complejidad L. Se dice que un algoritmo tiene tiempo de ejecución polinómico cuando, utilizando las variables de entrada en una fórmula polinomial, se puede determinar el tiempo de ejecución del mismo.

4.2.7. Complejidad NP. “NP es el acrónimo en inglés de *Polinómico no determinista (Non-Deterministic Polynomial-time)*. Es el conjunto de problemas que pueden ser resueltos en tiempo polinómico por una máquina de Turing no determinista.” [19]

4.2.8. Complejidad NP-Completo.

...la clase de complejidad NP-completo es el subconjunto de los problemas de decisión en NP tal que todo problema en NP se puede reducir en cada uno de los problemas de NP-completo. Se puede decir que los problemas de NP-completo son los problemas más difíciles de NP y muy probablemente no formen parte de la clase de complejidad P. La razón es que de tenerse una solución polinómica para un problema de NP-completo, todos los problemas de NP tendrían también una solución en tiempo polinómico.

Como ejemplo de un problema NP-completo está el problema de la suma de conjuntos que se puede enunciar como sigue: dado un conjunto S de enteros, ¿existe un subconjunto no vacío de S cuyos elementos sumen cero? Es fácil verificar si una respuesta es correcta, pero no se conoce mejor solución que explorar todos los $2^n - 1$ subconjuntos posibles hasta encontrar uno que cumpla con la condición. [20]

4.2.9. Complejidad NP-Complejo o NP-Hard.

En teoría de la complejidad computacional, la clase de complejidad NP-hard es el conjunto de los problemas de decisión que contiene los

problemas H tales que todo problema L en NP puede ser transformado polinomialmente en H . Esta clase puede ser descrita como conteniendo los problemas de decisión que son al menos tan difíciles como un problema de NP. Esta afirmación se justifica porque si podemos encontrar un algoritmo A que resuelve uno de los problemas H de NP-hard en tiempo polinómico, entonces es posible construir un algoritmo que trabaje en tiempo polinómico para cualquier problema de NP ejecutando primero la reducción de este problema en H y luego ejecutando el algoritmo A .

Asumiendo que el lenguaje [problema] L es NP-completo,

1. L está en NP
2. L' en NP, $L' \leq L$

En el conjunto NP-Hard se asume que el lenguaje [problema] L satisface la propiedad 2, pero no la la (sic) propiedad 1.

La clase NP-completo puede definirse alternativamente como la intersección entre NP y NP-hard. [21]

4.3. ANÁLISIS DEL PROBLEMA VRP

El problema de enrutamiento de vehículos está basado en dos problemas bien estudiados [22], como lo son.

4.3.1. Problema del agente viajero. Este problema es uno de lo más conocidos en el área de los problemas de búsqueda. Grosso modo trata de un vendedor, que por razón de su trabajo, debe recorrer diferentes ciudades para ofrecer sus productos, dicho trayecto lo debe realizar sin pasar dos veces por la misma ciudad; el problema consiste en encontrar la ruta óptima para el desplazamiento del agente viajero.

Dos definiciones más formales son:

- (Traducción libre) “Dado un conjunto de n nodos y las distancias para cada par de nodos, encontrar un trayecto que minimice el total de la

distancia visitando cada nodo exactamente una vez. La distancia del nodo i al nodo j es la misma que la del nodo j al nodo i ” [23]

- “Dado un grafo conexo K_n con pesos en sus aristas C_{uv} , encontrar el *Circuito de Hamilton* más corto en K_n ” [24]

El problema del agente viajero tiene variaciones, cada una de esas nuevas definiciones se conoce como instancia; una de las más conocidas es la definición asimétrica del problema del agente viajero, cuya única diferencia con la anterior (también conocida como simétrica) es que la distancia de un nodo a otro no es conmutativa, o sea la distancia del nodo i al nodo j no es la misma que la del nodo j al nodo i [25], o simplemente cuando en vez de un grafo se tiene un digrafo. [26]

Una variante de este problema es el conocido como multidimensional, en el que existe, no uno, sino varios caminos para llegar de una ciudad a otra, el problema ahora consiste en “encontrar la trayectoria óptima no sólo considerando la distancia que hay de una ciudad a otra, sino también en seleccionar un camino corto de entre todas las posibilidades.” [27]

La instancia de Multi Problema de agente viajero (conocido también como MTSP *Multi Traveling Salesman Problem*) es la generalización de las instancias anteriores, en esta se cuenta con un conjunto de vendedores (y no con uno solo) y un lugar de llegada, estos vendedores deben visitar un subconjunto de los clientes comenzando y terminando en un depósito común. Cada cliente debe ser visitado exactamente una vez por un vehículo. [28]

Debido a que cada ciudad puede visitarse una sola vez, el problema del agente viajero cae dentro de los problemas de optimización combinatoria [29], donde la solución se encuentra dentro del conjunto de permutaciones de todas las ciudades, cada posibilidad debe tener un valor, el cual puede estar basado en el número de kilómetros, facilidad de acceso, o cualquier otro ítem que sirva para determinar la conveniencia o no de un trayecto con respecto a otro. A medida que el número de ciudades va en aumento así también el número de cálculos necesarios para hallar la solución óptima, lo

que enmarca el problema de enrutamiento de vehículos en la categoría de los NP-Complejo, suponiendo esto que en la medida que el número de ciudades crezca, el espacio en memoria (y el tiempo que toma encontrar la solución) aumenta más rápidamente.

4.3.2. Problema de la mochila (Bin Packing Problem). Se tiene un conjunto de objetos y un recipiente (o mochila), el problema consiste en maximizar la cantidad de objetos que se pueden almacenar dentro del recipiente, generalmente tanto los objetos como el recipiente son de forma rectangular. [30]

Formalmente se puede definir:

- (Traducción libre) “Dado: n objetos para ser ubicados en recipientes de capacidad L . El objeto i requiere l_i unidades de capacidad de recipiente. Objetivo: determinar el mínimo número de recipientes necesarios para acomodar todos los n objetos”
“El problema de la mochila es NP completo cuando se formula como un problema de decisión. Cuando se hace como un problema de optimización es considerado como NP complejo” [31]
- (Traducción libre) “Matemáticamente la formulación del problema podría ser como sigue: dado un conjunto finito de elementos $E = \{e_1 \dots e_n\}$ con pesos asociados $W = \{w_1 \dots w_n\}$ tal que $0 \leq w_i \leq w(\text{recipiente})$. Dividir E en N subconjuntos tal que la suma de pesos en cada partición sea el máximo $w(\text{recipiente})$ y que N sea el mínimo. [32]

4.3.3. Definición de VRP. El problema que afronta VRP es el de diseñar un conjunto de rutas óptimas para una flota de vehículos para servir a un conjunto de usuarios. A diferencia del MTSP el VRP cuenta con requerimientos de servicio que deben ser cumplidos, la flota de vehículos es limitada y los vehículos tienen capacidad limitada.

Como se menciona arriba el VRP se fundamenta en los problemas de TSP y BPP, por lo que al igual que estos cae dentro de los problemas NP complejos y de los problemas de optimización combinatorial. La solución

de este problema requiere la utilización de heurísticas, de técnicas de optimización estándar o de metaheurísticas, como Simulated Annealing y Tabu Search, las cuales son las más utilizadas en este tipo de problemas.

4.3.3.1. Definición formal de VRP. Para la definición formal se remite a la expuesta en la página de Internet *VRP Web* en la cual hay una completa bibliografía acerca del tema.

El VRP es un problema combinatorial cuyo conjunto base son las aristas de un grafo $G(V, E)$. La notación usada para este problema es como sigue:

- $V = \{v_0, v_1, \dots, v_n\}$ es el conjunto de vértices donde:
 - Se supone que el depósito está ubicado en v_0 .
 - $V' = V / \{v_0\}$ se tomará como el conjunto de n ciudades.
- $A = \{(v_i, v_j) | v_i, v_j \in V; i \neq j\}$ son el conjunto de arcos.
- C es una matriz de costos o distancias no negativas c_{ij} entre usuarios v_i y v_j .
- d es un vector de demandas de los usuarios.
- R_i es la ruta para el vehículo i .
- m es el número de vehículos (todos iguales). Una ruta será asignada a cada vehículo.

