



Universidad Autónoma de Querétaro
Facultad de Ingeniería
Maestría en Matemáticas Aplicadas

**IMPLEMENTACIÓN DE UN ALGORITMO DE OPTIMIZACIÓN
BASADO EN UN MODELO MATEMÁTICO DEL
FLUJO DE TRÁNSITO VEHICULAR.
ESTUDIO DE CASO:
AVENIDA CONSTITUYENTES
EN LA CIUDAD QUERÉTARO, QRO**

TESIS

Que como parte de los requisitos para obtener el grado de
Maestro en Matemáticas Aplicadas

Presenta:

Lic. Gabriela Inzunza Gómez

Dirigido por:

Dr. Eric Moreno Quintero

SINODALES

Dr. Eric Moreno Quintero
Presidente

Dra. Rocio Peniche Vera
Secretario

Dr. Roberto de la Llata Gómez
Vocal

M.I. Oscar A. Rico Galeana
Suplente

M.C. Luz Angélica Gradilla Hernández
Suplente

Dr. Gilberto Herrera Ruíz
Director de la Facultad

[Firma]
Firma

[Firma]
Firma

[Firma]
Firma

[Firma]
Firma

[Firma]
Firma

[Firma]
Firma

Dr. Luis Gerardo Hernández Sandoval
Director de Investigación y
Posgrado

Centro Universitario
Querétaro, Qro.
Junio del 2008

México

RESUMEN

La programación de semáforos en las ciudades es una herramienta útil para mitigar los impactos en los tiempos de recorrido sobre las redes urbanas, en la seguridad vial y en el medio ambiente. En la actualidad la mayoría de los sistemas de semáforos se planean a partir de aforos de flujo fijos, sin considerar la respuesta que los usuarios tienen ante los cambios en la programación misma de estos semáforos, lo cual modifica los supuestos iniciales con los que se realiza la programación. En este trabajo se considera el problema de optimizar la programación, considerando como parámetros los tiempos de verde y la duración del ciclo de cada semáforo, sobre una red urbana, mostrando la forma en que la optimización puede anticipar el impacto en los patrones del tránsito vehicular debidos a los decisiones de ruta de cada uno de los usuarios de la red vial. Esto se logra incluyendo un modelo de equilibrio Wardropiano en la red, en el cual cada usuario busca su mejor ruta, como una de las restricciones del problema de optimización. El método heurístico denominado Algoritmos Genéticos es propuesto como método de solución a este problema, utilizando el tiempo total de viaje a través de la red como la función a optimizar, basado en los flujos de equilibrio que se generan para las diferentes programaciones de los semáforos. El procedimiento es aplicado en un caso de estudio en la ciudad de Querétaro. El comportamiento de los algoritmos es analizado con respecto a los parámetros del método de Algoritmos Genéticos.

Palabras clave: control semaforizado, equilibrio de Wardrop, optimización, algoritmo genético.

SUMMARY

Traffic signals programming is a well known and useful tool to control impacts on travel delay on urban networks, road safety and environment. Nowadays, traffic signals programs are usually designed on the basis of fixed traffic volume counts without considering users' response to changes in traffic signals, which implicitly alters the assumptions supporting the programming itself. This work addresses the problem of finding the optimal traffic signal program taking as parameters the green and cycle times in each traffic signal on an urban network, and showing how the optimization procedure anticipates the effect on the vehicle flows resulting from the routing decisions of the users traveling in the network. The task is done using a Wardropian equilibrium model, where each user looks for his (her) best route, as a constraint in the optimization problem. The heuristics named Genetic Algorithms is proposed as a solution method to the problem, using the users' total travel time on the road network as the performance function to optimize, and based on the equilibrium flows coming from the different traffic signals programs. The procedure is applied in a study case in the city of Querétaro. The algorithm behavior is analyzed with respect to the parameters of the Genetic Algorithms method.

Key words: traffic signal control, Wardrop equilibrium, optimization, genetic algorithm.

A mi hija, María Isabel.

AGRADECIMIENTOS

En la preparación de este trabajo se recogieron las opiniones desinteresadas del Dr. Eric Moreno Quintero y de la Dra. Rocío Peniche Vera de la Facultad de Ingeniería de la Universidad Autónoma de Querétaro, así como del Dr. Roberto de la Llata Gómez, del Centro Queretano de Recursos Naturales y del M.I. Oscar Rico Galeana y de la M. en C. Luz Angélica Gradilla Hernández del Instituto Mexicano del Transporte.

INDICE

	Página
Resumen	i
Summary	ii
Dedicatorias	iii
Agradecimientos	iv
Índice	v
Índice de cuadros	vi
Índice de figuras	vii
I. INTRODUCCIÓN	1
II. REVISIÓN DE LITERATURA	17
Teoría de asignación de tránsito y modelos de equilibrio	17
Problema de diseño de red. Modelo de optimización binivel	26
III. METODOLOGÍA	31
Formulación del problema	31
Proceso	40
Factibilidad de las restricciones	42
Metodología de solución	42
Diseño del cromosoma	44
Codificación de cromosomas y representación de restricciones	45
Inicialización	46
Evaluación del comportamiento	46
Selección	46
Entrecruzamiento	47
Mutación	47
Pseudocódigo del procedimiento	47
IV. RESULTADOS Y DISCUSIÓN	56
Características de la red vial	56
Resultados	60
Conclusiones	66
LITERATURA CITADA	67

ÍNDICE DE CUADROS

Cuadro		Página
1.1	Costo promedio de infraestructura carretera (miles de pesos del 2005 por km)	2
1.2	Escenario 1: Tiempo de recorrido conforme la distribución de los automóviles en la red cambia en las distintas rutas.	4
1.3	Tiempo de recorrido conforme la distribución de los automóviles en la red cambia en las rutas del escenario 2.	5

INDICE DE FIGURAS

Figura		Página
1.1	Ejemplo de la paradoja de Braess.	3
1.2	Escenario 2. Ejemplo de la paradoja de Braess	4
1.3	Ejemplo de equilibrio de Wardrop.	10
1.4	Velocidades de circulación en las rutas 1 y 2	11
1.5	Tiempo de recorrido en los caminos.	12
3.1	Diagrama de intersección a dos fases en intersecciones de cuatro ramas.	32
3.2	Diagrama de intersección a tres fases en intersecciones de cuatro ramas.	32
3.3	Diagrama de intersección a tres fases en intersecciones de cuatro ramas.	33
3.4	Relación entre capacidad y longitud de ciclo.	36
3.5	Diagrama de tiempos de ciclo para S fases.	37
3.6	Diagrama de flujo del procedimiento de solución.	41
4.1 a)	Mapa de la red vial de Querétaro	57
4.1 b)	Subred considerada	57
4.2	Mapa de Intersecciones Av. Constituyentes con calle de Los Plateros y con calle Ejército Republicano.	59
4.3	Mapa de Intersecciones Av. Constituyentes con Av. Cimatario y con calle 20 de Noviembre al norte y calle Nevado de Toluca al Sur.	60
4.4	Mapa de Intersecciones Av. Constituyentes con Av. Luis Pasteur.	61
4.5	Evolución de la Mejor Solución.	63
4.6	Desviación del Tiempo Total de Viaje	64
4.7	Evolución de la Mejor Solución con respecto a P_m	65
4.8	Desviación del Tiempo Total de Viaje con respecto a P_m	65

I. INTRODUCCIÓN

El crecimiento urbano y el acceso masivo a vehículos automotrices particulares en la mayoría de las ciudades medias y grandes del mundo, así como el ritmo de vida moderno que implica el traslado de personas a muchos y diversos lugares en un mismo día, han generado una tasa de crecimiento en los volúmenes de tránsito muy superior al índice de crecimiento de la vialidad. Esto ha agravado el problema del congestionamiento de tránsito vehicular en el ámbito mundial, cuyos impactos más conocidos son el deterioro de la calidad del aire, el aumento de accidentes de tránsito y un aumento exagerado del tiempo dedicado a trasladarse en las redes viales urbanas.

Por ejemplo, estimaciones para la capital norteamericana, Washington, D.C. indican que entre 1990 y 2020, el número de vehículos-milla aumentará alrededor del 75%, mientras que el incremento en kilómetros de red vial crecerá sólo un 22%, con lo que el número de horas-vehículo de retrasos en el flujo vehicular urbano aumentaría en cerca de 480%. Asimismo se ha estimado que el 85% de los viajes interregionales tendrán lugar en vías congestionadas (Ghassan, & Benekohal, 2000).

De acuerdo al Anuario Estadístico Económico del Municipio de Querétaro, publicado por la Secretaría de Economía en el año 2005, en este municipio, el padrón vehicular creció de 62,600 a 189,751 unidades entre 1995 y 2003, lo cual representa un crecimiento superior al 200% en menos de 10 años. En Querétaro hay cerca de 2,000 kilómetros de vías, clasificadas como primarias, secundarias y rurales con poco más de 140 kilómetros asfaltados. Algunas de estas vialidades municipales –como las del Centro Histórico– se establecieron desde la fundación de la ciudad en el año 1531, con un trazo para la circulación y flujos diferentes a los actuales, lo que en la mayoría de los casos imposibilita su ampliación.

La solución al problema de la congestión vehicular difícilmente puede basarse en la ampliación de capacidad de las redes viales, como se ha visto en la práctica, ya que, aparte de la dificultad de construir en el ambiente urbano, donde muchas veces ya no hay espacio disponible, estas construcciones implican costos de enormes magnitudes; de acuerdo con el Instituto Mexicano del Transporte (2006), los costos promedio de construcción carretera se

presentan en el Cuadro 1.1, en miles de pesos al 2005, agregando que en la construcción urbana hay costos adicionales por iluminación, drenaje y señalización.

Cuadro 1.1 Costo promedio de construcción carretera (miles de pesos del 2005 por km)

Tipo de Carretera		Tipo de Terreno		
		Plano	Lomerío	Montañoso
Cuatro Carriles	(A4) (22 m. de corona)	22,647	43,028	73,601
Dos Carriles	(A2) (11 m. de corona)	14,720	22,647	45,293
Dos Carriles	(B) (9 m. de corona)	5,706	7,633	10,523
Dos Carriles	(C) (7 m. de corona)	4,817	5,706	8,597
Un Carril	(E) (4 m. de corona)	482	963	1,408

Fuente: IMT. (2006). Manual Estadístico de Sector Transporte, 2006.

Instituto Mexicano del Transporte, Querétaro, México

Por otra parte, el hecho de que los automovilistas suelen tomar sus decisiones de ruta de modo individual, tratando de encontrar los mejores tiempos para llegar a sus destinos, puede entorpecer los esfuerzos de los planificadores urbanos que construyen y amplían la red vial. Un ejemplo de este resultado contradictorio lo ilustra la llamada Paradoja de Braess.

Esta paradoja, atribuida al matemático alemán Dietrich Braess, establece que la adición de capacidad extra a una red de caminos, cuando los usuarios deciden individual e independientemente el uso de sus rutas, puede, en algunos casos, disminuir el desempeño global de la red. La paradoja puede describirse como sigue: *“Sea una red vial donde en cada nodo hay un número dado de automóviles iniciando su viaje y con un destino determinado. Bajo estas condiciones se desea estimar la distribución de los flujos de tránsito en la red. El que una calle sea preferida a otra depende no sólo de la calidad del camino, sino también de la densidad del flujo vehicular que tiene. Si cada conductor elige la ruta más conveniente para él, los tiempos totales de traslado en la red no necesariamente son mínimos. Aún más, existen ejemplos en los que la construcción de un nuevo tramo en la red podría redistribuir el tránsito en un patrón que resulta en tiempos individuales de recorrido mayores”* (Wikipedia, 2007a).

