Un Modelo de Asignación de Recursos a Rutas en el Sistema de Transporte Masivo Transmilenio

A Model for Resource Assignment to Transit Routes in Bogota Transportation System Transmilenio

Sergio Duarte, Ing., David Becerra, Ing., Luis Fernando Niño, PhD.
Laboratorio de Investigación en Sistemas Inteligentes, ALGOS-UN
Universidad Nacional de Colombia - Sede Bogotá
(srduartet, dcbecerrar, lfninov) @unal.edu.co

Recibido para revisión 28 de Noviembre de 2007, aceptado 14 de Febrero de 2008, versión final 28 de Febrero de 2008

Resumen—En este trabajo se presenta un modelo basado en algoritmos genéticos, teoría de colas y teoría de grafos para la planeación de sistemas masivos de transporte. Entre las características principales del modelo se propone: *i)* El modelamiento real de la troncal de las Américas del sistema de transporte masivo Bogotano Transmilenio; *ii)* Un preprocesamiento de datos utilizando teoría de grafos para caracterizar las rutas más cortas entre todas las combinaciones posibles de estaciones origen y destino *iii)* Utilización de algoritmos genéticos para optimizar el tiempo que gasta un usuario en el sistema Transmilenio por medio de la asignación de buses y frecuencias de salida. iv) La simulación de eventos por medio de distribuciones de Poisson y Erlang, simulando el evento de llegada de un bus Transmilenio a una estación determinada y el tiempo de espera de arribo del próximo bus respectivamente.

Además se desarrolla una metodología experimental para validar el modelo propuesto y estudiar los resultados otorgados por el mismo.

Palabras Clave—Sistema de Transporte, Teoría de Colas, Distribución Erlang, Grafos, Dijkstra, Algoritmos Genéticos.

Abstract—In this work, a model based on genetic algorithms, queue theory and graph theory for route planning in a mass transportation system is presented. Most important features of the proposed approach are i) the modeling of the Americas line in the mass transportation system Transmilenio in Bogota; ii) Data preprocessing using graph theory to characterize the shortest routes between all the possible combinations of destination and source stations; iii) the optimization of travel time by route assignment using genetic algorithms iv) the simulation of events using the Poisson and Erlang distributions, corresponding to bus arrival at specific stations and to user´s waiting time.

Additionally, an experimental methodology was developed to validate the proposed approach.

Keywords—Transportation System, Graph Theory, Queue Theory, Genetic Algorithm, Dijkstra.

I. INTRODUCCIÓN

El desarrollo de estrategias de planeación de rutas en los sistemas de transporte masivo es fundamental para la optimización de costos del negocio y del cumplimiento de necesidades por parte del usuario de los sistemas de transporte. La implementación de dichas estrategias está enfocada a evitar costos operacionales elevados, disminuyendo la proporción de estos costos respecto a los logísticos [1].

La optimización de operaciones logísticas de transporte es una de las áreas de planeación que más auge ha tenido durante la última década debido al incremento de necesidades de crear soluciones y optimizaciones a los problemas de distribución, así mismo se ha visto soportada por la implementación de nuevas tecnologías que contribuyen a la realización de soluciones más integrales, confiables y dinámicas.

El desarrollo de nuevas tecnologías partiendo de sistemas inteligentes para la planeación de la operación de sistemas de transporte son herramientas útiles para la optimización de funciones y costos operativos de un sistema de transporte. [2]

Se espera que al aplicar las nuevas tecnologías de procesamiento de información, comunicaciones, control y electrónica, los sistemas masivos de transporte creen caminos, vehículos y usuarios "más inteligentes". Se espera también que la aplicación de estas tecnologías mejore la operación de los sistemas de transporte al proveer rutas más eficientes a los viajeros. [2]

En este trabajo se aplicó un algoritmo genético como metodología de optimización para la asignación de frecuencias y número de buses para las rutas determinadas en la operación del sistema de transporte masivo Transmilenio en la troncal de las Américas. Se simuló el comportamiento del sistema partiendo de la teoría de colas, herramientas estadísticas y probabilísticas. La simulación del comportamiento de usuario se trabajó a partir de teoría de grafos.

El documento está organizado como sigue: En la sección 2 algunos fundamentos de distribuciones y teoría de colas es brevemente explicado. En la sección 3 se analiza la lógica de la aproximación planteada. En la sección 4 se explica la implementación a partir de algoritmos genéticos. En las últimas secciones se desarrollar una metodología de experimentación para validar la significancia y precisión del modelo, seguida de los resultados y correspondientes conclusiones.

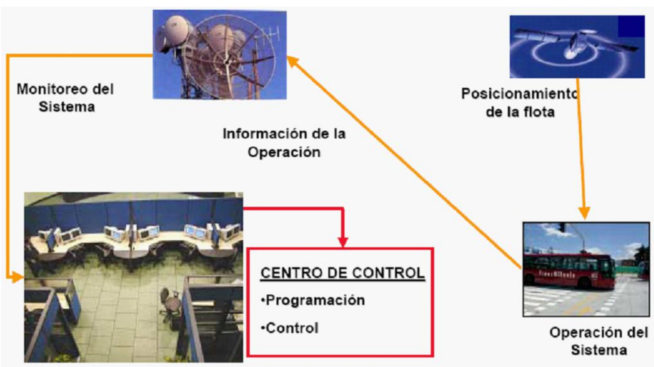

Figura 1. Sistema de programación y control de Transmilenio (tomado de [5]).