Cuando $c_{ij} = c_{ji}$ para todo $(v_i, v_j) \in A$ se dice que el problema es simétrico y entonces en general podemos reemplazar A con el conjunto de aristas $E = \{(v_i, v_j) | v_i, v_j \in V; i < j\}$.

»Con cada vértices v_i en V' está asociado una cantidad q_i de artículos a ser entregados por un vehículo. Así el VRP consisten en determinar un conjunto m de rutas que minimice el costo total, empezando y finalizando en un depósito, cada vértice en V' es visitado exactamente una vez por un vehículo. [33]

El VRP que cae dentro de la solución anterior es conocido como CVRP (Capacitated Vehicle Routing Problem), y es la instancia de VRP más extendida.

4.3.3.2. Parámetros de VRP.

- Flota de vehículos: Característica de los vehículos que prestan el servicio de transporte, es homogénea cuando los buses tienen las mismas características, y heterogéneas en caso contrario.
- Forma del bien transportado: Si se conoce las características de los bienes transportados, igualmente pueden ser homogéneos o heterogéneos.
- Depósito: Lugar a donde llega los vehículos.
- Número de depósitos: Indica cuantos depósitos existen, puede ser como en el caso de Hariti, que se cuente con depositos diferentes para la entrada y salida.
- Tipo de ruta: La ruta puede ser circular, el punto de partida es el mismo de llegada, o abierta, ambos puntos son diferentes.
- Función objetivo: La función que se busca optimizar, pueden influir variables de tiempo, distancia, etc.
- Número de vehículos: Cantidad de vehículos que se debe tener una ruta, muchas veces esta variable también se quiere optimizar.
- Ventanas de tiempo: Caso en que el cliente sólo está disponible en un intervalo de tiempo y es obligación del vehículo llegar al cliente durante ese intervalo (ventana) de tiempo.
- Clientes: Se debe determinar si el usuario se mueve con el tiempo o siempre cuenta con la misma ubicación.

4.4. PROBLEMA DE ENRUTAMIENTO DE VEHÍCULOS ESCOLARES COMO VRP

El enrutamiento de transporte escolar es, por supuesto, un problema tipo VRP. El depósito sería el garaje en donde se guarda el vehículo al terminar su turno, en el caso de estar repartiendo a los estudiantes hacia sus casas; o el colegio en caso de estar recogiendo a los estudiantes para llevarlos a su centro educativo. Los nodos serían pues, los diferentes estudiantes que toman el servicio de transporte.

Se podría pensar que el problema encaja perfectamente dentro de los TSP, pero teniendo en cuenta que se tienen limitantes de capacidad en los vehículos, queda encasillado como un problema de CVRP.

Al ser este problema del tipo CVRP, es obviamente un problema NP-Complejo, por lo que no puede resolverse con algoritmos “tradicionales”, sino con algoritmos heurísticos.

4.5. HEURÍSTICA

Algoritmo que se comporta bien en la mayoría de los casos, pero no existe forma de medir la exactitud de la respuesta. Dicha respuesta se puede obtener en un tiempo computacional menor que el utilizado por aquella que es la óptima.

4.6. METAHEURÍSTICA

Las metaheurísticas son algoritmos mas avanzados, que utilizan las heurísticas para su implementación, hacen énfasis en la exploración de todo el conjunto de soluciones, para así no caer en óptimos locales, como suele suceder con las heurísticas, obteniendo así mejores respuestas.

(Traducción libre) “Las metaheurísticas, en su definición original, son métodos para orquestrar una interacción entre procedimientos de mejora local y estrategias de más alto nivel para crear un proceso capaz de escapar de los óptimos locales y realizar una búsqueda robusta de una solución” [34]

4.6.1. Búsqueda Tabú. Heurística propuesta originalmente por Fred Glover en el año de 1986, en varios casos las soluciones muy cercanas a las óptimas, sino eran las mejores. Este éxito con la búsqueda tabú (TS de ahora en adelante) aumentó su popularidad vertiginosamente en los últimos años. Uno de los problemas ejemplo más utilizados para mostrar las bondades de este algoritmo es el ya mencionado CVRP.

Algo que caracteriza esta metaheurística de las demás y la no utilización de técnicas de aleatorización, optando por el uso de “técnicas de memoria adaptativa y de estrategias especiales de resolución de problemas” [35]

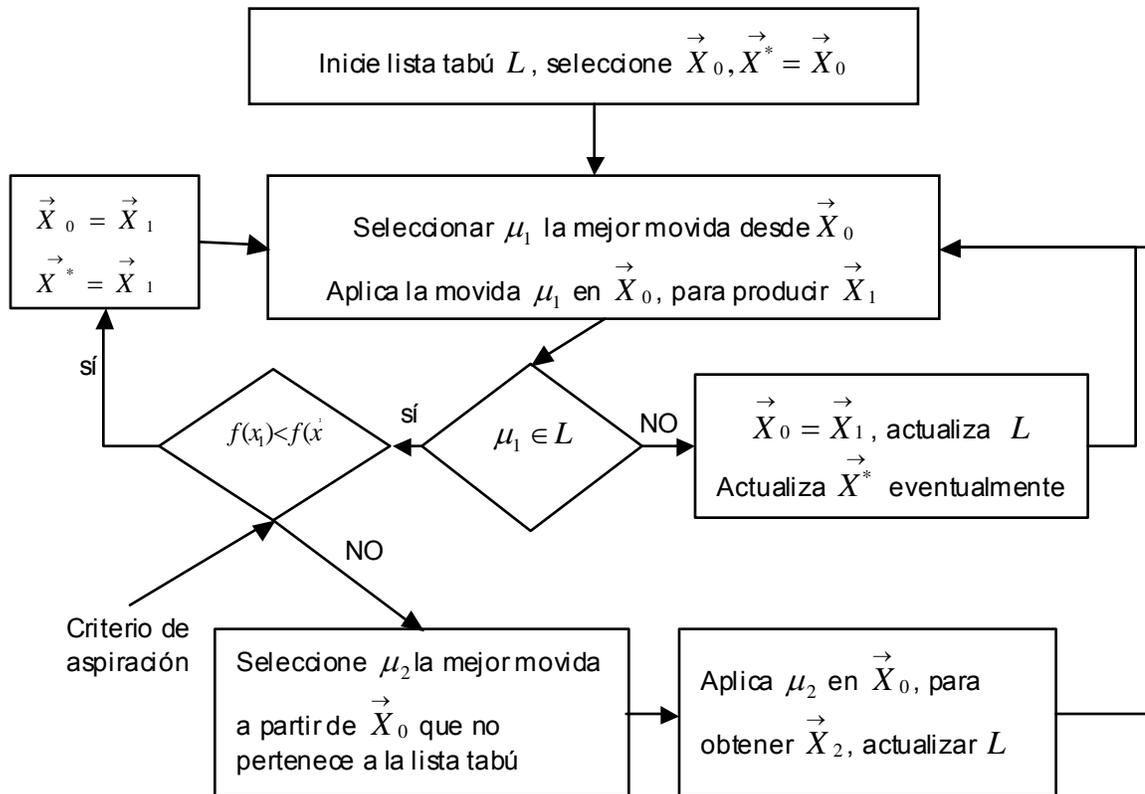
“La Filosofía de la búsqueda tabú es derivar y explotar una colección de estrategias inteligentes para la resolución de problemas, basadas en procedimientos implícitos y explícitos de aprendizaje” [36].