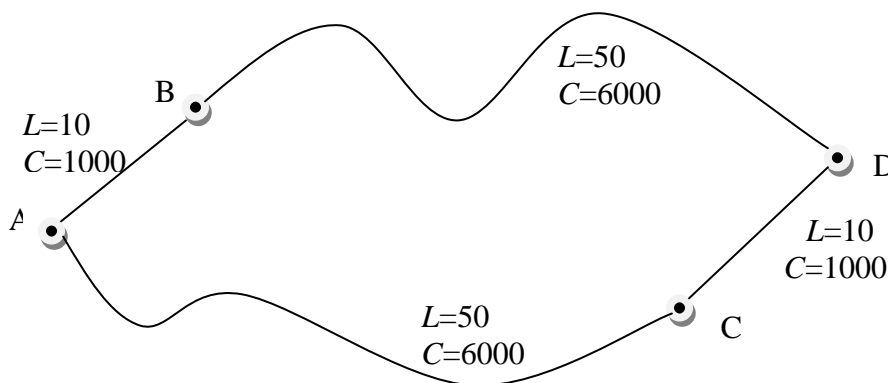
Feather (2005) presenta un análisis práctico de esta paradoja en un modelo simple de red vial, ilustrado en la Figura 1.1, en donde el origen se define en A y el destino en D. Bajo la hipótesis de que todos los caminos tienen un límite máximo de velocidad de 60mph (una milla por minuto), y en donde todos los usuarios manejarán al límite de velocidad, si se puede. Finalmente se supone que cada camino tiene capacidad de C de automóviles por hora y longitud L millas.

- Si el flujo F (en automóviles por hora) es menor que la capacidad C , fluye de manera libre y éste puede tomar L minutos en recorrer el camino. Conforme la ruta se satura habrá más congestión, lo que hace que el tiempo de recorrido sea mayor.
- Si el flujo F tiende a ser igual que la capacidad de la vía, empezará a haber interrupciones de flujo, siendo el tiempo de recorrido, proporcional tanto a la cantidad de tránsito como a la longitud del camino.

De esta manera se propone el siguiente algoritmo para la función de tiempo T de recorrido :

$$T = \begin{cases} L + F/1000 & \text{Si } F < C \\ LF/C & \text{Si } F \approx C \end{cases}$$

Figura 1.1. Ejemplo de la paradoja de Braess



Suponiendo un flujo de 6,000 automóviles por hora que desear viajar de A a D, existen dos rutas, una por B y otra por C. Cada ruta consiste en una sección de alta capacidad y otra de pequeña capacidad. La longitud en tiempo de la ruta para ir de A a D depende de la cantidad de tránsito en cada ruta y por lo tanto de la distribución de los autos en la ruta. En el

Cuadro 1.2 se muestran los cálculos resultantes de distribuir en distintas proporciones a los 6000 vehículos por las dos rutas posibles.

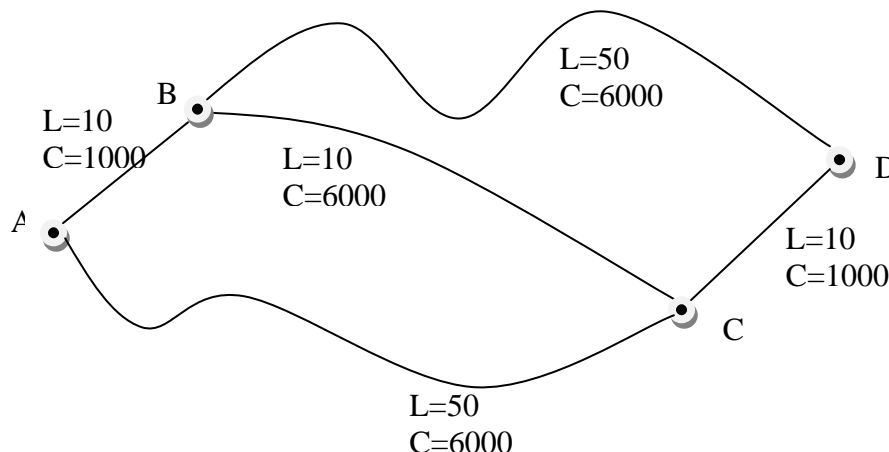
Cuadro 1.2. Escenario 1: Tiempo de recorrido conforme la distribución de los automóviles en la red cambia en las distintas rutas.

Autos por hora		Tiempo de recorrido de A a D	
vía B	vía C	vía B	vía C
0	6000	60	110
1000	5000	65	101
2000	4000	74	92
3000	3000	83	83
4000	2000	92	74
5000	1000	101	65
6000	0	110	60

Se observa que si el flujo es distribuido de manera asimétrica, una ruta será más rápida que la otra; los usuarios que van por la ruta más lenta cambiarán a la otra tratando de mejorar su tiempo de traslado. El sistema se estabiliza cuando la mitad del tránsito va en cada ruta con un tiempo total de 83 minutos. Esta estabilización corresponde al equilibrio de Wardrop.

El ejemplo continúa, suponiendo que los administradores de la red vial deciden construir un nuevo tramo de B a C con alta capacidad, como se ilustra en la Figura 1.2 Esta situación será el Escenario 2.

Figura 1.2. Escenario 2. Ejemplo de la paradoja de Braess



Bajo el escenario 2, existen tres rutas sensibles de A a D y el tiempo de traslado en cada una depende del tránsito en todas ellas. Inicialmente, un mayor número de usuarios utilizan la nueva vía debido a que está diseñada para hacer más rápido el recorrido, sin embargo, esto incrementa el tránsito haciendo el flujo más lento, lo cual provoca que los usuarios se distribuyan en las otras vías llegando al equilibrio de Wardrop cuando se distribuyen en la misma proporción en las tres diferentes rutas, con un tiempo de traslado total de 92 minutos, como se observa en los cálculos del cuadro 1.3. Pero este tiempo de equilibrio resulta 9 minutos mayor que el tiempo de equilibrio del Escenario 1 (casi un 11% más) cuando no existía la vía BC, con lo que queda ilustrada la llamada paradoja de Braess.

Cuadro 1.3. Tiempo de recorrido conforme la distribución de los automóviles en la red cambia en las rutas del escenario 2.

Autos por hora			Tiempo de recorrido de A a D		
vía ABD	vía ABCD	vía ACD	vía ABD	vía ABCD	vía ACD
0	6000	0	0	130	0
500	5000	500	105.5	125	105.5
1000	4000	1000	101	114	101
1500	3000	1500	96.5	103	96.5
2000	2000	2000	92	92	92
2500	1000	2500	87.5	81	87.5
3000	0	3000	83	0	83

En un contexto más general de planeación, está la Postura de Lewis-Mogridge formulada en 1990, que asegura que, a medida que se construyen más vialidades, más tránsito surge para ocuparlas (Wikipedia, 2007b). Según esta afirmación, el aumento de velocidad logrado en nuevas vialidades tiende a desaparecer en un lapso corto. A veces, las nuevas arterias ayudan a reducir embotellamientos, pero en muchos, se observa que la congestión simplemente se traslada a otra intersección.

Otro resultado que se relaciona con la problemática tratada por Braess es la Paradoja de Downs-Thomson (1992). Downs trata de explicar la dificultad de remover la congestión de horas pico de las carreteras. Afirma que la ampliación en las vías de tránsito en una red trae consigo tres efectos inmediatos: a) los usuarios que usan rutas alternativas empiezan a utilizar la vía expandida; b) aquellos usuarios que viajaban anteriormente en tiempos fuera de pico (ya sea antes o después de ellos) empiezan a viajar en las horas pico; y c) el transporte

público deja de usarse para volver a usar vehículos particulares. Estos tres efectos pueden, en algunos casos, empeorar el comportamiento general.

La postura de Lewis-Mogridge y la paradoja de Downs-Thomson no necesariamente concluyen que la construcción de nuevos caminos no está justificada, sino más bien que su desarrollo exige considerar el sistema en su totalidad.

Mientras que en las décadas pasadas se hizo un esfuerzo por desarrollar nuevas vías de transporte terrestre, la época actual tiende al desarrollo de sistemas de optimización de tránsito vehicular sobre las vías ya construidas. El sistema de transporte, tiene su propio y natural ciclo de vida.

Inicialmente el objetivo es la construcción de vías de tránsito vehicular generando una red entre los diferentes lugares de la ciudad, o del país, en su caso. Sin embargo, en la mayoría de las ocasiones, los flujos en estas vías son de magnitudes máximas, aumentando los requerimientos de mantenimiento y ocasionando tránsito lento y congestionamiento. Considerando que se tiene una gran demanda de rehabilitación y mantenimiento del sistema de vías de transporte generalmente con escasos presupuestos ; que existe una gran presión por parte de la sociedad, debido a los costos monetarios asociados al retraso en tiempo por congestionamiento vehicular; que existe una reducida capacidad de espacio para la construcción de nuevas vías vehiculares y que el ampliar las vías no siempre resulta en una mejora en el rendimiento de la red vial, se concluye que es necesario favorecer el desarrollo de sistemas que, con apoyo de la modelación matemática y los métodos de optimización, permitan la mejora del flujo vehicular en las redes viales urbanas.

ALGUNAS SOLUCIONES

El control del congestionamiento vial se ha abordado mediante diversas estrategias, como el desarrollo y mejora del transporte público, la creación de carriles dedicados sólo a autobuses, los días de no circula para ciertas placas, etc.

Según la CEPAL (2001), las principales medidas para controlar la congestión de tránsito provienen de dos fuentes:

Por el lado de la oferta, las principales medidas para controlar el congestionamiento vial son mejoras y mantenimiento adecuado en la red vial, mejoras en las intersecciones y la

administración eficiente de la red de semáforos, la cual debe de considerar ciclos estables y establecimiento de fases de tal manera que los vehículos puedan circular a velocidades específicas minimizando las interrupciones en el viaje por detenciones en las intersecciones semaforizadas. La coordinación puede variar en el curso del día en función del flujo de vehículos con el fin de dar preferencia a aquellos que usan el espacio más eficientemente, como es el caso de los del transporte público.

Por el lado de la demanda, las medidas de control también consideran la participación de los usuarios, a través de una mejora en la educación vial, el corrimiento de horarios de las actividades, las campañas para respetar los lugares de estacionamiento, etc.

Entre estas medidas, el control semaforizado de las redes viales ha sido uno de los primeros intentos de control a gran escala, y sigue teniendo a la fecha un gran potencial para acomodar los flujos vehiculares en las calles y avenidas de modo que se reduzcan los impactos negativos del tránsito.

La forma en que los semáforos han operado a lo largo de la historia ha evolucionado, incorporando los avances tecnológicos de cada época, desde el primer semáforo operado manualmente en Londres en el año de 1868 (Wikipedia, 2007c) hasta aquellos que responden en tiempo real a las condiciones del tránsito vehicular en forma coordinada dentro de un sistema. Aunque este último, llamado sistema de control adaptativo en línea, nunca se ha llevado al cabo totalmente, el esfuerzo de investigación ha generado tecnologías nuevas que muchas ciudades han usado para implantar los sistemas de control computarizado que operan hoy en día.