II. BACKGROUND

A Distribución de Poisson

La distribución de Poisson es una distribución de probabilidad discreta, que expresa la probabilidad de un número de eventos ocurriendo en un tiempo fijo si estos eventos ocurren con una tasa media conocida, y son independientes del tiempo desde el último evento.

La distribución de probabilidad está dada por:

$$f(k;\lambda) = \frac{e^{-\lambda}\lambda^{k}}{k!} \qquad \text{(Eq. 1)}$$

donde e es 2.71828, $k!$ es el factorial de k, λ es el número esperado de ocurrencias durante un intervalo dado.

A.1 Proceso de Poisson

Una variable es de Poisson cuando el número de eventos que ocurren en un intervalo temporal o espacial de tamaño dado cumple las siguientes condiciones [3]:

- El número de eventos que ocurren en el intervalo es independiente del número de los que ocurren fuera del mismo.
- Existe un intervalo lo suficientemente pequeño, de tamaño h, para el que la probabilidad de que en el mismo ocurra un sólo evento es proporcional al tamaño del intervalo, es decir es λh, siendo por tanto λ la probabilidad de que ocurra un evento en un intervalo de tamaño unidad.
- La probabilidad de que en cualquier intervalo de tamaño h suficientemente pequeño ocurran dos o más eventos, es prácticamente despreciable.

B. Distribución Erlang

La distribución Erlang es una distribución de probabilidad continua con dos parámetros k y θ cuya función de densidad para valores $x > 0$ es.

$$f(x) = \lambda e^{-\lambda x}\frac{(\lambda x)^{k-1}}{(k-1)!} \quad \text{(Eq. 2)}$$

La distribución Erlang es el equivalente a la distribución gamma con el parámetro k=1,2, ..etc y $\lambda = 1/\theta$. Para k= 1 eso es la distribución exponencial.

Se utiliza la distribución para describir el tiempo de espera hasta el suceso número k en un proceso de Poisson. Es decir que esta distribución mide el tiempo que transcurre entre un suceso y el m-ésimo siguiente. La distribución tiene dos parámetros *Erlang(m,b)* donde b es la media de una distribución exponencial y m es el número de sucesos que se cuentan[4].

C. Teoría de Colas

La teoría de las colas es el estudio matemático de las colas o líneas de espera. La formación de colas es, por supuesto, un fenómeno común que ocurre siempre que la demanda efectiva de un servicio excede a la oferta efectiva.

Dentro de las características claves de un sistema de colas podemos encontrar. [6]

- Una población de clientes, que es el conjunto de los clientes posibles.
- Un proceso de llegada, que es la forma en que llegan los clientes de esa población.
- Un proceso de colas, que está conformado por (a) la manera que los clientes esperan para ser atendidos y (b) la disciplina de colas, que es la forma en que son elegidos para proporcionarles el servicio.
- Un proceso de servicios, que es la forma y la rapidez con la que es atendido el cliente y por último un proceso de salida. Ver figura 2.

Figura 2. Sistema de Cola genérico (tomado de[7])

III. APROXIMACIÓN PROPUESTA

Se desarrollo un algoritmo genético que evoluciona poblaciones de configuraciones de frecuencia y número de buses para unas rutas preestablecidas en un sistema de transporte de pasajeros masivo. El objetivo de la implementación y modelo desarrollado es generar una solución óptima al problema de planear la asignación de los buses a las diferentes rutas y su frecuencia de salida.

A. Simulación de un Sistema Real

Se modelo el sistema de transporte de Transmilenio[5] en la troncal de las Américas, se utilizaron las 18 estaciones comprendidas entre el Portal de las Américas y la estación Jiménez. De igual forma se utilizó las especificaciones técnicas de carga de pasajeros de los buses Transmilenio (160 pasajeros por bus) y las rutas actuales que cubren la troncal F1, F14, F19, F23, F60, F61 y F91.

B. Objetivos de Optimización

En todo sistema de teoría de colas, es necesario maximizar la satisfacción del cliente minimizando los costos de operación del sistema, es decir que se debe establecer un balance equilibrado y óptimo entre las consideraciones cuantitativas de costes y las cualitativas de servicio. En una solución al problema de la asignación de rutas se espera no solo en minimizar el tiempo que los clientes pasan en el sistema, sino también en minimizar los costos totales de aquellos que solicitan el servicio y de quienes lo prestan.

El modelamiento del sistema es realizado por medio de la simulación de los tiempos de espera por servicio y los tiempos de servicio. Es decir que se intenta minimizar el tiempo total que el cliente gasta en el sistema, esto se logra minimizando el tiempo que el usuario gasta en su viaje y el tiempo que el usuario gasta esperando la llegada de un bus.

Se ha introducido una heurística de las rutas óptimas para llegar de un origen a un destino determinado, ésta heurística es el producto de un preprocesamiento y análisis de rutas en grafos que representan el sistema de transporte.

Lo anterior supone que el usuario es "inteligente" y tomará aquellas rutas que lo lleven a su destino en el menor tiempo posible, por tanto disminuiremos el tiempo que el usuario gasta en el sistema, maximizando satisfacción y minimizando costos del sistema.