Grosso modo la búsqueda tabú consiste en:

Se tiene un conjunto de elementos cuya ordenación de distinta forma altera el resultado total, por lo que se quiere encontrar una ordenación de tal forma que el conjunto dé un resultado óptimo. Para ello, se halla una solución inicial. Acto seguido se hallan los distintos intercambios posibles, de las todas las posiciones de los elementos que constituyen el conjunto inicial, dichos intercambios pueden realizarse de distinta forma, la utilización de uno y otro alterará el resultado final. Cada uno de los posibles intercambios se organizan de mayor a menor según mejoren o empeoren el resultado final. Se aplica el mejor resultado, dicho intercambio se almacena en una lista conocida como tabú, para que dentro de los próximos n (donde n es el tamaño de la lista tabú, que en realidad actúa más o menos como una cola) iteraciones no se vuelva aplicar dicho intercambio, así sea el mejor, en cuyo caso se tomará el segundo de la lista, que de también estar en la lista tabú, se seguirá con el siguiente, y así sucesivamente. Existe un caso en el que

se viola la no aplicación de un movimiento que está tabú, y es cuando ese movimiento genera un resultado total, mejor que todos los encontrados hasta ese momento, dicha violación se conoce como Criterio de Aspiración. Con la lista tabú, se busca no caer en óptimos locales y explorar el más el mundo de la solución, lo que se conoce como diversificación.

Figura 3. Algoritmo Búsqueda Tabú.



4.7. DESCRIPCIÓN DEL PROBLEMA

4.7.1. Definición de términos.

- **Paradero:** Ubicación en el que es recogido y dejado el estudiante por el vehículo transportador.
- **Trayecto recogida:** Trayecto que el vehículo realiza con la finalidad de llevar a los estudiantes al centro educativo.

- **Trayecto distributivo:** Trayecto que el vehículo realiza con la finalidad de distribuir a los estudiantes en sus viviendas después de la jornada escolar.
- **Recorrido:** Trayecto que realiza el vehículo empezando en el garaje, pasando por los paraderos asignados para recoger estudiantes, hasta la institución educativa, o viceversa.
- **Recorrido libre:** Comprende el trayecto en el cual el bus se encuentra desocupado dentro del recorrido. Este puede ir desde el garaje hasta el ingreso en el vehículo del primer estudiante dentro del trayecto de recogida, o desde el último estudiante que abandona el vehículo hasta el ingreso en el garaje.
- **Ruta o Recorrido ocupado:** Trayecto en el cual el vehículo se encuentra con estudiantes a bordo.

4.7.2. Problema de enrutamiento de transporte escolar.

Determinar el conjunto de recorridos que minimicen la distancia a desplazarse por los diferentes estudiantes en las rutas, utilizando sólo los vehículos estrictamente necesarios, es decir, aquellos cuyas capacidades sumadas sea igual al número de estudiantes o a lo sumo quedar un único bus con algunos puestos desocupados.

4.7.3. Definición matemática del problema. Determinar el conjunto de recorridos óptimos que minimicen el número de elementos y la distancia total del viaje de una flota de vehículos de transporte escolar, los cuales sirven a un centro educativo en la movilización de sus estudiantes desde y hasta sus residencias. Se parte de un número mínimo de buses escolares que suplen a la totalidad de estudiantes.

Se trabaja con un grafo dirigido que se obtiene del análisis de la malla vial y la ubicación geográfica de cada uno de los estudiantes. La malla vial tiene cómo atributos el sentido de las vías y su longitud, permitiendo generar un grafo dirigido entre todos los posibles puntos de parada de la ruta.

El problema se resume en:

1. Se tiene un grafo dirigido $G = \{V, A\}$ en donde:
 - $V = \{0, 1, \dots, n\}$ es el conjunto de vértices.
 - $\{1, 2, \dots, n-1\}$ conjunto de vértices que representa a los estudiantes.
 - 0 representa el vértice de partida.
 - n representa el punto de llegada.
 - $A = \{(vi, vj) \mid vi \in V, vj \in V; i \neq j\}$ es el conjunto de arcos.

2. c es una matriz con los costos (distancias) c_{ij} entre los vértices v_i y v_j .
 - Hay vértices en donde el costo que tiene ir en un sentido, no es el mismo si el recorrido es en sentido contrario.
 - $c_{ij} > 0$ cuando existe un arco entre i y j , de lo contrario $c_{ij} = \max$. Cuando $i = j$ entonces $c_{ij} = 0$.

Matriz c

	0	1	2	3	4	5	6	7	8
0	0	10	11	18	16	20	25	50	10
1	20	0	15	30	16	16	11	12	18
2	20	15	0	9	18	16	15	16	30
3	16	16	15	0	3	18	16	20	9
4	20	18	16	10	0	3	18	15	15
5	16	30	18	18	10	0	3	18	18
6	30	15	18	15	18	10	0	16	20
7	9	16	30	16	30	18	30	0	50
8	3	18	9	18	9	30	9	3	0

3. $F = \{1, \dots, k\}$ es la flota de vehículos, en donde k es el número total de vehículos.
4. D es un vector de demanda en donde D_i indica la demanda para el nodo i . Es decir cuantos estudiantes se recogerán en un punto determinado.

Vector D

i	D
0	2
1	3
2	1
3	1
4	2
5	1
6	1
7	3
8	2

5. C es un vector de capacidad en donde C_i indica la capacidad del vehículo i .

Vector C

i	C
0	4
1	4

6. x es una matriz de flujo que indica si un vehículo atraviesa un arco en la solución encontrada. Toma los valores de 1 si el arco $(i, j) \in A$ y 0 de lo contrario.

Matriz x

	0	1	2	3	4	5	6	7	8
0		1		1					
1					1				
2						1			
3			1						
4								1	
5							1		
6									1
7									1
8									

7. Los vehículos se desplazan a una velocidad promedio por la malla vial v_{ruta} .

8. El tiempo que el vehículo permanece libre se denota t_{libre} .
9. El tiempo que el vehículo permanece ocupado se denota t_{ruta} .
10. La duración total de la ruta es $t_{recorrido} = t_{ruta} + t_{libre}$.
11. o es una matriz que contiene los identificadores de las rutas. Esta matriz es similar a la matriz x pero en vez de contener 1 cuando un arco está en la solución óptima, o_{ij} contiene el identificador del vehículo que sirve al nodo j proveniente del nodo i .

Matriz o

	0	1	2	3	4	5	6	7	8
0		1		2					
1					1				
2						2			
3			2						
4								1	
5							2		
6									2
7									1
8									

4.7.1. Función objetivo.

$$\min \sum_{i \in V} \sum_{j \in V} c_{ij} x_{ij}$$

s.a.

$$(1) \quad \sum_{i \in V} x_{ij} = 1 \quad \forall j \in V \setminus \{0, n\}$$

$$(2) \quad \sum_{j \in V} x_{ij} = 1 \quad \forall i \in V \setminus \{0, n\}$$

En (1) se estipula que la sumatoria de todos los elementos de una fila es igual a uno sin tener en cuenta la primera y la última fila. Lo mismo sucede

con las columnas (2). Esta restricción indica que a un vértice solo le llega un vehículo y de él solo puede salir un vehículo, lo cual permite definir que todas las rutas son disyuntas.

$$(3) \quad \sum_{i \in V} x_{i0} = 0$$

$$(4) \quad \sum_{j \in V} x_{nj} = 0$$

La sumatoria de la primera columna tiene que ser igual a cero (3), esto indica que ningún vehículo regresa al punto de origen.