En la actualidad, la mayoría de los controles ajustan los ciclos y tiempos de verde con base en estadísticas históricas de los flujos observados, asignando tiempos de verde fijos que pueden programarse para distintas horas del día y para distintos días de la semana, presentando una coordinación estable, conforme parámetros estimados; esto es, no cambian en función del flujo de tránsito real, de esta manera, la sincronización se hace obsoleta cuando los parámetros de movimiento de tránsito, con los que fueron programados, presentan variaciones significativas, haciéndose evidente la necesidad de continuar con los esfuerzos en la implementación de controles de tránsito en tiempo real.

En los controles semaforizados de tránsito, el diseño de los tiempos de verde, las longitudes de los ciclos y la coordinación generalmente siguen un proceso de optimización para reducir los impactos, pero el resultado final es un plan de tiempo fijo para cierto periodo del día. La desventaja de estos procedimientos de optimización es el supuesto implícito de que los flujos calculados para el diseño de las señales óptimas seguirán sin cambio luego de su implantación. Bullock and Urbanik (2000) afirman que esta suposición pierde validez si la implementación de la señalización óptima mejora sustancialmente los tiempos de traslado en alguna ruta, ya que los automovilistas que usan rutas alternas en la red vial podrían cambiarse a la ruta mejorada, con lo que los flujos y el nivel de congestión aumentarían en esa ruta mejorada.

La modelación del impacto que los cambios en la señalización de los semáforos puede tener en los flujos vehiculares que circulan en las distintas arterias de la red urbana puede enmarcarse dentro de la Teoría de Asignación de Tránsito. En este contexto, dada una red de transporte (vialidad urbana), donde la señalización de los semáforos se modela explícitamente, se supone que los usuarios eligen sus rutas en la red hasta llegar a un equilibrio en el que los viajeros se redistribuyen en la red vial de modo que para cualquier par origen-destino, el tiempo total de viaje es el mismo sin importar la ruta elegida. Este equilibrio bajo congestión es el llamado equilibrio de Wardrop, que se explica enseguida.

Para cualquier usuario de una red vial, buscar la ruta más corta para ir a su destino resulta natural. Esta búsqueda plantea un problema matemático ya resuelto por la investigación de operaciones en los años 1950. El conocido algoritmo de Dijkstra, que encuentra la ruta más corta en una red formada por nodos y arcos, conociendo las longitudes de los arcos, es un típico ejemplo de este desarrollo. El modelo mencionado puede dar la respuesta a los viajeros que buscan la mejor ruta, si no fuera porque no se trata de un único viajero, sino de muchos más usando la misma vía, y cuando éstos comparten el camino, comienzan a estorbarse entre sí, obligando a todos a reducir la velocidad a la que se puede circular, con lo que la optimalidad de la ruta más corta pierde validez y la congestión empieza a manifestarse.

En 1952, J. G. Wardrop, un ingeniero británico dedicado al análisis de transporte, descubrió el equilibrio que se da en una red vial cuando los usuarios buscan la mejor ruta para ir a su destino, repartiéndose entre las diversas rutas alternativas hasta llegar a un

estado en el cual se tiene la mejor solución posible respecto de la congestión. El equilibrio de Wardrop se enuncia en dos principios (Ortúzar y Willumsen, p. 304-305):

1er. Principio de Wardrop: *En condiciones de equilibrio, en una red congestionada, el flujo se acomoda de modo que todas las rutas utilizadas en un par Origen-Destino dado tienen el mismo costo mínimo, mientras que las rutas no elegidas tienen costos iguales o mayores.*

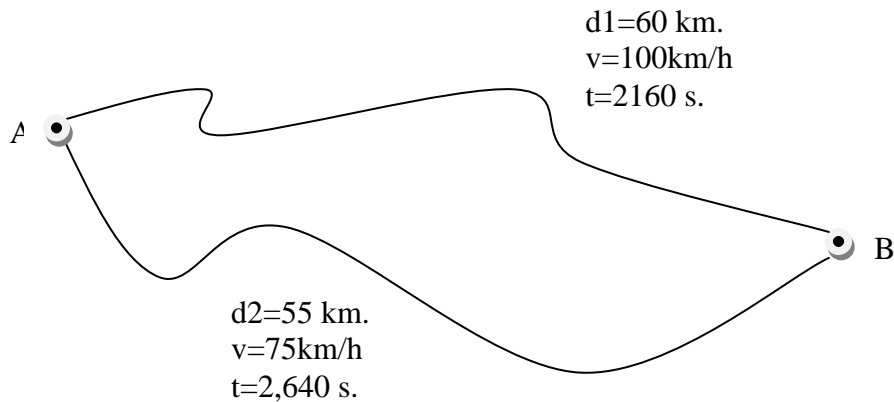
En este enunciado, el costo de la ruta suele interpretarse como el tiempo de recorrido y el principio indica que en equilibrio ningún usuario puede ahorrar tiempo en su viaje si individualmente decide cambiar a otra ruta alternativa. En equilibrio, el efecto de la congestión es mínimo para todos los usuarios. El costo mínimo que describe este primer principio es llamado también Óptimo del Usuario (OU), pues es el resultado final de las decisiones independientes de muchos usuarios, cada uno buscando mejorar su propio tiempo de recorrido.

2do. Principio de Wardrop: *En condiciones de equilibrio social, en una red congestionada el flujo debería acomodarse de manera que el costo total de los usuarios sea mínimo.*

El segundo principio tiene un sentido más bien normativo, en contraste con el carácter individualista de las decisiones de los usuarios en el primer principio. El segundo principio de Wardrop lleva el enfoque del planificador del sistema de transporte, que busca minimizar la suma de los tiempos de recorrido (el costo total) de todos los usuarios en la red vial, ya que así minimiza los impactos de tránsito (gasto de combustible, emisiones de contaminantes, ruido, etc.) y logra un óptimo desde el punto de vista social. El equilibrio logrado en este enfoque es llamado Optimo Social (OS).

Para analizar mejor la situación, conviene examinar un caso sencillo, donde los usuarios tienen que elegir entre dos rutas alternativas para alcanzar su destino (IMT,2007). Considérese el diagrama de la Figura 1.3, para ir del origen A al destino B, donde se muestran las dos rutas posibles, sus distancias, las velocidades de circulación y los tiempos de recorrido en condiciones de flujo libre, es decir, cuando los conductores pueden elegir la velocidad que más les convenga. La ruta 2 es más corta en distancia (55 km.), que la ruta 1, que es más corta en tiempo (36 min.).

Figura 1.3. Ejemplo de equilibrio de Wardrop.



Supongamos ahora que a partir del tiempo de recorrido de flujo libre, por cada auto que entra al camino el recorrido aumenta en 4 segundos en la ruta 1 y en 5 segundos en la ruta 2. Si circulan N vehículos en la ruta 1, su tiempo de recorrido aumenta como sigue:

$$\begin{aligned}
 \text{tiempo 1} &= 2160 \text{ s.} + (4 \text{ s.})N \\
 &= \frac{3}{5} \text{ h.} + \frac{1}{900 \text{ hr}} N \\
 &= \frac{3}{5} + \frac{N}{900} \text{ (en horas)}
 \end{aligned}$$

y como $\text{Velocidad} = \text{distancia} / \text{tiempo}$, la velocidad de circulación de la ruta 1 es:

$$\text{Vel1} = \frac{60 \text{ km.}}{\left(\frac{3}{5} + \frac{N}{900}\right) \text{ h}} = \frac{54,000}{540 + N} \text{ km/h}$$

del mismo modo, con N vehículos circulando en la ruta 2, el tiempo de recorrido es:

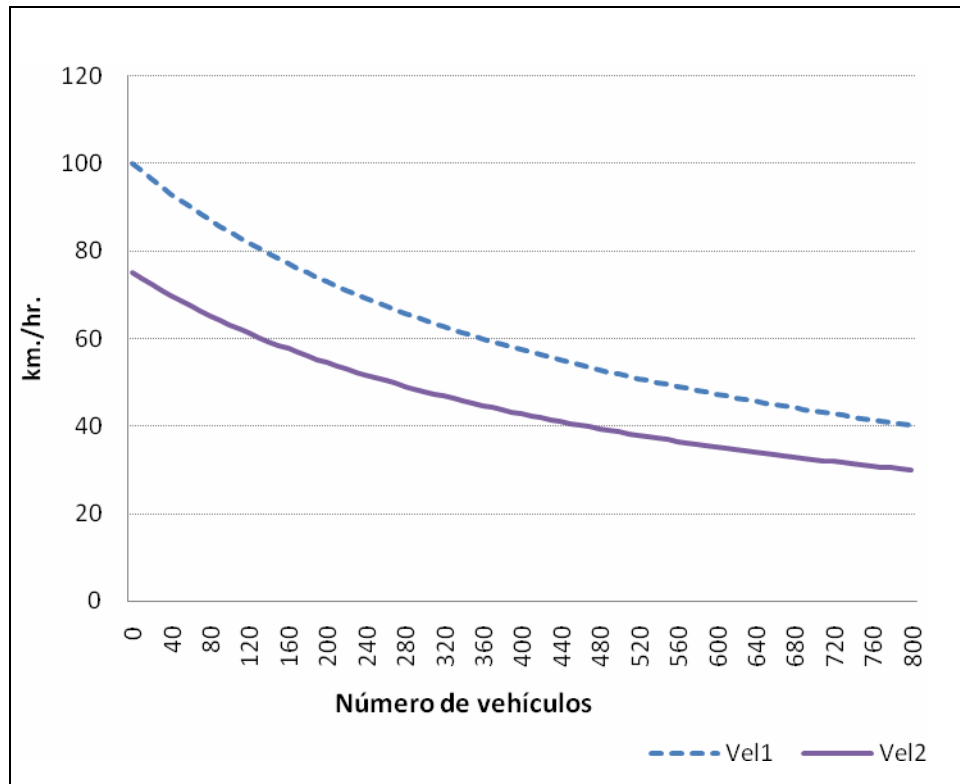
$$\begin{aligned}
 \text{tiempo 2} &= 2640 \text{ s.} + (5 \text{ seg.})N \\
 &= \frac{11}{15} \text{ h} + \frac{1}{720 \text{ h}} N \\
 &= \frac{11}{15} + \frac{N}{720} \text{ (en horas)}
 \end{aligned}$$

y la velocidad de circulación en la ruta 2 resulta:

$$Vel2 = \frac{55km.}{\left(\frac{11}{15} + \frac{N}{720}\right)h} = \frac{39,600}{528 + N} km/h$$

La figura 1.4 muestra la variación de estas velocidades a medida que N aumenta para las dos rutas:

Figura 1.4. Velocidades de circulación en las rutas 1 y 2.