Figura 3. Portal de las Américas, Transmilenio (tomado de [5])

C. Arribo de Usuarios

La distribución utilizada para asignar a los pasajeros, da mayor prioridad a estaciones de mayor afluencia (Portal Américas, Banderas, Ricaurte, Av Jimenez), dado que en situaciones reales existen estaciones con un mayor volumen de usuarios, caso de los portales e intersecciones en el caso del sistema de trasporte masivo modelado Transmilenio. En el modelo propuesto cada estación corriente tiene asociada una probabilidad de:

$$p = \frac{1}{e + (e_{portales} \times 2)} \quad \text{(Eq. 3)}$$

Y para estaciones portales o especiales la probabilidad asociada es de:

$$p = \frac{3}{e + (e_{portales} \times 2)} \quad \text{(Eq. 4)}$$

donde e es el número de estaciones y $e_{portales}$ es el número total de portales o estaciones especiales consideradas en el sistema.

D. Utilización del servicio

D.1 Modelamiento de eventos por procesos de Poisson

El arribo de buses Transmilenio a una estación determinada se modela como un proceso de Poisson, ya que el número de Transmilenios que paran en una estación dada dentro de la Troncal Américas en las horas de mayor tráfico es un claro ejemplo que posee las características de una distribución de probabilidad de Poisson.

El promedio (media) de los arribos de los buses Transmilenio por hora de gran tráfico puede estimarse a partir de las frecuencias de salida de los buses de Transmilenio de un portal dado, es decir, que la media se calcula como la distancia en tiempo de salida de cada Transmilenio a cumplir su ruta. Esta media está determinada por.

$$media = \frac{TR_i}{NB_i}, \text{ (Eq. 5)}$$

donde TR_i es el tiempo óptimo de recorrido de la ruta i, y NB_i es el número de buses asignado a la ruta i.

Es claro que la simulación posee las siguientes características:

- El número de buses Transmilenios que llegan en un determinado intervalo a un tiempo específico es independiente del momento en que el intervalo de tiempo ocurre durante la hora de gran tráfico.
- El número de llegadas de un bus Transmilenio en cualquier intervalo de tiempo no depende del número de arribos de cualquier otro intervalo de un segundo.

D.2 Modelamiento de tiempos de espera Erlang distribuidos

Los eventos que ocurren independientemente con alguna rata promedio de ocurrencia pueden ser modelados con procesos de Poisson. El tiempo de espera entre k ocurrencias de los eventos son distribuidos Erlang. Es decir que el evento independiente de la llegada de un bus a una estación dada dentro de la troncal de las Américas son eventos modelados con procesos de Poisson, y el tiempo de espera que debe aguardar un pasajero del sistema para abordar el siguiente bus que le sirve dentro de la ruta óptima escogida es un evento Erlang distribuido.

El modelamiento es guiado respecto a el evento esperado y la media de espera, donde la media de espera está definida por la frecuencia de salida de los buses de la ruta que cada usuario del sistema espera, y el evento k se define como el número de buses que arriban (evento) antes que el usuario pueda tomar su servicio.

D.3 Modelamiento de exceso de cupo de servicio.

La utilización del servicio (tomar el bus Transmilenio) depende de la capacidad de cupos libres que tenga la ruta asignada en el momento de presentarse el evento de arribo a una estación.

Este modelamiento es realizado por la probabilidad:

$$p = \begin{cases} = \dfrac{NB_i * 160}{NP_i} & if(NP_i > NB_i * 160) \\ = 1 & if(NP_i <= NB_i * 160) \end{cases} \text{ (Eq. 6)}$$

donde NP_i corresponde al número de usuarios que esperan una ruta determinada y NB_i corresponde al número de buses asignados a esa ruta.

Es importante anotar que esta probabilidad determina la variable k en la distribución Erlang, donde k se define como el número de sucesos necesarios o arribos de buses Transmilenios necesarios para la utilización del servicio.

IV. ALGORITMO GENÉTICO PROPUESTO

Los algoritmos genéticos han emergido como una muy buena aproximación para la solución de problemas industriales, de servicios u otros sistemas complejos que son modelados por simulaciones en el computador y que necesitan ser optimizados en términos de su diseño estructural y políticas operacionales [9, 10].

Los algoritmos genéticos son métodos sistemáticos para la resolución de problemas de búsqueda y optimización que aplican a estos los mismos métodos de la evolución biológica: selección basada en la población, reproducción sexual y mutación.

De forma explícita, el proceso de evolución es modelado como una sucesión de frecuentes cambios en los *genes*, que constituyen las soluciones análogas a *cromosomas*. El espacio de soluciones posibles es explorado aplicando transformaciones a éstas soluciones candidatas tal y como se observa en los organismos vivientes: *cruce*, *inversión*, *mutación*.

Las soluciones codificadas en un cromosoma - el genotipo - *compiten* para ver cuál constituye la mejor solución (aunque no necesariamente la mejor de todas las soluciones posibles). El ambiente, constituido por las otras posibles soluciones, ejercerá una presión selectiva sobre la población, de forma que sólo los mejor adaptados (aquellos que resuelvan mejor el problema) sobrevivan o donen su material genético a las siguientes generaciones, igual que en la evolución de las especies.