La sumatoria de la última fila debe ser cero (4), lo que muestra que ningún vehículo va a ir desde el punto de destino a cualquier otro.

$$(5) \quad \sum_{j \in V \setminus \{n\}} x_{0j} = K$$

$$(6) \quad \sum_{i \in V \setminus \{0\}} x_{in} = K$$

La sumatoria de todos los campos de la primera fila sin el último, indica el número total de rutas que salen hacia algún paradero (5). De igual manera, la sumatoria de la última columna sin el primer elemento, indica el número de rutas que llegan al destino final de la ruta (6).

$$(7) \quad x_{0n} = 0$$

La restricción (7) evita que una ruta pueda ir directamente del lugar de origen al destino sin pasar por algún paradero.

$$(8) \quad \sum_{i \in S} \sum_{j \in S} x_{ij} \geq r(S) \quad \forall S \subseteq V \setminus \{0, n\}, S \neq \emptyset$$

	0	1	2	3	4	5	6	7	8
0		1		1					
1					1				
2						1			
3			1						
4								1	
5							1		
6									1
7									1
8									

La restricción (8) es aplicable al caso en el que todos los vehículos tienen la misma capacidad. S es un conjunto de clientes cualquiera el cual es atravesado por un número de arcos no menor que $r(S)$ (indica el número mínimo de vehículos que son necesarios para servir S),

Por último se tiene una restricción para cuando las capacidades de los vehículos son diferentes, para tal caso, lo que se busca es que en el momento en que en la matriz o aparezca el vehículo k , se determina cual es la demanda para ese vértice acumulándose. La sumatoria debe ser menor o igual a la capacidad del vehículo k .

5. SOLUCIÓN PROPUESTA

5.1. LEVANTAMIENTO DE REQUERIMIENTOS

La solución del problema de enrutamiento requiere la implementación de los requerimientos abajo mencionados.

- Una interfaz a través de la cual se pueda interactuar con el cliente, encargada de leer la información de: los estudiantes (ubicación, identificación, servicio), la del colegio (ubicación e identificación) y las de los vehículos de la flota de buses con las que cuenta el colegio, la malla vial del municipio en el cual se encuentra el centro educativo.
- Una herramienta que despliegue la ruta final sobre la malla vial del municipio, identificando sitios donde detenerse y el orden en que debe realizarse el viaje.

5.1.1. Identificación de actores.

- Estudiante. Es al fin y al cabo el usuario primario del servicio, es él el que toma la ruta y al que se le quiere mejorar su medio de transporte disminuyendo la distancia por él recorrida. Su responsabilidad es informar su identificación, su ubicación y el tipo de servicio a prestar. También es responsabilidad tomar el bus en la parada para ser recogido y dejado por el bus.
- Colegio. Es el sitio a donde llegarán los diferentes estudiantes y de donde saldrán los días escolares. Tiene como responsabilidad suministrar su identificación y su ubicación en la malla vial del municipio en donde se prestará el servicio. Es el usuario del aplicativo y es él que va a interactuar con él, por lo que también es su responsabilidad recoger la información suministrada por todos los estudiantes que tomarán el servicio de transporte escolar e ingresarlos al sistema.

- Vehículo. Es el encargado de transportar desde la parada origen hasta el colegio en el trayecto recogida, y del colegio a la parada en el trayecto distributivo. Es su responsabilidad suministrar la información de su identificación (placa) y su capacidad.

5.2. DISEÑO

5.2.1.1. Módulo de lectura de entradas. El aplicativo Hariti tiene como entrada un conjunto de Shapes (archivos .shp generados por herramientas como ArcGIS), por lo que se hace necesario un módulo que tome dichos archivos y los transforme de tal manera que puedan ser entendidos por el aplicativo. Este paso es relevante ya que la geocodificación arroja un shape como respuesta. Dicho proceso de geocodificación es un proceso manual que transforma la dirección de un estudiante a su respectiva coordenada espacial, más adelante se volverá a ella.

5.2.1.2. Módulo de fragmentación o segmentación. Módulo encargado de tomar los diferentes estudiantes y dividirlos según las diferentes capacidades de los vehículos. Con la finalidad de reducir el problema, se dividen los estudiantes en fragmentos en donde a cada fragmento le será asignado un vehículo y este vehículo el responsable de transportar a cada estudiante del fragmento.

El proceso de segmentación en Hariti consiste en dividir a los diferentes estudiantes de un plantel educativo en grupos según sus direcciones de vivienda, de tal manera que cada uno pueda ser atendido por un bus. Debido a lo anterior el grupo n , atendido por un bus de capacidad m , no puede ser de más de m estudiantes.

El algoritmo de segmentación combina la heurística *KMeans* con la de *Metrópoli Montecarlo*. La primera es utilizada para determinar los diferentes segmentos en que se divide la población estudiantil según su dirección de vivienda y la segunda para determinar una buena asignación de capacidades a cada uno de los segmentos hallados.

5.2.1.2.1 Descripción del algoritmo. Los pasos de ejecución del algoritmo son los siguientes.

1. Se carga la lista de estudiantes. Cada uno tiene su respectiva ubicación (x, y) la cual indica la dirección de su vivienda.
2. Se aplica la heurística KMeans para segmentar los diferentes estudiantes. El número de segmentos viene dado por la cantidad de buses disponibles. Al terminar este paso se deben tener n segmentos (donde n es el número de buses) y cada segmento debe tener asignado un conjunto de estudiantes (Figura 4).

Se aplica la heurística *Kmeans* para segmentar el grupo de estudiantes. El número de segmentos que se obtiene al final del proceso viene dado por la cantidad de buses (n) que suplen a todos los alumnos, los cuales a su vez son incluidos en cada uno de los segmentos generados (ver Figura 4).

3. A cada uno de los segmentos generados se le asigna aleatoriamente un vehículo, junto con la capacidad del mismo (Figura 5).
4. Se eliminan todos los estudiantes de todos los segmentos, quedando únicamente los diferentes centroides de cada uno (Figura 5).
5. Se reasignan los estudiantes a los diferentes segmentos, teniendo en cuenta el que tenga su centroide más cerca y respetando la capacidad del segmento específico. Si un estudiante tiene el segmento con centroide c_{ij} más cerca de él que cualquier otro; y siempre y cuando dicho segmento tenga capacidad; el estudiante se le es asignado. En caso contrario, se le asignará al segundo centroide más cercano, y así sucesivamente. El orden en que son asignados los estudiantes también es aleatorio, para garantizar igualdad de condiciones para todos los estudiantes.

Figura 4. Segmentación con KMeans (1000 estudiantes – 30 segmentos).

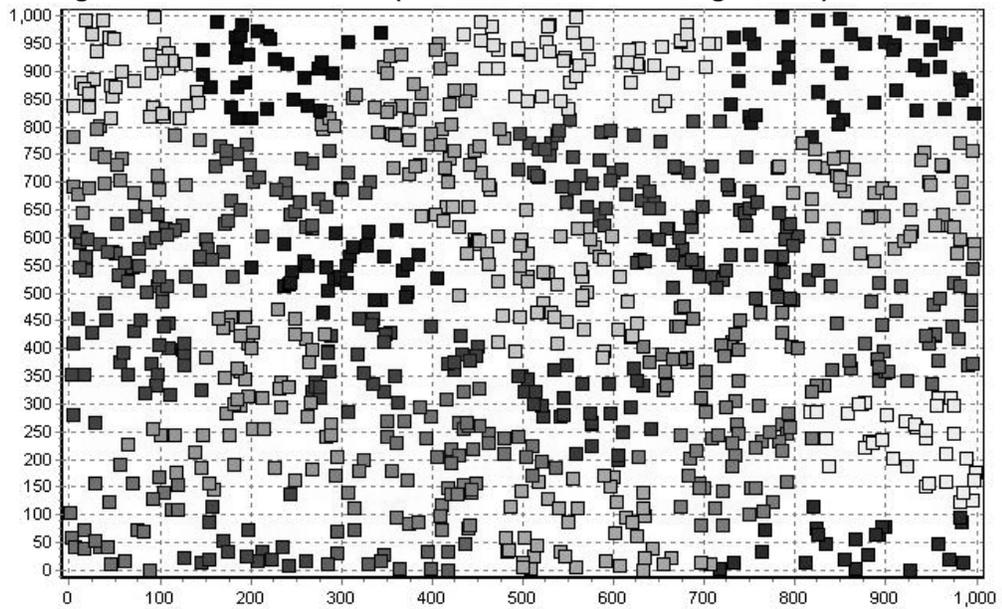
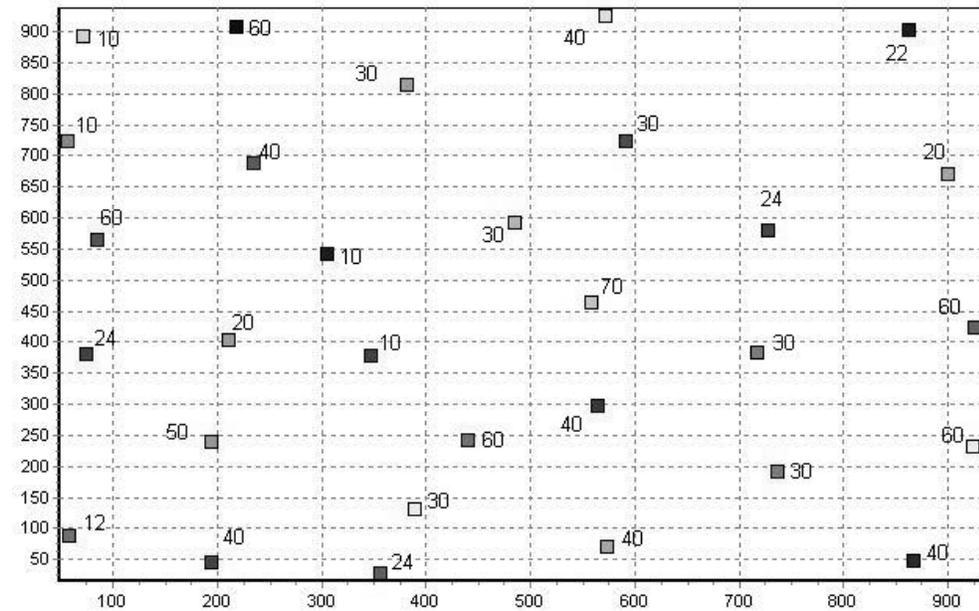