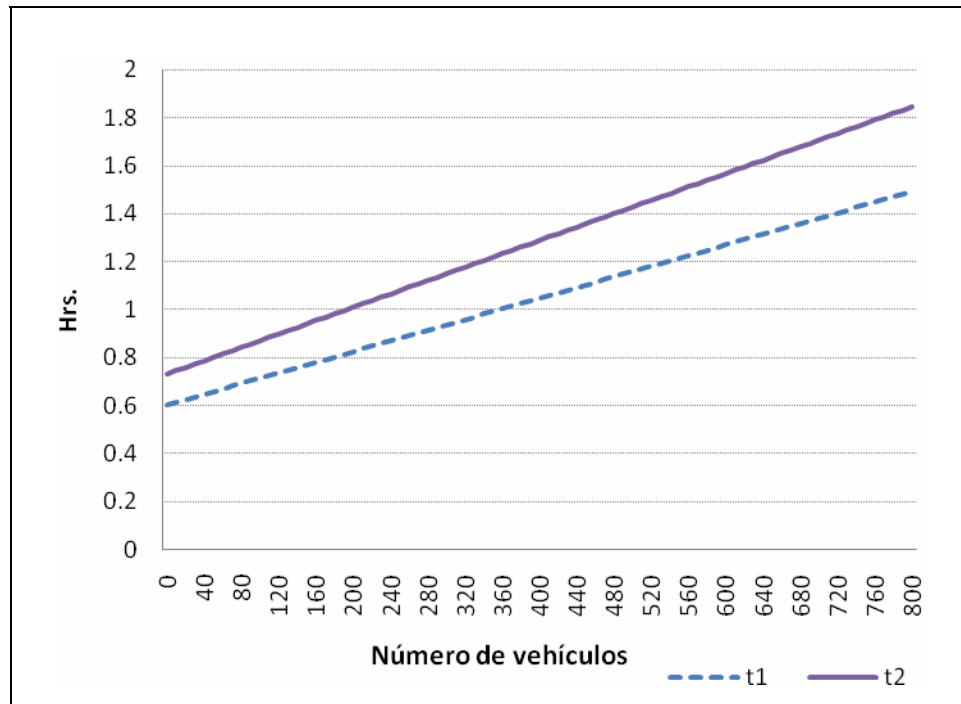


Considerando al tiempo de recorrido como el “costo” (horas) de usar las rutas, su valor está dado por:

$$C_1 = \frac{3}{5} + \frac{N}{900}; \quad C_2 = \frac{11}{15} + \frac{N}{720}$$

donde C_1 es el costo de usar el camino 1 y C_2 el correspondiente al camino 2; la Figura 1.5 ilustra cómo varían estos tiempos a medida que aumenta el número de usuarios en las rutas.

Figura 1.5. Tiempo de recorrido en los caminos.



La situación descrita plantea a los usuarios el problema de elegir entre la ruta 1 o la ruta 2. Comenzando por la ruta más rápida (ruta 1), un único usuario tardaría 36 minutos (0.6 hrs.) en recorrerla, pero mientras más usuarios entran al camino, el tiempo que tardaría cualquiera de ellos aumenta como se ve en la figura 1.5. Así por ejemplo, con 360 autos en la ruta 1, el recorrido es de una hora. Algo semejante pasaría con la ruta 2.

Si suponemos que hay 750 usuarios que desean viajar de A hacia B, la pregunta es entonces: ¿cuántos deben de ir por la ruta 1 y cuántos por la ruta 2 para que los automovilistas se estorben lo menos posible entre ellos? Esta pregunta lleva al concepto de equilibrio en los flujos en el cual la congestión vial llega a un valor mínimo, idea que se formaliza en los principios de Wardrop.

El óptimo del usuario busca minimizar el costo individual de cada usuario de la red vial; la idea corresponde al primer principio de Wardrop, donde ningún usuario individual de la red puede mejorar sus propios costos usando rutas alternativas. La función objetivo para el este problema fue propuesta por Beckman, McGuire y Winsten en 1956. El problema de optimización consiste en minimizar la suma de las áreas bajo las curvas de costo de los tramos de la red, que para el ejemplo se expresa como sigue:

$$MinW = \int C_1(u)du + \int C_2(u)du$$

s.a.

$$N_1 + N_2 = N$$

$$N_1, N_2 \geq 0$$

Donde C_1 y C_2 son las funciones de costo en las rutas 1 y 2, N_1 y N_2 son los volúmenes vehiculares en las rutas 1 y 2.

Reemplazando las funciones de costo del ejemplo presentado e integrando, el problema del óptimo del usuario tiene el siguiente planteamiento:

$$MinW = \int C_1(u)du + \int C_2(u)du = \frac{3}{5}N_1 + \frac{N_1^2}{1800} + \frac{11}{15}N_2 + \frac{N_2^2}{1440}$$

s.a.

$$N_1 + N_2 = 750$$

$$N_1, N_2 \geq 0$$

Cuya solución óptima es $N_1=470$; $N_2=280$, y los correspondientes costos de usuario $C_1(470)=C_2(280)=1.122h.$, esto es, si 470 vehículos eligen la ruta 1 y el resto elige la ruta 2, cualquier automovilista tardará el mismo tiempo de 1.122 h. sin importar la ruta elegida; éste es el estado de equilibrio que logra el óptimo del usuario. El costo total es la suma de los costos de todos los usuarios, esto es: $750(1.122) = 841.50$ h.

Por otra parte, el óptimo social busca minimizar los impactos negativos de los flujos vehiculares, lo cual se modela minimizando el costo total asociado a los flujos que circulan en la red vial, y corresponde al segundo principio de Wardrop. Así el problema de optimización social es:

$$MinW = C_1(N_1)N_1 + C_2(N_2)N_2$$

s.a.

$$N_1 + N_2 = N$$

$$N_1, N_2 \geq 0$$

Reemplazando las correspondientes funciones de costo del problema, resulta:

$$MinW = \frac{3}{5}N_1 + \frac{N_1^2}{900} + \frac{11}{15}N_2 + \frac{N_2^2}{720}$$

s.a.

$$N_1 + N_2 = 750$$

$$N_1, N_2 \geq 0$$

Cuya solución óptima se da en $N_1=443$ y $N_2=307$, con los correspondientes costos de usuario en $C_1(443)=1.093$ h. y $C_2(307)=1.159$ h. El costo total mínimo es de $443(1.093)+307(1.159) = 840$ h. Nótese que en este equilibrio social los automovilistas de la ruta 2 parece que anteponen el interés colectivo al individual, pues aceptan un tiempo de recorrido más largo (1.159 h.) comparado al óptimo del usuario calculado anteriormente (1.122 h.). Claramente, el óptimo social difícilmente ocurre de manera espontánea, y en la práctica siempre requiere la intervención del planificador o de la autoridad vial para imponer peajes, prohibiciones a la circulación, discriminación de tipo de vehículo y otras medidas limitantes que buscan lograr ese óptimo social.

En el caso del control semaforizado, un cambio en los tiempos de verde, en los ciclos o en la coordinación de los semáforos que pertenecen a la red, romperá el equilibrio del usuario al alterar los tiempos de recorrido en las rutas, lo que a su vez hará que cambien sus rutas para mejorar sus tiempos y por tanto cambien los flujos en la red. El modelado de la condición de equilibrio del usuario debe ser usado como parte del proceso de optimización de las señales del sistema semaforizado, concretamente, representándose como una restricción del problema de optimización, con la intención de tener una estimación anticipada del impacto que los cambios en la señalización tendrán en los flujos circulando en la red.

El problema de determinar los tiempos óptimos de señalización y su programación (coordinación, tiempo de verde y duración del ciclo) en la red semaforizada, prevé la respuesta de equilibrio de los automovilistas y es un caso particular del llamado Problema de Diseño de Red (Network Design Problem, NDP) (Yin, 2000). En un NDP el objetivo es mejorar el desempeño de una red de transporte dada con respecto a algunas variables de diseño discretas, continuas o mixtas. Si además se considera al mismo tiempo la respuesta de los usuarios a los cambios de señalización, se tiene lo que se llama un problema de optimización con estructura bi-nivel. En el nivel superior, se tiene la función objetivo del planificador para “diseñar la red” eligiendo las variables de diseño en busca de un óptimo de

tipo social. En el nivel inferior, los automovilistas responden al diseño cambiando su elección de las vías a utilizar de manera de alcanzar el óptimo del usuario, lo cual podría no coincidir con los flujos que esperaría el planificador.

Formalmente, el problema de optimización del planificador en el nivel superior tiene como una de sus restricciones el problema del equilibrio del usuario en el nivel inferior. Esta clase de problemas bi-nivel es uno de los retos actuales en el campo de la optimización, debido a que las funciones objetivo que se manejan en los problemas de diseño de red en general no son suaves y pudieran no tener derivadas en todo su dominio, además de que las regiones factibles del problema suelen ser no-convexas. Debido a esto, se tiene la posible existencia de óptimos múltiples, así como una gran dificultad para desarrollar métodos eficientes y robustos para encontrar incluso los óptimos locales.

El problema de optimización así presentado, en donde la función objetivo planteada puede ser no lineal y no necesariamente diferenciable, y en donde las restricciones involucran variables tanto discretas como continuas, invita a buscar la aplicación de métodos de optimización heurísticos tales como los Algoritmos Genéticos (AG).

Los AG nacen en los años 1970, con las contribuciones de John Holland, son métodos heurísticos aplicados a la búsqueda de óptimos. Son llamados así porque se inspiran en la evolución biológica y su base genético-molecular. Estos algoritmos hacen evolucionar una población de soluciones someténdola a acciones aleatorias semejantes a las que actúan en la evolución biológica (mutaciones y recombinaciones genéticas), así como también a una selección de acuerdo con algún criterio, en función del cual se decide cuáles son los individuos más adaptados, que sobreviven, y cuáles los menos aptos, que son descartados. Los algoritmos genéticos son eficaces en los casos en los que la función objetivo pudiera ser no diferenciable o bien cuando la región de factibilidad es no convexa, aunque su uso es posible con cualquier tipo de función. En general, el esquema del algoritmo esta basado en los siguientes pasos:

1. Iniciar con una población de soluciones al problema, las cuales son generadas de manera aleatoria y representadas por una estructura de datos adecuada.
2. Evaluar cada una de las soluciones, y asignarle una puntuación según lo bien que se adapten.

3. Escoger de la población la parte que tenga una puntuación mayor
4. Mutar (cambiar) y entrecruzar (combinar) las diferentes soluciones de los elementos seleccionados, para reconstruir la población.
5. Repetir un número predeterminado de veces, o hasta que se haya encontrado la solución deseada.

Algunos otros modelos para la optimización de semáforos han sido desarrollados basados en el control adaptativo, el cual responde en tiempo real a las condiciones de tránsito vehicular con base en información de las intersecciones viales cercanas. Para implementar este tipo de controles se requiere contar con un diseño óptimo del sistema de comunicaciones de señal entre los semáforos. Entre los principales trabajos desarrollados bajo este enfoque se encuentra el Sistema RHODES desarrollado desde 1991 (Head, Mirchandani, and Shelby, 2005), en la Universidad de Arizona, el cual ha sido implementado en intersecciones en Tempe, Arizona; Tucson; Seattle, Washington, y Santa Clara, California a manera de prueba piloto; el modelo **OPAC**, implementado en algunos sistemas en la ciudad de Reston, VA y el modelo **RTACL** implementado en Chicago, IL; entre otros.

El modelo que se presenta en el presente trabajo apunta a la optimización de tiempo de transporte total en una vía. Este objetivo se modela minimizando el costo total asociado a los flujos que circulan en la red vial, considerando como parte de las restricciones, el equilibrio del usuario como una función de los programas de operación de los controles semaforizados.

HIPÓTESIS DEL TRABAJO

Es posible determinar la programación en el control semaforizado que minimice los tiempos de traslado dentro de una red vial, integrando la respuesta de los usuarios a la “impedancia” que representa la programación de los semáforos en la elección de sus rutas. Los tiempos óptimos de señal pueden ser encontrados a partir de la modelación de esta conducta basada en el equilibrio Wardropiano y de las bondades de los Algoritmos Genéticos, en combinación con software especializado de transporte (TransCad) que encuentra esos equilibrios del usuario.