La utilización de estos algoritmos nos ofrece varias ventajas para el problema de asignación de buses y frecuencias en un sistema de trasporte dada la idoneidad de los algoritmos genéticos para encontrar soluciones en espacios de búsquela grandes y complejos, que suelen ser característicos en el problema tratado dado el alto grado de libertad que conlleva la asignación de frecuencias a las distintas rutas junto al alto número de posibilidades en la asignación de buses en las mismas.

Adicionalmente, los algoritmos genéticos representan un método de búsqueda idóneo dada su naturaleza paralela que permite explorar el espacio de búsqueda en varias direcciones, esta característica evita estancamientos al encontrar soluciones sub-óptimas como en otros métodos en los que es necesario comenzar de nuevo todo el proceso al estar en uno de estos estados. Para el problema de trasporte tratado, este comportamiento es de gran utilidad dado que nos permite estudiar soluciones alternas que pueden llegar a ser de gran utilidad en la toma de decisiones cuando surgen condiciones externas al problema.

El algoritmo genético propuesto para el problema de asignación de buses y frecuencias al sistema de trasporte tratado se bosqueja en la figura 1. El proceso comienza con la asignación de parámetros del algoritmo genético, entre los parámetros se encuentra el grado de la mutación, el número total de usuarios en el sistema, número de buses, capacidad de los buses, número de rutas, tiempo total de viaje de cada ruta - teórica- , tamaño de la población, número de generaciones, entre otros. Posteriormente se inicializa la primera población de algoritmo genético y luego de forma cíclica se evoluciona las generaciones por medio de la aplicación de mutaciones a los individuos, la selección de los individuos que conforman cada nueva generación se implementa por el método de torneo.

En las siguientes secciones se explica en detalle los componentes del algoritmo genético para el problema de asignación e buses y frecuencias por medio de la especificación del individuo, la función de *fitness*, los operadores genéticos y el procedimiento de selección.

A. Definición del Individuo

El primer paso a la hora de implementar un algoritmo genético consiste en decidir la estructura que tendrá el cromosoma de cada individuo. De la definición correcta de esta estructura dependerá en gran medida el comportamiento del algoritmo.

Para el presente enfoque se codificó el cromosoma como un arreglo de igual longitud al número de rutas, de esta forma cada posición del arreglo contiene la programación correspondiente a una ruta. Dicha programación consiste en el número de buses el cual se codifica por medio de un número entero, y a la frecuencia de salida de dichos buses o la distancia en tiempo entre los buses de dicha ruta, el cual se representa por medio de un número de doble precisión. En la Figura 2 se muestra un esquema del genotipo del individuo:

El rango posible para el parámetro buses de cada ruta es *[1, n-r]*, donde n es el número total de buses establecido para el sistema y r representa el número de rutas, este intervalo simplemente nos dice que todas las rutas deben tener asignado por lo menos un bus. Para el parámetro frecuencia de cada ruta, el rango de valores posibles es: $[t_i/b_i, \ ]$, donde t_i es el tiempo máximo de viaje de la ruta i y b_i es el número de buses actualmente asignado a la ruta i. Con este intervalo garantizamos que al menos exista un bus para una ruta dada en el sistema.

Es importante recalcar, que el cromosoma representa en última instancia una potencial solución de planeación para el sistema de trasporte definido.

B. Inicialización de la Primera Población

La inicialización de la primera población del algoritmo genético se lleva a cabo por medio de un proceso completamente aleatoria. En el proceso, el cual es aplicado a cada individuo, se asigna cada uno de los buses a una ruta que es elegida siguiendo una distribución uniforme, no obstante se garantiza que ninguna ruta posea al menos un bus asignado.

Es importante mencionar que junto a la inicialización de la primera población, se lleva a cabo un proceso de inicialización de los orígenes y destinos de los pasajeros en el sistema por cada generación de individuos del algoritmo genético. Para este fin se construye una matriz de tamaño $e \times e$ - donde e representa el número de estaciones en el sistema - y se genera de forma aleatoria el origen y destino para cada uno de los pasajeros en el sistema - el número de pasajeros es un parámetro de entrada del modelo-, de esta forma en la posición $M_{i,j}$ de la matriz, contamos el número de pasajeros que parten de la estación e_i y se dirigen a la estación e_j. Naturalmente no se admite que el origen sea igual al destino, lo cual produce que la diagonal de esta matriz sea siempre cero.