Figura 5. Centroides después del proceso de KMeans.



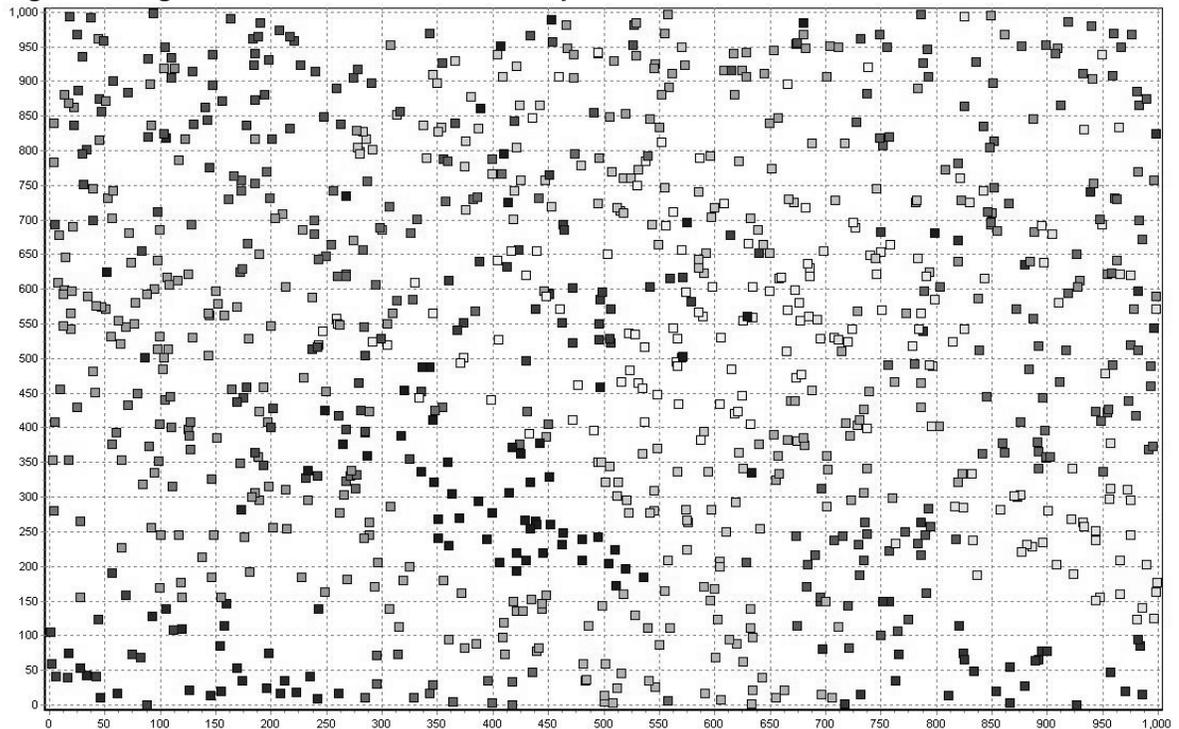
6. Se realiza la sumatoria de las distancias euclidianas (*sumaDistancias*) de los estudiantes a sus respectivos centroides.

Dentro de cada segmento se realiza la sumatoria de las distancias euclidianas desde cada punto a su centroide. Luego se suman todos los resultados de todos los segmentos.

7. Se repiten por un número determinado de veces los pasos desde el 3°. Si *sumaDistancias* es mayor que la nueva suma de distancias, entonces se actualiza *sumaDistancias* con la nueva sumatoria de distancias y se almacena la nueva distribución de estudiantes y capacidades.

Al finalizar el algoritmo se obtienen segmentos con capacidades correspondientes a la de los buses, y con los estudiantes asignados a los mismos (Figura 6).

Figura 6. Segmentación con KMeans con capacidad.



El número de iteraciones del algoritmo determina la precisión del mismo. En todas las pruebas realizadas, después de 750 iteraciones no se obtenían mejoras significativas, por lo que 1000 iteraciones es un buen número de repeticiones.

5.2.1.2.1. Diagrama de clase.

Ver Anexo A.

5.2.1.3. Módulo geocodificación. En este módulo se trabaja con la malla vial y la base de datos que tiene la información de los estudiantes. ArcGIS 9.1 (Herramienta utilizada) cuenta con un geocodificador que permite obtener las coordenadas geográficas de una dirección sobre una malla vial.

Debido a que ArcGIS está diseñado para trabajar con direcciones de ciudades norteamericanas, es necesario adaptar la nomenclatura de los estudiantes a dicho formato. Las ciudades norteamericanas utilizan un nombre para identificar cada calle acompañado del número de la casa, el tipo de vía y en algunas ocasiones el código postal y el nombre de la ciudad. (Tabla 2)

Tabla 2. Estructura dirección Norteamérica.

Ave	Aurora	572	02457812	Salt Lake City
Tipo de vía	Nombre calle	Número casa	Cod. Postal	Ciudad

De esta manera, se puede hacer una geocodificación a partir del nombre de la calle con el tipo de vía, estos dos datos reducen el campo de búsqueda a una sola calle que tiene como longitud una cuadra, el dato adicional es el número de la casa, él cual si es par está a un costado y si es impar está al otro de la vía.

En el caso colombiano es necesario adaptar la nomenclatura para poder hacer un buen uso del geocodificador. Con respecto a la denominación de la malla vial de Bogotá, se tiene el formato como se muestra en la Tabla 3.

Tabla 3. Ejemplo direcciones colombianas.

1	0	99	100	53A-119A	S	Kr 53A	
From left	From right	To left	To right	Street	Pedir	Street name	Sufdir

En donde:

- From left: indica el número de la primera casa que se encuentra a la izquierda de la calle.
- From right: indica el número de la primera casa que se encuentra a la derecha de la calle.
- To left: indica el número de la última casa que se encuentra a la izquierda de la calle.
- To right: indica el número de la última casa que se encuentra a la derecha de la calle.
- Street: indica la calle.
- Predir: indica si la primera componente de la dirección se encuentra al norte, sur, este u oeste.
- Street name: indica el nombre del a vía.
- Sufdir: indica si la segunda componente de la dirección se encuentra al norte, sur, este u oeste de la ciudad.