II. REVISIÓN DE LITERATURA

Conscientes de los problemas de tránsito vehicular presentes actualmente en el Municipio de Querétaro y los que se proyectan para el futuro cercano, las administraciones municipales ha hecho esfuerzos permanentes por encontrar soluciones. El Departamento de Control y Semaforización de la Dirección de Tránsito Municipal (SSPM), que depende directamente de la Secretaría de Seguridad Pública Municipal, se encuentra trabajando permanentemente en la coordinación de las señales en vías secundarias. Asimismo, a través de la SSPM, el Municipio recientemente solicitó al Conacyt, como parte del Plan Estratégico de Tránsito y Vialidad, la realización de estudios especializados; algunos de los objetivos de esta solicitud son la regularización del tránsito vial, la identificación de puntos viales de conflicto y la tipificación de la problemática de señalización de la red vial. La literatura de control señalizado del tránsito reporta que en algunos países desarrollados se tienen implementadas soluciones basadas en tecnología de punta que dan una respuesta inmediata a las condiciones de tránsito vehicular que se presentan. El enfoque de esta investigación esta dirigido a la programación de ciclos en semáforos basados en modelos de optimización de tránsito vehicular. Como se mencionó en el capítulo I, esta modelación tiene como marco la Teoría de Asignación de Tránsito, considerando el impacto que los cambios en los parámetros de los semáforos pueden tener en los flujos vehiculares que circulan en las distintas arterias de la red urbana.

Teoría de Asignación de Tránsito y Modelos de Equilibrio

La Teoría de Asignación de Tránsito tiene como objetivo predecir los flujos dentro de una red de transporte, que en este caso será la red vial urbana, asociados con ciertos escenarios hipotéticos, para generar estimaciones de las rutas de viaje.

La planeación del transporte urbano ha tenido un proceso evolutivo. A principios de los años 50's los estudios se enfocaban en la investigación del desarrollo de nuevas vías y al pronóstico del tránsito en el corto plazo, los métodos utilizados en la primera etapa de la planeación estaban enfocados en el pronóstico de poblaciones futuras y su distribución, en el análisis del comportamiento de los movimientos de personas y en las características de la población (poseedores de automóviles o usuarios de transporte público). Los modelos que se enfocaban en la planeación de redes viales, no consideraban rutas individuales e ignoraban

el hecho que el tiempo de viaje depende del flujo o volumen del tránsito de las rutas congestionadas y que múltiples vías pueden servir para un par origen - destino determinado (O-D).

Con el avance tecnológico y computacional, los métodos de planeación evolucionaron rápidamente hasta llegar a métodos de gran sofisticación, considerando más factores en la modelación del tránsito vehicular. Los modelos más evolucionados son conocidos como Modelos de Equilibrio y consideran el tiempo de viaje como una función del volumen de automóviles en la vía, resultando flujos de rutas y tiempos de viaje consistentes entre sí. Los algoritmos de equilibrio de flujo son iterativos en el proceso de asignación y el proceso de cálculo del tiempo de viaje.

Dependiendo del propósito del estudio de transporte, los modelos que se han propuesto contienen diferentes componentes del sistema de transporte (patrones de desarrollo de vialidades, políticas de control, distribución y generación de viajes, etc.), diferentes niveles de agregación de la realidad física (modelos macroscópicos y microscópicos), diferentes horizontes de planeación (desde modelos basados en aforos hasta el uso de sistemas de manejo de tránsito en tiempo real) y pueden estar basados en diferentes principios (modelos estadísticos, modelos de optimización, modelos de simulación).

El problema de encontrar el patrón de flujo de equilibrio sobre una red urbana dada es conocido como Asignación de Tránsito. Desde la perspectiva de comportamiento, los métodos de Asignación de Tránsito son el resultado de la agregación de elecciones de ruta individual de los viajeros. Wardrop (1952) propuso un modelo de equilibrio¹ cuya hipótesis fundamental es que los usuarios de la red de tránsito vehicular tienen información perfecta referente a los atributos de las rutas alternativas para un determinado origen-destino, de tal suerte que cada uno de ellos elige la ruta que minimiza el tiempo de viaje. De esta manera la red llega a un equilibrio en el que los automóviles se distribuyen en las vías de modo que

¹ El equilibrio de Wardrop expresa el mismo concepto del Equilibrio de Nash en Teoría de Juegos, es la situación en la que ningún jugador puede cambiar unilateralmente de estrategia sin salir perjudicado. Ambos principios fueron desarrollados por separado.

para cualquier par origen-destino, el tiempo total de viaje es el mismo no importando la ruta elegida.

El modelo de equilibrio de Wardrop utiliza un proceso iterativo para alcanzar una solución; en cada iteración los flujos de las rutas en la red son calculados incorporando los efectos de las restricciones de capacidad de cada ruta y los tiempos de viaje como funciones del flujo.

La hipótesis del modelo de Wardrop puede ser realista mediante la implementación de sistemas de información al usuario respecto al congestionamiento de las diversas vías en tiempo real, estos sistemas se encuentran funcionando ya en algunas ciudades del mundo.

Un modelo de equilibrio alternativo fue propuesto por Daganzo y Sheffi (1977) conocido como el Equilibrio Estocástico, el cual está basado en la hipótesis de que los usuarios tienen información imperfecta acerca de las condiciones de las rutas de la red, en este caso los usuarios consideran que pueden mejorar su utilidad esperada eligiendo una ruta diferente. Este modelo produce resultados más reales cuando no se cuenta con sistemas de información a los usuarios respecto a la situación real del tránsito vehicular, permitiendo ponderar las rutas conforme a su uso regular clasificándolas como rutas “más atractivas” y “menos atractivas”; las rutas “menos atractivas” son aquellas con menor flujo vehicular. Un método convergente utilizado con frecuencia para resolver el equilibrio en este modelo es el Método de Promedios Sucesivos (Sheffi and Powell, 1982; Sheffi, 1985).

Los dos modelos de equilibrio planteados operan a partir de una matriz de flujos que indican el volumen de tránsito entre cada par origen - destino (O-D) y están basados en el tiempo de traslado en rutas alternativas. El tiempo de traslado es considerado como una función del volumen o grado de saturación de la ruta y es conocido como función de impedancia. Una de las funciones que con más frecuencia es utilizada en el cálculo del equilibrio es la llamada BPR (Bureau of Public Roads), la que establece que el tiempo de traslado esta dado por:

$$t = t_f \left[1 + \alpha \left(\frac{v}{c} \right)^\beta \right]$$

En donde:

- t : tiempo de viaje sobre la ruta congestionada.
- t_f : tiempo de viaje considerando flujo libre en la ruta.
- v : volumen de la ruta.
- c : capacidad de la ruta.
- α, β : parámetros de calibración.

Mientras que diferentes formulaciones de esta función han sido propuestas (Branston, 1976; Davidson, 1966), la función BPR (Traffic Assignment Manual, BPR, 1964) es muy adecuada para ser utilizada en los modelos de asignación de tránsito. Con una elección correcta de los parámetros, esta función puede representar una gran variedad de relaciones flujo-retraso (incluso las relaciones representadas por las otras funciones que se han propuesto).

Wardrop presenta su modelo de equilibrio en 1952 y sólo cuatro años más tarde Beckman et. al. (1956) proponen un planteamiento matemático riguroso para expresar este principio formalmente, cuya resolución se plantea mediante un algoritmo matemático, sin embargo, tomó muchos años el planteamiento de algoritmos que fueran viables en su implementación práctica (LeBlanc, 1973).

Beckman et. al. comparan el problema de asignación de equilibrio con problemas de equilibrio en el contexto mecánico. Ellos demuestran que, bajo las hipótesis de que el costo, C_a , de cualquier ruta a , es una función del flujo x_a , y de que las funciones del comportamiento de la ruta son crecientes, entonces los flujos x_a , satisfaciendo el principio de Wardrop, son únicos e iguales a aquellos que satisfacen:

$$\min z(x) = \sum_a \int_0^{x_a} C_a(u) du$$

El método utilizado para resolver este problema es el algoritmo de combinación convexa, originalmente propuesto por Frank y Wolfe en 1956 (Frank and Wolfe, 1956), el cual es un procedimiento para resolver problemas de programación cuadrática con

restricciones lineales. LeBlanc (1973) aplica el método Frank-Wolfe obteniendo un algoritmo eficiente para resolver este problema.

El costo de viaje, C_a , puede incluir varios componentes, reflejando el tiempo de viaje, el número de “altos”, la seguridad, el combustible gastado, etc. Algunos de estos componentes pueden ser expresados en términos del tiempo de viaje t_a , por lo que es común utilizar éste, en lugar del costo de viaje C_a .

El problema del equilibrio de Wardrop se formula como sigue:

$$\min z(x) = \sum_{ij} \int_0^{x_{ij}} t_{ij}(u) du$$

$$s.a. \quad D(j, s) + \sum_{ij} x_{ij}^s = \sum_k x_{jk}^s, \quad s = 1, 2, \dots, p$$

$$j = 1, 2, \dots, n$$

$$j \neq s$$

$$x_{ij}^s \geq 0$$

En donde:

n = El número de nodos en la red vial

p = El número de orígenes y destinos en la red.

x_{ij} = El flujo total en el arco (i, j) .

x_{ij}^s = El flujo en el arco (i, j) con destino s .

t_{ij} = Representa la relación entre el flujo y el tiempo de viaje del arco ij , representa la forma en que cambia el tiempo de recorrido con el flujo en el tramo.

$D(j, s)$ = El flujo que se origina en el nodo j con destino s .

La primera de las ecuaciones en las restricciones es la ecuación de conservación de flujo y las demás ecuaciones son restricciones de no negatividad de flujos.

LeBlanc demostró que este problema de optimización es convexo respecto a los flujos x_{ij} , sin embargo, es un problema no convexo con respecto a los flujos de la ruta x^s_{ij} , lo cual implica que la condición de equilibrio no es única en relación con los flujos de la ruta.

Zangwill, en 1969, desarrolla la expansión de primer orden de $z(x)$ alrededor de x^n , dado un vector de flujo factible, x^n (que satisface la ecuación de conservación de flujo y la restricción de no negatividad):

$$z(y) = z(x^n) + \nabla z(x^n + \theta(y - x^n))(y - x^n) \quad \text{para } 0 \leq \theta \leq 1$$

Considera que una aproximación lineal conveniente para $z(y)$ se genera al asignar $\theta=0$ con lo cual:

$$z(y) = z(x^n) + \nabla z(x^n)(y - x^n)$$

Con el fin de encontrar una dirección de búsqueda en la que la función objetivo disminuya su valor, resuelve el siguiente problema de programación lineal:

$$LP: \quad \min \quad z(y) = z(x^1) + \nabla z(x^1)(y - x^1)$$

Después de algunas manipulaciones de esta última ecuación y una vez removidos los términos constantes, produce la siguiente función objetivo:

$$LP: \quad \min \quad \nabla z(x^n)y$$

$$s.a. \quad D(j, s) + \sum_{ij} x_{ij}^s = \sum_k x_{jk}^s \quad s = 1, 2..p$$

$$j = 1, 2..n$$

$$j \neq s$$

$$x_{ij}^s \geq 0,$$

El vector solución y^n a este problema es una solución factible del problema no lineal $LP: \quad \min \quad z(y) = z(x^1) + \nabla z(x^1)(y - z^1)$. De esta manera, concluye que la dirección $d^n = y^n - x^n$ es la dirección de búsqueda de un valor menor de z .

En el estudio de Zangwill, se observa que la región factible es convexa ya que se genera por las ecuaciones conservación de flujo las cuales son lineales. La búsqueda de la

dirección de descenso d^n está determinada por la solución de un problema lineal planteado en la ecuación anterior, por lo que y^n se encuentra en la frontera de la región factible. De acuerdo con este método, el vector solución de la siguiente iteración, x^{n+1} se encuentra en el segmento de recta definido por x^n y y^n . Se puede observar que la determinación de la dirección de descenso, automáticamente genera una cota para la línea de búsqueda.