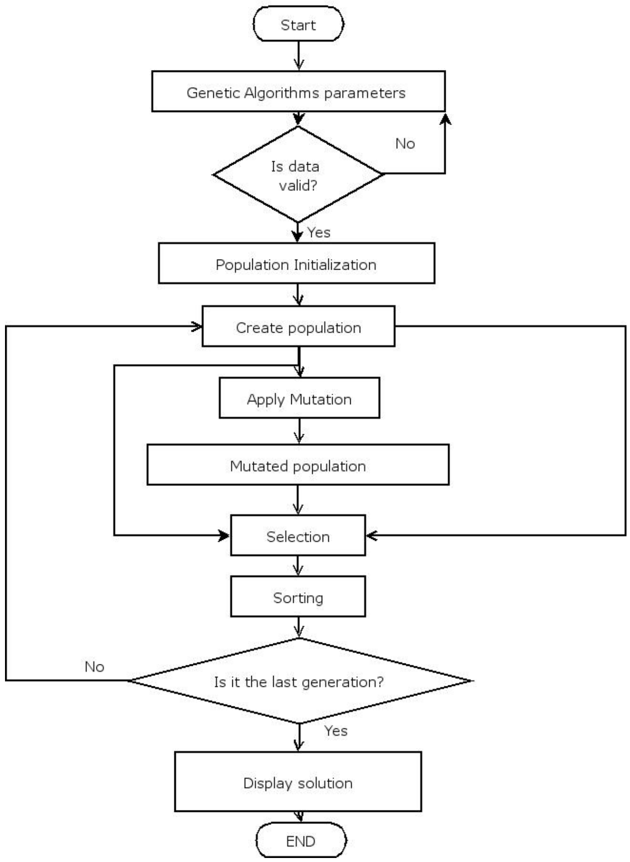

Figura 4: Diagrama de flujo del modelo

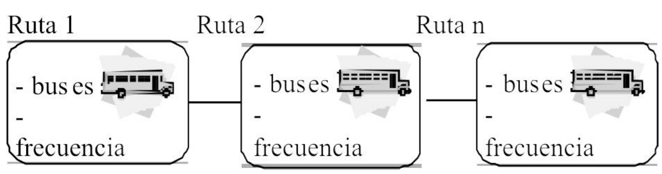

Figura 5: Representación del cromosoma

C. Definición de los Operadores Genéticos

Uno de los puntos clave del éxito de los métodos que utilizan algoritmos genéticos se encuentra en la habilidad de usar información encontrada en los genes de una población para utilizarla con la idea de crear unos nuevos candidatos más aptos. En el presente enfoque, este objetivo es alcanzado a través de un proceso de mutación, en el cual se llevan a cabo variaciones en individuos de la población original con el fin de explorar el espacio de búsqueda y encontrar mejores soluciones por medio de un criterio de selección.

La mutación que se propone consiste en elegir de forma aleatoria un gen en el individuo - de forma uniformemente distribuida- y añadir al gen (ruta) seleccionar un número de buses aleatorio entre el rango [1, *max*], donde *max* es un parámetro del algoritmo genético que nos permite controlar el grado de variación en los individuos mutados. Dado que existe un número establecido de buses en el sistema, se hace necesario equilibrar el número total de buses, por lo que se elige de forma aleatoria otro gen y se resta al número de buses contenido en éste el número de buses sumado al primer gen elegido. En el caso de que la diferencia realizada resulte en un número negativo o en cero, se intenta computar la resta con un gen diferente, si luego de intentar con todos los genes restantes no es posible terminar la operación, el individuo mutado se toma como no valido y se reinicia completamente el proceso de mutación seleccionando otro individuo de forma aleatoria.

En el caso en que el proceso haya sido exitoso - esto es equilibrar el número de buses en el sistema -, se procede a actualizar la frecuencia de salida en las dos rutas afectadas. Dicha actualización se efectúa por medio de la siguiente fórmula:

$$f = \frac{T_i}{b_i} \qquad \text{(Eq. 7)}$$

donde T_i es el tiempo total de recorrido de la ruta i y b_i es el nuevo número de buses asignado a dicha ruta. Con esta actualización garantizamos que al menos halla un bus en el sistema para la ruta especificada.

Es importante mencionar que para valores pequeños del parámetro *max* se lleva a cabo una exploración gradual del espacio de búsqueda mientras que valores grandes conllevan a saltos más grandes en la búsqueda.

D. Función Fitness

En el algoritmo genético buscamos la programación de buses y frecuencias que minimicen el tiempo que gastan los usuarios en el sistema, este tiempo puede ser cuantificado por medio de la siguiente ecuación:

$$t_{total} = \sum_{i=0}^{p} (t_v + t_e) \qquad \text{(Eq. 8)}$$

donde p es el número total de pasajeros del sistema en un periodo de tiempo discreto, t_v es el tiempo de viaje de cada uno de los pasajeros y t_e es el tiempo de espera del usuario para recibir el servicio. A continuación se describe como se cálculo cada uno estos términos.

D.1 Calculo del tiempo de viaje por medio de técnicas de grafos

Para calcular el tiempo de viaje se asumió que cada uno de los pasajeros selecciona el conjunto de rutas que representa el camino más corto para llegar a su destino. Esta hipótesis es congruente con la intención del usuario por minimizar su tiempo de viaje y con el razonamiento que lleva a cabo para satisfacer dicho objetivo.

Para obtener el tiempo de viaje se diseñó un grafo con las rutas seleccionadas para el problema, donde los nodos del grafo representan las estaciones y los arcos las conexiones existentes entre las estaciones, los cuales definimos como *caminos*. Los costos de los arcos se definen como el tiempo que tardamos de ir entre las estaciones que el arco conecta, de forma adicional a este tiempo, se agrega el conjunto de posibles rutas que cubren el recorrido entre el par de nodos - estaciones- analizadas.

La utilización de una estructura de grafo es idónea para representar la información del problema de una forma concisa y eficiente, ya que tenemos acceso a los tiempos de viaje entre cualquier origen y destino y aún más importante, podemos conocer las rutas que el usuario debe tomar para llegar a su destino con el tiempo de viaje más corto, información que es también necesaria para calcular el tiempo de espera promedio de los usuarios en el sistema.