Como se tiene una dirección compuesta, en donde el número de la casa no puede estar relacionado solo con un número, la tabla de datos de la malla vial se tuvo que modificar para quedar como se muestra en la tabla 4.

Tabla 4. Tabla de la malla vial modificada.

119A-1	119A-0	119A-99	119A-100	53A-119A	S	Kr 53A	
From left	From right	To left	To right	Street	Pedir	Street name	Sufrir

De igual forma, las direcciones de los estudiantes se deben colocar en un formato similar para lograr su ubicación geográfica. Como se muestra en la tabla 5, se debe intercambiar en nombre de la vía con el número sobre la calle.

Tabla 5. Antes y después de las direcciones.

Antes	Después
CI 73 12-22	12-22 CI 73

Una vez los datos están en el formato adecuado, se genera un *Address Locator* en *ArcCatalog* que utilice como datos de referencia la malla vial acondicionada. El siguiente paso en la geocodificación es utilizar la herramienta *Geocoding* de *ArcMap* para obtener las coordenadas x e y de los estudiantes y generar el shape de puntos para su posterior análisis.

5.2.1.4. Módulo metaheurística. Para resolver los problemas VRP existen varias metaheurísticas por las que se puede atacar. La literatura recomienda, sobre todas las demás, las conocidas como Tabu Search y la Simulated Annealing.

La responsabilidad del presente trabajo es la implementación de la metaheurística Tabu Search (TS de ahora en adelante).

Como se menciona anteriormente, la metaheurística TS es sencilla en su implementación y converge rápidamente. La rapidez de la convergencia depende sobre todo en la técnica de diversificación utilizada, así como el método utilizado para obtener los vecinos.

Al proceso de diversificación inherente a TS como lo es la lista Tabú, se pueden agregar otras técnicas, cuyo uso depende en gran medida del problema específico a solucionar.

La técnica de ubicación de vecinos que se utiliza fue el intercambio entre par de nodos. Si se tiene una ruta específica, por ejemplo: {1,2,3,4} se intercambian los diferentes pares de nodos adyacentes, el 1 con: 2, 3, 4; el 2 con 3, 4; y el 3 con 4. Obteniendo así un conjunto de vecinos a la solución actual. Luego se le asigna una ponderación para cada uno de estos vecinos, ordenándolas de menor a mayor (ya que se trata de un problema de minimización).

Para apoyar la técnica de diversificación de Tabú, se idea una nueva técnica que no se encuentra en la literatura actual. La técnica consiste en durante un número n de iteraciones se trata de buscar una ruta que mejore la solución actual, de encontrarla, se continúa normalmente con el algoritmo TS. Pero, de no encontrarse una mejor solución se eliminan del tope de la lista, ordenada de menor mayor con los costos de cada ruta vecina, los primeros m elementos, y se trata la lista resultante (que se comporta casi igual a una cola) como si fuera la obtenida por el método de obtención de vecinos.

Esta metaheurística se repite para cada uno de los fragmentos arrojados por el módulo de fragmentación.

5.2.1.5. Módulo de post-optimación o implosión. El módulo de post-optimización tiene como función la mejora de la ruta arrojada por la metaheurística implementada.

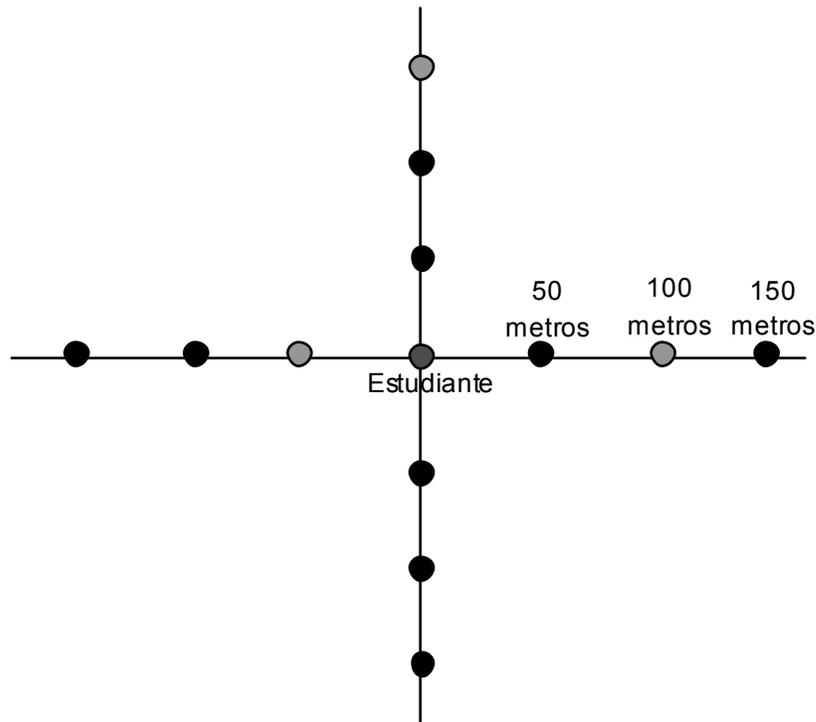
Para esto, se vale del hecho que los estudiantes se pueden desplazar un número de metros dados alrededor de su casa, acercándose al sitio por donde transita el bus. Existen otros casos en que los estudiantes cuentan con el servicio puerta a puerta, lo que significa que el bus debe recoger y dejar obligatoriamente al estudiante en su casa, esto se presenta para los estudiantes de más baja edad.

Para la determinación de la nueva ruta teniendo en cuenta esta nueva características se realiza el siguiente proceso.

- Se calculan n puntos alrededor de los estudiantes que se puedan mover de su casa (no cuentan con el servicio puerta a puerta), dichos puntos no deben significar un desplazamiento mayor a una cantidad m de metros para el estudiante. Estos puntos se obtienen de la división en $n-avos$ la circunferencia, en el caso de ser 4, se ubicarían los puntos en los grados 0° , 90° , 180° y 270° , a parte de de esto, se ubican puntos, siempre en estos grados, cada p metros. Por ejemplo si la distancia máxima a desplazarse por cada estudiante es de 200 metros se calculan puntos, según la distancia euclidiana, cada 50 metros hasta un número menor a 200, o sea, 150, y esto se repita para cada uno de los grados

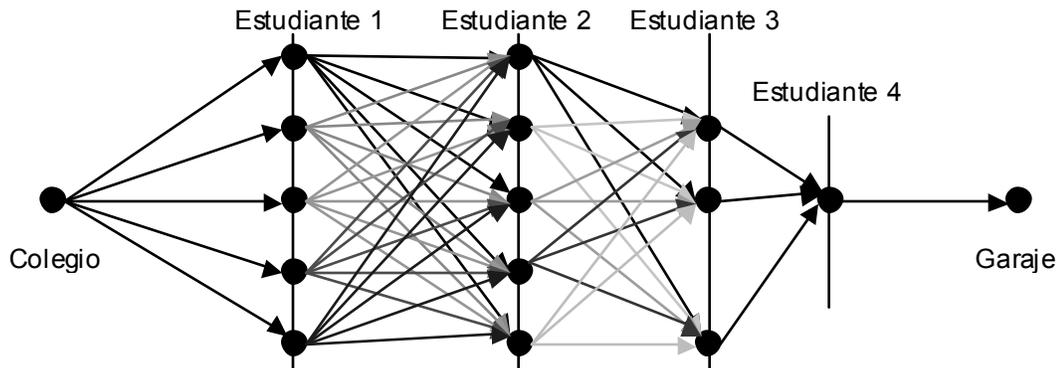
obtenidos de dividir la circunferencia en n grados. Con la herramienta ArcGIS se calculan los puntos alrededor del estudiante que más cerca estén de significar para el estudiante un desplazamiento de m metros (Figura 7), esta es la razón por la cual no se establece el punto a los 200 metros, ya que siempre la distancia real es un poco mayor que la distancia euclidiana, por lo que nunca la distancia del estudiante a un punto a 200 metros euclidianos, significa 200 metros reales de desplazamiento. Como se ve en la figura, puede que no exista un punto valido para uno de los grados, en este caso para 270° . A este conjunto de puntos obtenidos se le suma el estudiante, repitiendo este proceso para cada uno de los estudiantes que se puedan desplazar.