Con el fin de encontrar el siguiente punto (el vector de flujo de la siguiente iteración), $z(x)$ tiene que ser minimizada a lo largo de $d^n = (y^n - x^n)$:

$$\begin{aligned} \min z[x^n + \alpha(y^n - x^n)] \\ \text{sujeto a } 0 \leq \alpha \leq 1 \end{aligned}$$

Como los intervalos de búsqueda están segmentados, esta minimización se logra aplicando métodos de reducción por intervalos. Una vez que la solución se ha encontrado, el siguiente punto es calculado mediante:

$$x^{n+1} = x^n + \alpha(y^n - x^n)$$

La ecuación $\min z[x^n + \alpha(y^n - x^n)]$ revela que :

$$\nabla z(x^n) = \min \sum_{ijs} \frac{\partial z(x^n)}{\partial x_{ij}^s} y_{ij}^s$$

La derivada se obtiene por regla de la cadena:

$$\frac{\partial z(x_{ij})}{\partial x_{ij}^s} = \frac{\partial z(x_{ij})}{\partial x_{ij}} \frac{\partial x_{ij}}{\partial x_{ij}^s}$$

Es claro que:

$$\frac{\partial z(x_{ij})}{\partial x_{ij}^s} = t(x_{ij})$$

El vector x_{ij} puede ser escrito como $x_{ij} = \sum_s x_{ij}^s = x_{ij}^1 + x_{ij}^2 \dots + x_{ij}^s + \dots$, por lo que $\frac{\partial x_{ij}}{\partial x_{ij}^s} = 1$.

Si se define $c_{ij} = t(x_{ij} / x = x^n)$, el programa lineal LP puede ser escrito como:

$$\min \sum_s \sum_{ij} c_{ij} y_{ij}^s$$

Este último planteamiento es equivalente a minimizar el tiempo total de viaje sobre una red con tiempos fijos (no dependientes del flujo), esto es, encontrando la ruta más corta que conecta cada par origen destino y asignando el flujo a ésta.

El algoritmo presentado puede ser resumido como sigue:

Inicio.

1.- Desarrollar Asignación Todo o Nada² basada en $t_{ij} = t_{ij}(0)$. Esto produce el vector de flujo x^1 . Se asigna al número de iteración $n=1$.

2.- Actualizar el tiempo de viaje

$$t_{ij}^n = t_{ij}(x_{ij}^n)$$

3.- Encontrar la dirección de búsqueda desarrollando Asignación Todo o Nada con t_{ij}^n lo que produce el vector de flujo auxiliar y_{ij}^n

4.- Búsqueda de

$$\min_{0 \leq \alpha \leq 1} \sum_{ij} \int_0^{x_{ij}^s + \alpha(y_{ij}^n - x_{ij}^s)} t_{ij}(u) du$$

² La Asignación Todo o Nada es un algoritmo de ruta mínima. La ruta mínima representa la ruta de tiempo mínimo entre dos nodos a la cual se le asigna todo el volumen de tránsito entre los nodos en cuestión. Como el volumen y el tiempo de viaje son crecientes, los resultados de este método pudieran no converger.

5.-Determinar el valor x_{ij} de la siguiente iteración sobre la dirección encontrada:

$$x_{ij}^{n+1} = x_{ij}^n + \alpha(y_{ij}^n - x_{ij}^n)$$

6.- Prueba de convergencia, esto es, cuando el resultado óptimo sobre dos iteraciones consecutivas es menor a un número determinado, el cual en la mayoría de las ocasiones tiende a ser cero, el proceso termina considerando que se ha encontrado una solución.

Por otro lado, para resolver el modelo de Equilibrio Estocástico se aplica el método de promedios sucesivos (Powell y Sheffi, 1982); utilizando una secuencia predeterminada de tamaños de pasos de la siguiente forma:

$$\alpha_n = \frac{K_1}{K_2 + n}$$

en donde:

α_n : tamaño de paso.

n : el número de iteración.

K_1, K_2 : parámetros de la ecuación elegidos de tal manera que:

$$\sum_{n=1}^{\infty} \alpha_n \rightarrow \infty .$$

$$\sum_{n=1}^{\infty} \alpha_n^2 < \infty$$

La primera condición garantiza que la sucesión alcance valores deseados, no importando con qué valor se comience, mientras que la segunda condición garantiza que la varianza de la variable aleatoria disminuya en cada iteración. Una de las secuencias más simples del tamaño de paso que satisface ambas condiciones es

$$\alpha_n = \frac{1}{n} \quad (K_1 = 1, K_2 = 0)$$

Problema de Diseño de Red. Modelo de Optimización Bi-nivel.

Tradicionalmente, los modelos de Asignación tratan las intersecciones de manera simplista, los efectos del comportamiento de los semáforos en la asignación son tratados como penalizaciones fijas o en función al flujo. Cantarella et al. (1991) aseguran la existencia de una fuerte relación entre los patrones de flujo y la programación de los semáforos ya que los usuarios hacen opciones de ruta que varían con las estrategias de control, y su comportamiento asociado a las características de la red. Así mismo, Allsop (1974) reporta la interdependencia que existe entre los parámetros de las señales de tránsito y la asignación (la opción de ruta):

“Cuando toda o parte de la red está sujeta a control, la relación entre el costo de viaje y el flujo de tránsito en algunas o todas las rutas de la red depende de los parámetros de control, y éstos pueden ser utilizados para influenciar el número de viajes hechos a través de la red y de las rutas tomadas”.

En el llamado Problema de Diseño de Red con estructura bi-nivel, el objetivo es optimizar el desempeño de la red vial, en donde las variables de decisión son, en este caso, los parámetros de las señalizaciones (coordinación, tiempo de verde y duración del ciclo), considerando dentro de las restricciones la respuesta de los usuarios a los cambios de señalización. El problema es considerado pues, como un problema de optimización con dos niveles; en el nivel superior se encuentra la función de impedancia a optimizar y en un segundo nivel, como una restricción, se encuentra la respuesta de usuarios a los cambios en los parámetros de diseño de red buscando optimizar su propia función de impedancia.

Este problema puede ser resuelto como una aproximación global de optimización, suponiendo que el desarrollo del sistema depende de los patrones de flujo y de los parámetros de control, sujeta a la restricción de conservación de flujo y a las restricciones adicionales expresadas en el principio de equilibrio de Wardrop:

$$\min \sum_i x_i t_i(x_i, g)$$

s.a. Ecuaciones de conservación de flujo.

Condiciones que satisfacen el Equilibrio de Usuario.

En donde g representa la longitud del tiempo de verde, x_i el grado de saturación del arco i , y t_i , el tiempo de viaje en el arco i , el cual depende de x_i y de g .

La optimización global es difícil de alcanzar debido a la restricción no lineal presente. Una dificultad adicional es la relación no convexa existente entre la demora vehicular y la programación de las señales de tránsito. Una gama de métodos han sido propuestos para resolver este tipo de problemas, desde métodos heurísticos iterativos (Steenbrick, 1974; Allsop, 1974; Suwansirikul et. al., 1987), métodos de linealización (LeBlanc & Boyce, 1986; Ben Ayed et. al., 1988), métodos basados en sensibilidad (Friesz et. al., 1990, Yang, 1997), métodos basados en las condiciones de Karush-Kuhn-Tucker (Marcotte, 1981; Verhoef, 2002), métodos desarrollados a partir de sistemas de solución óptima (Dantzig et. al. 1979; Marcotte, 1981; Bergendorff et. al., 1997; Hearn and Ramana, 1998), métodos de funciones marginales (Meng et.al., 2001), métodos de curvas de nivel (Lawphongpanich y Hearn, 2004), Métodos de Búsqueda Estocástica (Friesz et. al., 1992; Ceylan & Bell, 2004) hasta métodos meta-heurísticos (Teklu, et. al., 2007).

Debido a la naturaleza del problema, los métodos heurísticos iterativos se han presentado como procedimientos comunes de solución. Allsop (1974) propone un algoritmo de asignación óptima iterativo el cual consiste en iteraciones sucesivas entre el procedimiento de optimización de señal y el programa de asignación de flujo de tránsito. Durante la fase de asignación, los parámetros de control de tránsito se suponen fijos y durante la fase de optimización de señal, el patrón de flujo de tránsito se supone fijo. El procedimiento iterativo no necesariamente converge a la solución óptima global (Harker y Friesz, 1984). Dickinson (1981) muestra que esto puede ocurrir cuando el costo total de la red se incrementa de una iteración a otra llevando a una solución no óptima.

El algoritmo iterativo puede ser formulado como el siguiente proceso de dos fases:

Fase 1: Asignación de Equilibrio de Usuario:

$$\min z(x) = \sum_i \int_0^{x_i} t(u, g) |_{g \text{ fijo}} du$$

s.a. ecuaciones de conservación de flujo.

Fase 2: Optimización de señal.

$$\min J[x(g), g]_{x \text{ fijo}}$$

s.a. viabilidad de los parámetros de control.

Los parámetros iniciales de las señales de tránsito se suponen fijos. Con este conjunto de parámetros de control se encuentra una asignación convergente. El flujo resultante es utilizado para optimizar los parámetros de las señales de tránsito basado en el método Webster (Webster y Cobbe, 1966). Este procedimiento se repite hasta que el cambio entre dos iteraciones sucesivas es menor que un valor de convergencia específico.

Teklu, Sumalee y Watling (2007), presentan la aplicación de métodos metaheurísticos denominados Algoritmos Genéticos (AG) como algoritmos de optimización en el Problema de Diseño de Red como una buena alternativa debido a que no requieren información del gradiente de la función objetivo y a que tienen capacidad de escapar de óptimos locales. Otra gran ventaja de aplicar AG en este tipo de problemas es que su marco de desarrollo está basado en simulación, la cual es utilizada en muchos métodos de diseño de señales (p ej en software especializado como Transyt, TransCad, Saturn) que no usan funciones matemáticas explícitas, evitando el uso de métodos de optimización analíticos.

Los algoritmos genéticos (Montana, T. Hussain, and T. Saxena, 2002) forman parte de la investigación en la inteligencia artificial y fueron diseñados para resolver problemas de optimización utilizando el modelo de evolución y la selección natural de los seres vivos.

Para comenzar, un AG debe definir dos cosas: a) la representación genética (llamada cromosoma) del dominio de soluciones y b) una función de aptitud (o sea la función objetivo a optimizar) para evaluar las soluciones del dominio. El esquema general de un AG es descrito por Hillier, F.S. y Lieberman, G.J. (2006) como sigue:

Inicio. Se genera una población de soluciones de prueba factibles, posiblemente de forma aleatoria, del problema de optimización y se evalúa la aptitud (el valor de la función objetivo) de cada elemento de esa población.

Paso Iterativo. Es un proceso aleatorio que está sesgado para elegir a los miembros más aptos (con mejores valores de la función objetivo) de la población en turno para

seleccionar algunos de ellos como padres y generar hijos. Estos padres se aparean combinando sus cromosomas (usualmente en forma de números binarios) al intercambiar de manera aleatoria fragmentos del mismo tamaño de su representación binaria con lo que se producen nuevas soluciones (hijos); esta operación es llamada entrecruzamiento (crossover). A veces se incluye una operación de mutación en la cual un 0 es reemplazado por un 1 o viceversa en alguna posición del cromosoma elegida aleatoriamente. Con los hijos generados y una parte de los miembros más aptos de la primera generación se forma una nueva población del mismo tamaño de la original para la siguiente iteración; el resto de miembros menos aptos de la población original es descartado. La siguiente iteración inicia evaluando la aptitud de cada elemento de la nueva población.