Los datos que se emplearon en la construcción del grafo se resumen en la Tabla 1.

TABLE I
DESCRIPCIÓN DE LAS RUTAS TOMADAS PARA EL ESTUDIO

Ruta	Orden de las estaciones que visita
F1	0,1,2,3,4,5,6….14,15,16,17
F14	0,2,4,8,13
F19	4,7,9,14
F23	0,1,4,12,14,15,17
F60	0,2,4,8,10,11,14,16,17
F61	0,2,4,7
F91	0,2,4,8,12,14,15,17

Las siete rutas que se incluyeron en el modelo se tomaron de las rutas existentes en el sistema de transporte masivo Transmilenio para la Troncal Américas, así como las estaciones existentes en el mismo sistema - específicamente de Portal Américas hasta Av. Jiménez-. De esta forma nuestro modelo tiene una gran relevancia práctica ya que es posible comparar los resultados de nuestro sistema con la planeación actual de Transmilenio.

La obtención de las rutas más cortas entre todas las combinaciones de nodos se obtuvo por medio del algoritmo de Dijsktra [8], el cual nos permite obtener de una forma eficiente los caminos más cortos de un nodo origen hacia el resto de nodos del grafo. Dicho algoritmo se aplicó tomando como nodos origen cada uno de las *e/2* - donde *e* es el número de estaciones- primeras estaciones, nótese que no es necesario aplicar el algoritmo de Dijsktra *e* veces ya que el costo de ir a la estación e_i a e_j es el mismo de e_j a e_i, así como las rutas a tomar.

Es importante aclarar que la aplicación del algoritmo de Dijsktra [8] hace parte de un pre-procesamiento de datos en el cual se busca construir la matriz que representa al grafo con las estaciones conectadas por medio de las rutas impuestas y los tiempos de los caminos más cortos así como las rutas que conllevan a dichos tiempos, por lo que en la aplicación del algoritmo genético no se utiliza Dijsktra, simplemente el grafo construido de forma previa. De esta forma el cálculo de la función de *fitness* es optimizada desde el punto de vista de tiempo de cómputo, lo cual contribuye sustancialmente a reducir la complejidad computacional del algoritmo genético. En la Figura 3 se muestra el grafo diseñado para las rutas seleccionadas, en la figura también se puede observar la aplicación del algoritmo de Dijsktra tomando como nodo origen el primer nodo.

Desde el punto de vista computacional, el grafo se puede representar por medio de una matriz en la que el elemento $m_{i,j}$ de la matriz contiene el tiempo de viaje de ir de la estacón i a la estación j y las rutas asociadas a dicho tiempo.

Finalmente el cálculo del tiempo de viaje de los usuarios del sistema de trasporte, se lleva a cabo por medio de la siguiente fórmula:

$$t_{viaje} = \sum_{i=0}^{e/2} \sum_{j=i+1}^{e} (m_{i,j} \times p_{i,j}) \qquad \text{(Eq. 9)}$$

donde $p_{i,j}$ se refiere a los elementos de la matriz que posee el número de pasajeros que viajan de la estación *i* a la estación *j*.

D.1 Cómputo del tiempo de espera

El tiempo de espera de los usuarios del sistema se calcula por medio de la distribución de probabilidad Erlang, la cual es ideal para modelar los tiempos de espera en sistema de colas, el cual es análogo a la espera que los usuarios llevan en las estaciones para poder abordar un bus.

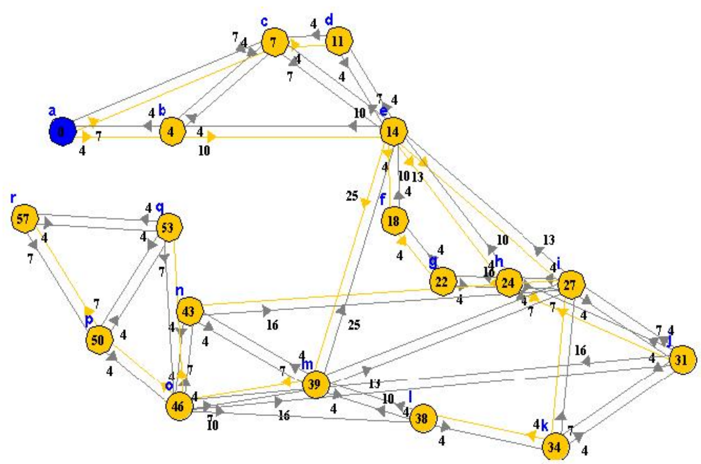

Figura 6: Grafo que representa las estaciones de la Troncal Américas y las rutas más cortas obtenidas a partir de 7 de las rutas del sistema Transmilenio.

Para calcular el tiempo de espera de los usuarios empleamos un vector de igual longitud al número de rutas, en el cual cada posición contiene la cantidad de usuarios que utiliza la ruta que representa dicha posición, esto es fácil de calcular a partir de la matriz que se inicializa en cada generación con el número de usuarios que viajan de la estación *i* a la estación *j*. Adicionalmente utilizamos un vector de la misma longitud al anterior, donde almacenamos la capacidad total de trasporte de cada una de las rutas, esta información se puede obtener a partir del número de buses que cada ruta tiene asignada por la capacidad de cada bus - valor que es un parámetro de entrada.