Figura 7. Puntos alrededor de un estudiante, 4 segmentos, 200 metros.



- El módulo anterior arroja una ruta que se obtiene de aplicar la metaheurística. Basados en este orden, en la ruta recogida de los estudiantes, se calculan las distancias de todos los puntos de dos estudiantes adyacentes. (Figura 8).

Figura 8. Obtención puntos alrededor de un estudiante.



- Una vez se tiene las matrices origen y destino de cada uno de estos segmentos (estudiante1 con estudiante2, estudiante 2 con estudiante 3, etc.) se procede a calcular con el algoritmo de Dijkstra el camino que minimice el recorrido total.

Al terminar el algoritmo, arroja una ruta en el mismo orden de la cual arrojo la metaheurística, pero con los caminos acortados por los desplazamientos de los estudiantes.

5.2.1.6. Módulo desplegador. Este módulo es el encargado de mostrar sobre la malla vial las rutas que deben tomar los diferentes buses, tanto en el trayecto recogida como el distributivo. Aparecen en diferentes colores para su mayor comprensión.

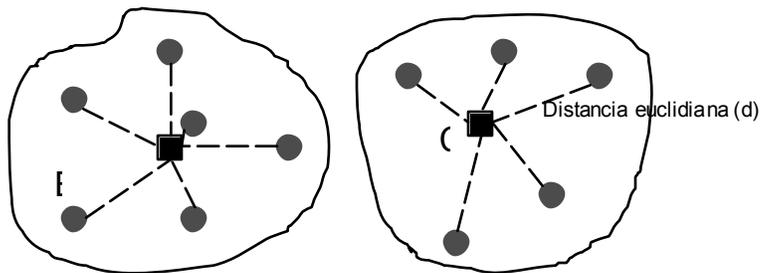
6. RESULTADOS

6.1. RESULTADOS SEGMENTACIÓN

Este módulo busca encontrar segmentos en los que cada uno mantenga la capacidad del vehículo que se asigna para servir sus estudiantes. Además se agrega a cada segmento la ubicación del colegio y del garaje para que en su generación se tenga en cuenta esos dos puntos comunes para todos los grupos.

La función objetivo es minimizar la sumatoria total de las distancias euclidianas (d) de cada uno de los estudiantes (E) al centroide (C) de cada uno de los segmentos, sin exceder la capacidad del segmento (S). En la Figura 9 se muestra un esquema del problema seguido del planteamiento matemático del modelo.

Figura 9. Modelo segmentación.



$$\min \sum_{i \in C} \sum_{j \in E} d(i, j) X_{ij}$$

sa

$$\sum_{j \in E} X_{ij} \leq S_i, i \in C$$

$$\sum_{i \in C} X_{ij} = 1, j \in E$$

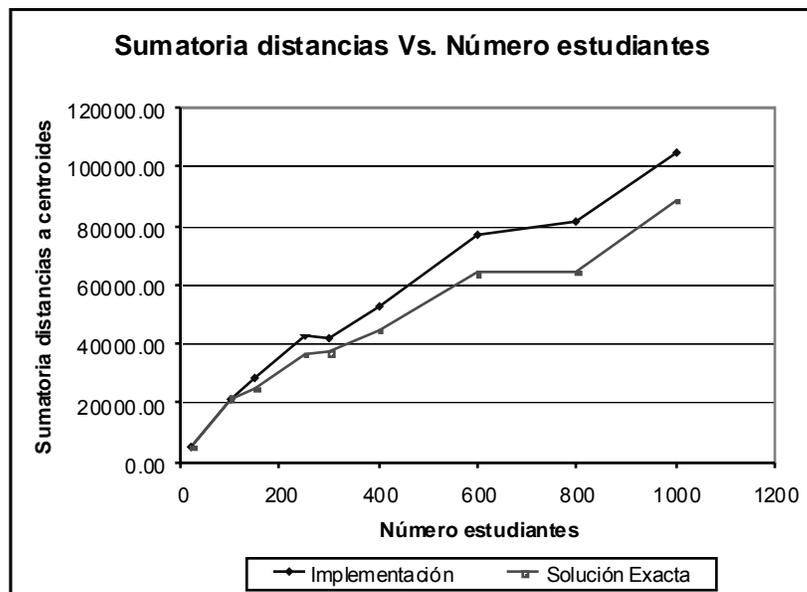
$$X_{ij} \in \{0,1\}$$

Para validar la solución propuesta se implementa en Xpress-IVE un algoritmo para conocer la solución exacta al problema. Esta herramienta se utiliza para resolver problemas de optimización en donde se tiene una función objetivo y un número de restricciones que aplican sobre el modelo. A continuación se muestran los resultados, se puede apreciar que a medida que el tamaño crece la calidad de la respuesta va disminuyendo, pero es preciso anotar que si el algoritmo implementado para la investigación se alarga, los resultados mejoran.

Tabla 6. Resultados validación segmentación

Num Estudiantes	Sumatoria distancia centroides		% del óptimo
	Implementación	Exacta	
20	5309.65	5244.3	1.25%
100	21741.43	21403.8	1.58%
150	28658.37	25241.8	13.54%
250	42361.09	36558.5	15.87%
300	42032.29	36981.7	13.66%
400	52960.81	44891.1	17.98%
600	77148.49	63793.7	20.93%
800	81144.12	64228.4	26.34%
1000	104624.89	88806.6	17.81%

Figura 10. Comportamiento del modelo.



6.2. COMPARACIÓN TABU SEARCH – SIMULATED ANNEALING

Uno de los objetivos de la presente investigación es comparar los resultados obtenidos entre la implementación de la metaheurística simulated annealing y la de tabu search.

Para que la comparación se ambos algoritmos se realice en condiciones iguales, se define el presente ambiente de pruebas:

- Datos de prueba: se utiliza la instancia ftv64.atsp.
- Características del computador utilizado: Intel Pentium 4 1.70GHz

Se ejecutaron varias veces las dos metaheurísticas para obtener datos comparables, se toma el tiempo y el costo de la solución final, en la tabla 7 y tabla 8, se muestran los resultados obtenidos y en la figura 11 se puede ver el comportamiento de las implementaciones en el tiempo.

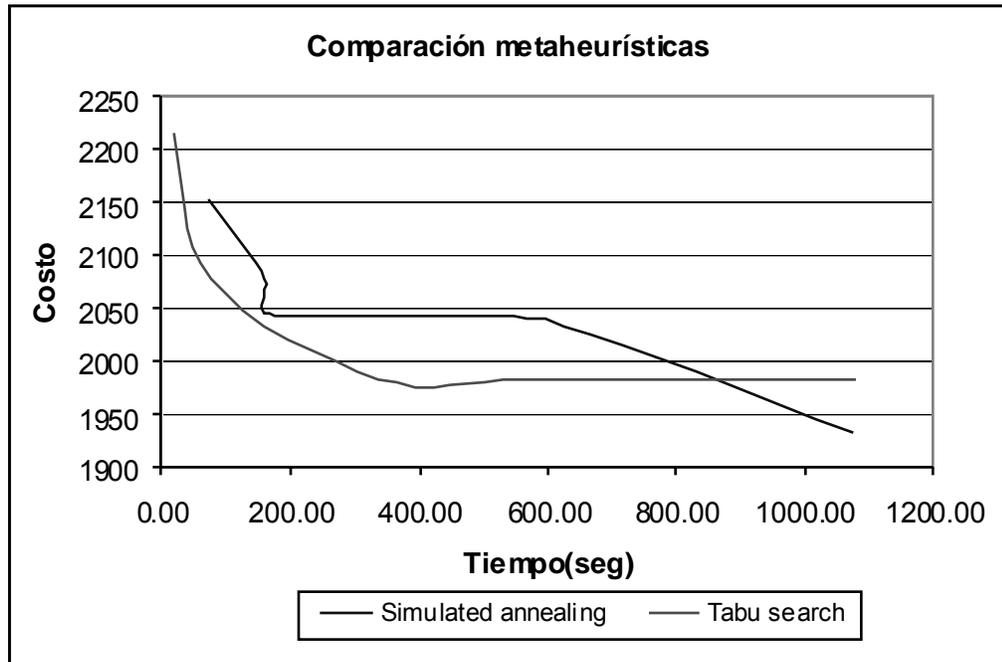
Tabla 7. Resultados para Simulated Annealing.