Regla de terminación. Es una regla para decidir cuándo detener el algoritmo, usualmente esto ocurre: a) si se encuentra una solución que satisface algún criterio de optimalidad; b) luego de un número fijo de generaciones evaluadas, c) si luego de un número prefijado de iteraciones ya no hay mejora a la mejor solución encontrada hasta ese punto.

En los métodos de diseño de señales, los AG han sido utilizados en la optimización de tiempos de señal, tiempos de verde y coordinación (Foy et. al,1992; Park et. al. 2000; Park & Yun, 2005), sin embargo, tales planteamientos no consideran la respuesta de los usuarios a cambios en la sincronización, además de que su aplicación se limita a redes viales muy pequeñas. Lee (1998), Taale & Van Zuylen (2003) y Ceylan & Bell (2004) utilizan AG en optimización de tiempo de señal anticipando los impactos del cambio en ruta de los usuarios, para optimizar la longitud de la fase y el tiempo de ciclo que minimicen el tiempo total de traslado. Lee (1998) presenta un comparativo entre AG y el método de Recocido Simulado con algoritmos de búsqueda local iterativa y muestra que la elección del algoritmo debe de hacerse en función de las características de suministro de la red y diferentes escenarios de demanda. Taale & Van Zuylen (2003) reportan resultados prometedores en el empleo de los AG en la optimización de tiempos de verde en el contexto del Problema de Diseño de Red aplicados en redes viales pequeñas. En su aproximación para solucionar el Problema de Diseño de Red, Ceylan & Bell (2004) usan la inversa de un índice de comportamiento definida como la suma ponderada de retrasos y altos de todo el flujo de

tránsito en la red, en donde las variables de decisión son el ciclo y longitud de verde, considerando los flujos en la ruta restringidos por el Equilibrio Estocástico (Daganzo y Sheffi, 1977). Utilizando una red pequeña, muestran que la optimización del problema bi-nivel con GA da resultados más eficientes que los algoritmos iterativos en términos de costos de viaje en sistemas amplios.

Teklu, Sumalee y Watling (2007) resuelven el problema de equilibrio de tránsito usando simulación de los modelos de asignación con énfasis en las características de las intersecciones, movimientos de vuelta y conflictos asociados a las mismas. Basado en modelos de flujos cíclicos (Robertson, 1969), el submodelo de simulación determina los retrasos en intersecciones de un número de estados de flujo, basado en estimaciones de demandas obtenidas del submodelo de asignación. El submodelo de asignación utiliza esta información para estimar las relaciones marginales entre el tiempo de viaje en cada arco y el flujo en ese arco y estas funciones son posteriormente regresadas al submodelo de asignación para recalculan las demandas de rutas. El submodelo de asignación utiliza el método de combinación convexa Frank-Wolfe para determinar el equilibrio en el patrón de flujo. Asimismo, para una función de tiempo-flujo de viaje el submodelo de asignación converge a una solución de equilibrio, el proceso mediante el cual los modelos de asignación y simulación son alternativamente resueltos son heurísticos. Es importante señalar que el uso extendido de esta heurística provee evidencia numérica de que puede ser asegurada una buena convergencia.

De esta manera, la organización de los flujos viales por medio de una programación adecuada en las señales de tránsito, se plantea como una solución viable a los problemas de congestión vial en las grandes urbes. Se ha demostrado que la optimización del tiempo de traslado debe buscarse considerando la respuesta que los usuarios tienen a los cambios en las señales de tránsito, esto es, a través de un problema de optimización bi-nivel (Yin, 2000). Este tipo de problemas poseen características que invitan a la aplicación de métodos heurísticos para su solución.

III. METODOLOGÍA

El equilibrio bajo congestión se refiere al fenómeno que ocurre cuando los usuarios eligen sus propias rutas para dirigirse de un destino a un origen determinado. Los usuarios se distribuyen sobre la vialidad hasta llegar el momento en que, para cualquier par origen-destino, todas las posibles rutas tienen el mismo costo, en este caso, en relación al tiempo de traslado. Al alterar los tiempos de recorrido en las rutas, mediante la modificación de los parámetros con los que operan los semáforos en las vías señalizadas, se rompe el equilibrio, ya que los usuarios cambiarán sus rutas para mejorar sus tiempos haciendo que cambien los flujos en la red.

Tomando en cuenta la relación que existe entre el cambio en la programación de los semáforos y los flujos de tránsito en la red, la condición de equilibrio es considerada como una restricción del problema de optimización.

En ese trabajo, se plantea el problema de determinar los tiempos óptimos de señalización y su programación (coordinación, tiempo de verde y duración del ciclo) en la red semaforizada, se plantea considerando la respuesta de los usuarios a los cambios de señalización, con base en el modelo desarrollado por Teklu, Sumalee y Watling (2007) que utiliza como un programa de optimización matemática bi-nivel con restricciones de equilibrio.

Formulación del Problema

Variables de decisión. Se consideran como variables de decisión, los tiempos de señal restringidos a los flujos asociados, que son resultado del Equilibrio del Usuario el cual considera el tiempo de viaje como función de los tiempos de señal dados.

Para una intersección, las variables de diseño de señal son:

- tiempo de verde para diferentes fases.
- tiempo de ciclo,
- desfase,

Se llama **fase** al intervalo de tiempo en el cual uno o más de los movimientos que se realizan en la intersección tiene asignado su derecho a paso a través de ellas. Se acostumbra mostrar los movimientos que se efectúan en cada fase en un gráfico al que se le llama diagrama de fases (Rico, 1989).

En las figuras 3.1, 3.2 y 3.3 se muestran los diagramas de fase de una intersección con dos, tres y cuatro fases respectivamente la cuales son utilizadas comúnmente en intersecciones de cuatro ramas (Rico, 1989):

Figura 3.1. Diagrama de intersección a dos fases en intersecciones de cuatro ramas.

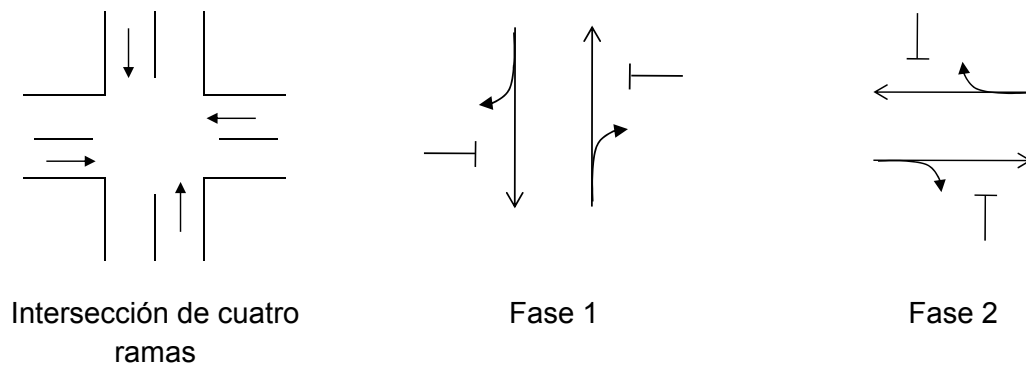


Figura 3.2. Diagrama de intersección a tres fases en intersecciones de cuatro ramas.

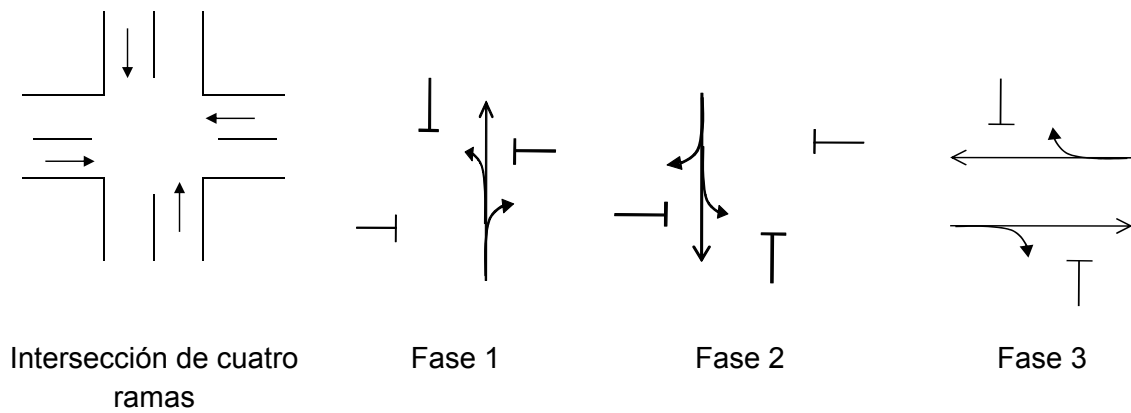
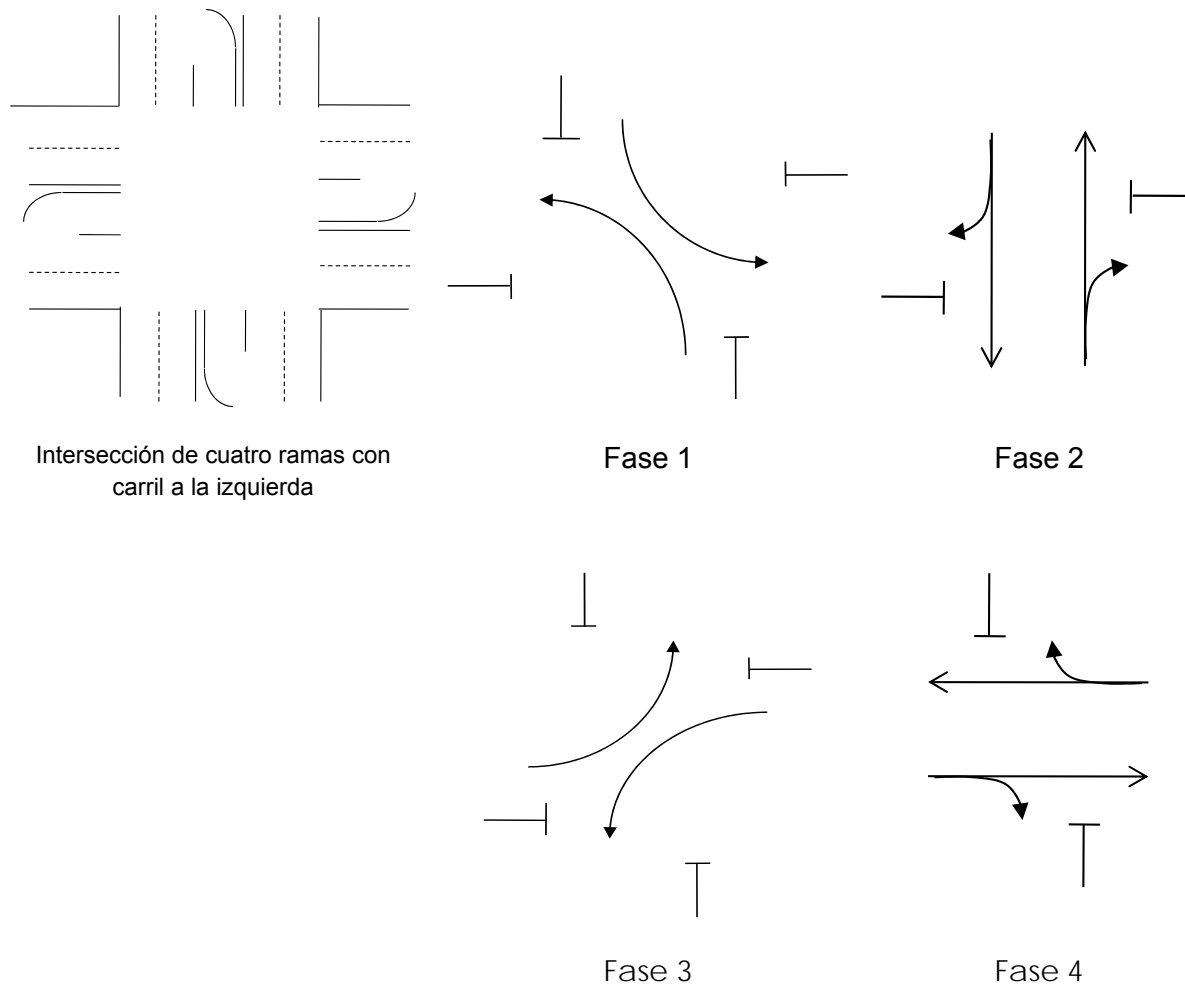


Figura 3.3. Diagrama de intersección a cuatro fases en intersecciones de cuatro ramas.