Con dicha información podemos emplear la distribución de Erlang. Para ello primero determinamos para cada usuario del sistema el número de intentos que debe realizar para poder tomar el servicio. Si la relación entre los cupos disponibles y la cantidad de personas que van a viajar en una ruta determinada es mayor o igual a 1, entonces podemos asumir que el usuario podrá utilizar el servicio en el primer intento. En caso contrario, el número de intentos está determinado por el número de eventos necesarios para obtener a partir de una distribución uniforme un valor aleatorio mayor a la razón mencionada.

Finalmente, con el anterior valor obtenido, nos es posible aplicar la siguiente fórmula para obtener el número total de espera:

$$t_{viaje} = \sum_{i=0}^{e/2} \sum_{j=i+1}^{e} erlang(k, k \times f(M_{i,j})) \quad \text{(Eq. 10)}$$

donde $f(M_{i,j})$ calcula la sumatoria de frecuencias de los caminos que se toman para ir de la estación *i* a la estación *j*.

V. MARCO EXPERIMENTAL

Se diseñó un marco experimental con el fin de analizar la consistencia y calidad de los resultados obtenidos por el modelo propuesto. Así mismo, se analizó el comportamiento de los algoritmos genéticos respecto a la evolución de las poblaciones con el fin de validar el modelo y la relevancia de sus resultados para el problema de programación de buses y frecuencias en un sistema de trasporte. Es también de gran importancia analizar la validez de las distribuciones de probabilidad elegidas para modelar el problema por medio del estudio de los resultados obtenidos en la simulación

Dentro de este marco experimental se desarrollo una aplicación empleando la tecnología JAVA, que implementa la simulación del algoritmo genético. Los experimentos se llevaron a cabo sobre un computador de escritorio con procesador Intel Core 2-Duo de 2.3Ghz y 1 Gb de memoria RAM.

Los datos de entrada del modelo son las rutas, las estaciones que recorre cada una de estas rutas, la cantidad de buses del sistema y la capacidad de cada bus. Los datos utilizados en nuestra experimentación fueron extraídos del sistema de trasporte Transmilenio, en particular se utilizaron las 18 estaciones de la Troncal Américas - entre Portal Américas y Av. Jiménez - y 7 de la rutas ofrecidas por Transmilenio en los horarios corrientes. La capacidad de los buses es de 160 pasajeros y se consideraron 400 buses en el sistema para una afluencia de usuarios de 100.000 personas.

Para el algoritmo genético se definió 50 como el tamaño de la población y se llevo a cabo la simulación para 20.000 generaciones. Los resultados de dicha simulación así como su análisis se presentan en la siguiente sección.

VI. RESULTADOS

En la figura 7 se observa el comportamiento del algoritmo genético a través de las generaciones, donde en la generación inicial el *fitness* del mejor individuo fue de 1432.32 minutos, mientras que el *fitness* del individuo de la última generación fue de 689.24 minutos. Los picos de la figura se explican por la utilización de distribuciones Erlang para el modelamiento de los tiempos de espera, ya que para cada generación correspondía una distribución diferente de los usuarios en las estaciones del sistema y las probabilidades de tiempo de espera variaban a través de las generaciones.

La convergencia del algoritmo genético nos muestra que nuestro sistema es capaz de modelar con significancia un sistema de transporte, ya que el algoritmo no se limitó a unas probabilidades dadas, sino que el sistema es dinámico a través del tiempo. En otras palabras, el modelamiento propuesto es capaz de simular un sistema de transporte en cualquier instante de tiempo, con modificaciones de variables que afectan el sistema.

En la figura 7 se muestra la mejor configuración obtenida para la asignación de buses y frecuencias para cada una de las rutas de la troncal de las Américas. Analizando éstos datos observamos que ésta información es congruente y directamente relacionada con la información obtenida por los grafos del sistema. Es decir que aquellas rutas que son más utilizadas por los usuarios y cuyo recorrido en tiempo es mayor, les fue asignado un mayor número de buses y su frecuencia se minimizó.

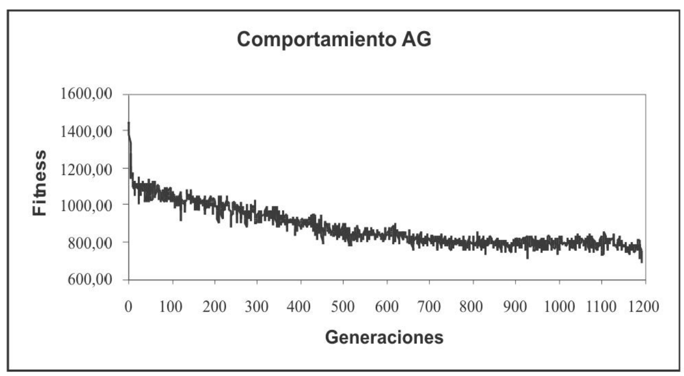

Figura 7: Comportamiento del algoritmo genético

TABLE II
ASIGNACIONES DEL MEJOR INDIVIDUO

RUTA	Num_buses	Frecuencia (min)
F1	97	1,4845
F14	54	1,5926
F19	39	1,6923
F23	94	1,234
F60	68	1,7353
F61	16	3
F91	32	3,625

VII. CONCLUSIONES

En éste trabajo se desarrollo una aproximación basada en algoritmos genéticos, teoría de colas y teoría de grafos para la optimización de la asignación de frecuencias y buses en rutas determinadas en un sistema de transporte masivo real. Aún cuando ciertas restricciones y discretizaciones del modelo real fueron tomadas, es claro que el modelo propuesto es un buen optimizador de asignación de buses a rutas dadas. Sin embargo es necesario trabajos más profundos que limiten las restricciones planteadas al sistema.