Tiempo(seg)	Costo	% del óptimo
73.95	2154.2	17.46%
160.68	2077.2	13.26%
167.81	2043.8	11.44%
446.93	2042.2	11.35%
624.31	2033.8	10.89%
1073.20	1932	5.34%

Tabla 8. Resultados para Tabu Search.

Tiempo(seg)	Costo	% del óptimo
17.28	2215	0.207742639
78.91	2076	0.131952017
336.38	1982	0.080697928
556.67	1982	0.080697928
753.28	1982	0.080697928
1078.67	1982	0.080697928

Figura 11. Resultados en costo y tiempo de las dos metaheurísticas.



De lo anterior se puede concluir que el algoritmo que mejor desempeño tiene es el de la búsqueda tabú, a pesar de que Simulated Annealing da buenos resultados solo logra mejorar la solución de Tabu Search luego de varios minutos.

REFERENCIAS

- [1] Ministerio de Transporte <URL:www.mintransporte.gov.co> Documento CONPES, No. 3348. 18 abril de 2005.
- [2] Seguí Pons, J. M., Ruiz Pérez, M., Guaita Mas, F., Escalas, F., Bauxà, A. (2003): “La planificación de rutas de transporte escolar a través de un SIG: El proyecto SIGTEBAL”, GeoFocus (Artículos), nº 3, p. 58-76. ISSN: 1578-5157.
- [3] Secretaría de educación Distrital. Beneficios para el sector educativo derivados del impuesto al alumbrado público [En línea]. 2003. Ítem 2.2. Desfase entre la ubicación de la oferta y la demanda. Disponible en Internet
<<http://pagina.sedbogota.edu.co/ExposicionAlumbradoEDAb21.pdf>>
- [4] Seguí Pons, J. M., Ruiz Pérez, M., Guaita Mas, F., Escalas, F., Bauxà, A. La planificación de rutas de transporte escolar a través de un sig: el proyecto sigtebal [En línea]. 2003. Revista internacional de Ciencia y Tecnología de la información geográfica. GeoFocus Artículos), nº 3, p. 58-76. ISSN: 1578-5157. 1. Introducción Disponible en Internet.
- [5] Ibídem.
- [6] Sustainable Transport - School travel strategy. School travel strategy. [Online]. 16 may 2005. Available from Internet.
<<http://www.wandsworth.gov.uk/Home/EnvironmentandTransport/LocalAgenda21/SustainableTransport/schooltravelstrategy.htm>>
- [7] Safe Routes to Schools. Case studies. [Online]. 2004. Available from Internet
<<http://www.saferoutestoschools.org.uk/index.php?f=casestudies.htm>>
- [8] Proyecciones de DAPD; 2011-2013: Proyecciones SAS-SED, con base en la tasa de crecimiento geométrica anual del período 1997-2010.
- [9] Secretaría de educación Distrital. Beneficios para el sector educativo derivados del impuesto al alumbrado público [En línea]. 2003. Ítem 2.2. Desfase entre la ubicación de la oferta y la demanda. Disponible en Internet
<<http://pagina.sedbogota.edu.co/ExposicionAlumbradoEDAb21.pdf>>
- [10] Ibídem.
- [11] Secretaría de educación Distrital. Campaña Ruta Pila 2005. [En línea]. Disponible en Internet
<http://www.transitobogota.gov.co/admin/contenido/documentos/Seg-Escolar-2_10_57_10.pdf>

- [12] Problemas de etiquetado: Complejidad computacional. Pedro Reyes Columé. Sevilla, 2002 [En línea]. Disponible en Internet: <http://www.personal.us.es/preyes/Invest/Tesis_PRC.pdf>
- [13] Ibídem.
- [14] Curso C++. Capítulo 0.1.1 Computación [En línea]. Disponible en Internet <http://www.zator.com/Cpp/E0_1_1.htm>
- [15] Máquina de Turing. Wikipedia. [En línea]. Disponible en Internet. <http://es.wikipedia.org/wiki/M%C3%A1quina_de_Turing>
- [16] Espacio logarítmico. Wikipedia. [En línea]. Disponible en Internet. <http://es.wikipedia.org/wiki/Espacio_logar%C3%ADtmico>
- [17] NL (Clase de complejidad). Wikipedia. [En línea]. Disponible en Internet. <http://es.wikipedia.org/wiki/NL_%28clase_de_complejidad%29>
- [18] Tiempo polinómico. Wikipedia. [En línea]. Disponible en Internet. <http://es.wikipedia.org/wiki/Tiempo_polin%C3%B3mico>
- [19] NP. Wikipedia. [En línea]. Disponible en Internet. <<http://es.wikipedia.org/wiki/NP>>
- [20] NP-completo. Wikipedia. [En línea]. Disponible en Internet. <<http://es.wikipedia.org/wiki/NP-completo>>
- [21] NP-hard. Wikipedia. [En línea]. Disponible en Internet. <<http://es.wikipedia.org/wiki/NP-hard>>
- [22] Vehicle Routing Problem – The Evolutionary way. [En línea]. Available from Internet. <<http://eden.dei.uc.pt/~jast/vrp/>>
- [23] TSPLIB. [En línea]. Available from Internet. <<ftp://ftp.zib.de/pub/Packages/mp-testdata/tsp/tsplib/tsplib.html>>
- [24] Problema del agente viajero. [Fuera de línea]. Encontrada en Internet. <http://www.pue.udlap.mx/%7Etesis/lis/lezama_m_r/capitulo3.pdf>
- [25] TSPLIB. [En línea]. Available from Internet. <<ftp://ftp.zib.de/pub/Packages/mp-testdata/tsp/tsplib/tsplib.html>>
- [26] Problema del agente viajero. [Fuera de línea]. Encontrada en Internet. <http://www.pue.udlap.mx/%7Etesis/lis/lezama_m_r/capitulo3.pdf>
- [27] Ibídem.
- [28] Módulo TSP. [En línea]. Disponible en Internet. <<http://www.fing.edu.uy/inco/cursos/ioce/material/Invitados/ruteo-parte1.ppt>>
- [29] Técnicas heurísticas para la resolución de problemas: computación evolutiva y redes neuronales. [En línea]. Disponible en Internet. <<http://geneura.ugr.es/~jmerelo/tutoriales/heuristics101/>>
- [30] Ibídem.
- [31] Bin Parking Problema. [En línea]. Disponible en Internet. <<http://www.cs.gsu.edu/~cscskp/Algorithms/NP/node11.html>>

- [32] The VRP Web. [En línea]. Disponible en Internet.
<http://neo.lcc.uma.es/radi-aeb/WebVRP/index.html?/Problem_Descriptions/BinPacking.html>
- [33] The VRP Web. [En línea]. Disponible en Internet.
<<http://neo.lcc.uma.es/radi-aeb/WebVRP/>>
- [34] Handbook of Metaheuristics. Glover F.W., Kochenberger G.A. 2002.
- [35] Búsqueda Tabú. Fred Glover. Belén Melián. [En línea]. Disponible en Internet. <<http://tornado.dia.fi.upm.es/caepia/numeros/19/glover.pdf>>
- [36] Ibídem.

ANEXOS

Anexo A. Diagrama de clases de fragmentación.