Las metodologías basadas en teoría de colas y grafos poseen significancia del mundo real y permiten el modelamiento efectivo de sistemas de transporte masivo de una forma fácil, lógica, estructurada y acorde a la realidad.

El uso de teoría de grafos para la determinación de la ruta óptima para cada una de las rutas establecidas en la troncal de las Américas es una heurística introducida como preprocesamiento de datos, que permitió el modelamiento del comportamiento del usuario de una forma eficaz, disminuyo la complejidad de la implementación y corrida del algoritmo propuesto y adicionalmente era congruente con los dos objetivos propuestos en la función de *fitness* de los individuos.

Los algoritmos genéticos para optimizar funciones extraidas del funcionamiento de sistemas de transporte masivo son herramientas fácilmente implementables y que logran modelar correctamente la función objetivo. Además, los algoritmos genéticos tienen la ventaja de producir un conjunto de familias de conformaciones de asignaciones de frecuencias muy cercanas al óptimo. Al obtener una familia de soluciones se tendrá mayor flexibilidad en el diseño e implementación real de las soluciones al sistema.

REFERENCIAS

[1] B. Blanchard, "Logistics Engineering And Management". Fifth Edition.Prentice Hall. 1998.
[2] J. Daza, J. Espinosa. "Hacia una Arquitectura Nacional para los sistemas Inteligentes de Transporte". Publicación Técnica No 251Sanfandila, Qro, 2004.
[3] V. Abraira, A. Pérez. "Métodos Multivariantes en Bioestadística". Ed. Centro de Estudios Ramón Areces. 1996.
[4] J. Serrano, "Computación Estadística, Generación de variables aleatorias", 2007.
[5] TransMilenio S.A, Transmilenio, http://www.transmilenio.gov.co.
[6] H. Moskowitz, G.P. Wright. "Investigación de Operaciones". Prentice_Hall Hispanoamericana S.A. 1991
[7] Urban operations Research, Logistical and Transportation Planning Methods, 1981.
[8] E.W. Dijkstra, "A Note on Two Problems in Connection with Graphs". Numerische Math. 1, 269-271, 1959.
[9] F. Azadivar, "Simulation Optimization Methodologies", Proceedings of the 1999 Winter Simulation Conference.
[10] S. Ólafsson, J. Kim, "Simulation Optimization", Proceedings of the 2002 Winter Simulation Conference.

Sergio R. Duarte. Obtuvo su título de Ingeniero de Sistemas en la Universidad Nacional de Colombia. Actualmente cursa su maestria en Linguisticas Computacionales en la Univerzita Karlova en Praga, República Checa. Sus áreas de trabajo e interés son la bioinformática, las linguisticas computacionales, el aprendizaje de máquina y la inteligencia artificial.

David C. Becerra. Obtuvo su título de Ingeniero de Sistemas en la Universidad Nacional de Colombia. Actualmente es estudiante de segundo semestre de la Maestría de Ingeniería de Sistemas y Computación en la misma Universidad. Su área de estudio es la bioinformática y el aprendizaje de máquina.

Luis F Niño. Es Ingeniero de sistemas de la Universidad Nacional de Colombia con Master en Ciencias Matemáticas de la Universidad Nacional de Colombia y Universidad de Memphis con énfasis en Ciencias de la Computación y Doctor de Filosofía en Ciencias Matemáticas de la Universidad de Memphis. Las áreas temáticas en las que trabaja son Sistemas Inteligentes, Aprendizaje de Máquina, Sistemas Inmunes Artificiales, Bioinformática, Aplicaciones Biomédicas de Sistemas Inteligentes. Actualmente se desempeña como profesor asociado en el Departamento de Ingeniería de Sistemas en la Universidad Nacional de Colombia.

Universidad Nacional de Colombia Sede Medellín
Facultad de Minas

Escuela de Ingeniería de Sistemas

Grupos de Investigación

Grupo de Investigación en Sistemas e Informática

Categoría A de Excelencia Colciencias
2004 - 2006 y 2000.

GIDIA: Grupo de Investigación y Desarrollo en Inteligencia Artificial

Categoría A de Excelencia Colciencias
2006 – 2009.

Grupo de Ingeniería de Software

Categoría C Colciencias 2006.

Grupo de Finanzas Computacionales

Categoría C Colciencias 2006.

Centro de Excelencia en Complejidad

Colciencias 2006

Escuela de Ingeniería de Sistemas
Dirección Postal:
Carrera 80 No. 65 - 223 Bloque M8A
Facultad de Minas. Medellín - Colombia
Tel: (574) 4255350 Fax: (574) 4255365
Email: esistema@unalmed.edu.co
http://pisis.unalmed.edu.co/