Se llama **ciclo** al intervalo de tiempo en el cual se efectúa la secuencia completa de fases.

La longitud del ciclo es un factor que afecta notablemente la operación del tránsito vehicular a través de una intersección. De manera general podemos afirmar que el ciclo deberá ser mayor en la medida en que aumentan los volúmenes de demanda.

Se debe tener especial cuidado y no utilizar ciclos extensos en casos de volúmenes bajos, ya que se fomenta la desobediencia del usuario ante su desesperación de tener impedido su derecho de paso cuando no observa vehículos que circulen en los otros movimientos. Esta desobediencia puede ocasionar accidentes (Rico, 1989).

Ciclos cortos aplicados a volúmenes grandes aumentan las demoras a los vehículos que esperan pasar por la intersección, debido a que los tiempos perdidos entre fases por arranque y despejes de los movimientos, los cuales son fijos para cualquier longitud del ciclo, representan una proporción de tiempo mayor entre menor sea el ciclo (Rico, 1989).

Ante lo delicado de esta situación es importante elegir un tiempo de ciclo adecuado a la demanda vehicular. Este problema planteado de manera analítica fue resuelto por Webster (1966) minimizando las demoras de los vehículos que esperan a pasar por la intersección, la cual se resume a continuación:

El ciclo óptimo se obtiene minimizando la demora total con respecto al tiempo de ciclo y a su división. Mediante una simulación del comportamiento del tránsito, Webster encuentra la siguiente expresión para la demora promedio por vehículo en el acceso de una intersección:

$$d = \frac{c(1-\lambda)^2}{2(1-\lambda x)} + \frac{x^2}{2q(1-x)}$$

en donde:

d = igual demora promedio por vehículo en el acceso particular de un intersección.

c = tiempo de ciclo.

λ = porción de ciclo que es verde efectivo, para la fase bajo consideración.

q = flujo o intensidad de tránsito en un acceso particular del intersección.

s = flujo de saturación (igual a 525 veces la anchura del acceso en metros) en vehículos ligeros por hora.

x = grado de saturación. Es la relación entre flujo real y el flujo máximo que puede pasar por un acceso, a través de la intersección; está dado por: $x = \frac{q}{\lambda s}$.

La demora total para cada acceso de la intersección, por unidad de tiempo, es el producto de la demora promedio por vehículo, por el flujo. Por consiguiente, la demora total para la intersección en conjunto está dada por:

$$D = \sum_{i=1}^n d(q_i)$$

A través del desarrollo matemático, resultado de diferenciar la ecuación de la demora total con respecto al tiempo de ciclo, Webster llega a la siguiente expresión para el ciclo óptimo:

$$C_o = \frac{KL + 5}{1 - Y}$$

En donde K varía entre 1.24 y 1.98 de acuerdo con el valor de la saturación de flujo promedio y la relación de flujo a flujo de saturación.

La expresión general para todos los casos es:

$$C_o = \frac{1.5L + 5}{1 - Y}$$

En donde:

C_o = tiempo óptimo de ciclo. Es el tiempo de ciclo que produce la menor demora a todos los vehículos utilizan la intersección.

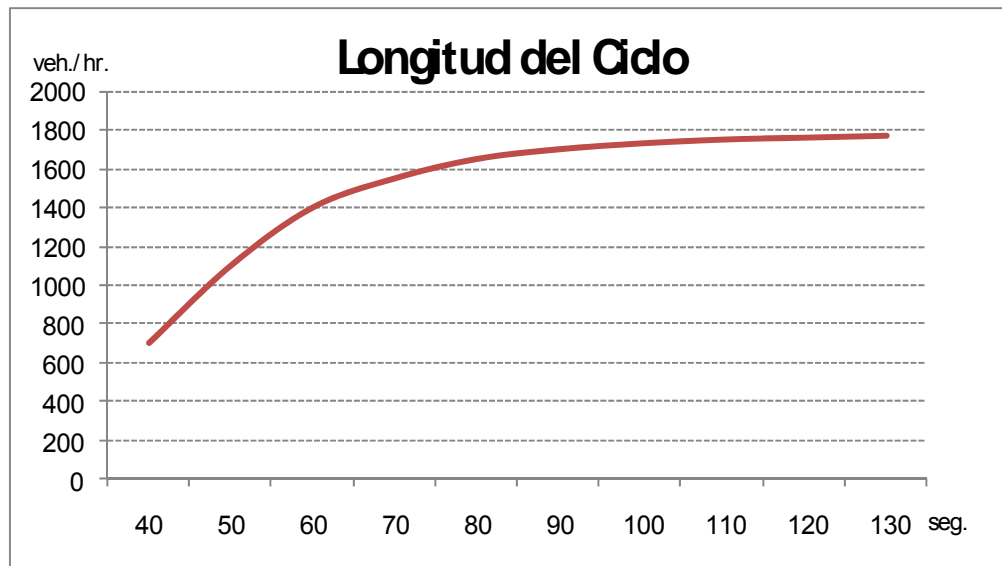
L = tiempo total perdido por ciclo. Es la suma de tiempos perdidos en cada fase incluyendo aquellos períodos cuando todas las luces indican rojo o ámbar-rojo.

Y = suma para toda la intersección, de los valores y correspondientes a cada fase siendo y la relación entre flujo real y el flujo de saturación (q/s) para una fase dada.

El rango de valores aceptados para el ciclo óptimo está dado por $40 \leq C \leq 120$ seg.

La longitud del ciclo también debe ser mayor a medida que existan más de dos fases (Rico, 1989), sin embargo es conveniente insistir en que la longitud del ciclo máxima utilizar es de 120 segundos, ya que se ha observado que ciclos mayores producen incrementos insignificantes a la capacidad de la intersección y en cambio incrementan aceleradamente la demora total, esto se muestra en la figura 3.4.

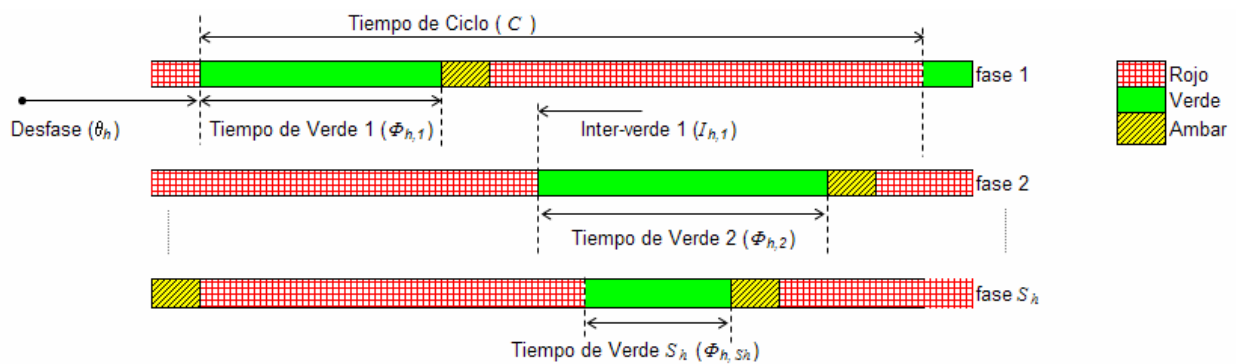
Figura 3.4. Relación entre capacidad y longitud de ciclo (Rico, 1989).



El **desfase** se define como la diferencia entre los comienzos de verde entre dos intersecciones sucesivas. En una intersección, el **período entre fases**, se define como el intervalo entre el final del tiempo de verde de una fase hasta el inicio del tiempo de verde de la fase inmediata posterior.

El diagrama de tiempos de ciclo es la forma gráfica de mostrar los tiempos de inicio y fin de cada fase y su situación con respecto a los tiempos de las otras fases. En el sentido horizontal se representa al tiempo, generalmente medido en segundos, y en el sentido vertical se colocan los números con los que se etiquetan los distintos grupos de semáforos para distinguir los entre sí.

Figura 3.5. Diagrama de Tiempos de Ciclo para S fases. (Teklu, et. al., 2007)



Función Objetivo. La función objetivo a minimizar en el primer nivel es el tiempo total de recorrido (TT) sobre la red elegida. TT es definido como la suma de los productos de los flujos y tiempos de viaje sobre el total de las rutas de toda la red. En las intersecciones, cada movimiento de vuelta es separado como un arco de longitud cero, que tiene un retraso considerado en TT . TT es influenciado por la programación de señal ψ y por el patrón del flujo $q^*(\psi)$ en la red.

En el segundo nivel, como una restricción, se encuentra la respuesta de usuarios a los cambios en los parámetros de diseño de red buscando optimizar su propia función de impedancia.

Planteamiento. Matemáticamente, el problema se plantea como:

$$\text{Min}_{\psi \in \Omega_0} TT(\psi, q^*(\psi)) = \sum_{a=1}^L q_a t_a(\psi, q^*(\psi)) \quad (1)$$

$$\text{sujeto a: } \psi(C, \theta, \phi) \in \Omega_0 \quad (2)$$

$$\text{esto es: } C_{\min} < C < C_{\max} \quad (3)$$

$$0 < \theta_h < C - 1 \quad (4)$$

$$\phi_{h,r \min} < \phi_{h,r} < \phi_{h,r \max} \quad (5)$$

$$C = \sum_{r=1}^{S_h} \phi_{h,r} + \sum_{r=1}^{S_h} I_{h,r} \quad \forall h \quad (6)$$

y sujeto a las restricciones del Equilibrio de Usuario, las que se presentan como:

$$\text{Min } z(x) = \sum_{i,j} \int_0^{x_{i,j}} t_{ij}(u) du \quad (7)$$

sujeto a:

$$D(j, s) + \sum_i x_{ij}^s = \sum_k x_{ik}^s \quad s = 1, 2, \dots, p \quad j = 1, 2, \dots, n \quad j \neq s \quad (8)$$

$$x_{ij}^s > 0, \quad s = 1, 2, \dots, p \quad (9)$$

En donde:

TT = Tiempo total de viaje en la red.

q_a = flujo en el arco a ($a = 1, 2, \dots, L$).

C = longitud común del ciclo.

L = Número de arcos en la red.

n = Número de nodos en la red.

p = Número de orígenes y destinos.

x_{ij} = Flujo total en el arco (i,j) .

x_{ij}^s = Flujo total en el arco (i,j) con destino s .

$D(j, s)$ = Flujo originado en el nodo j con destino s .

θ_h = Desfase³ de la intersección h , es un elemento del vector θ de desfases.

$\phi_{h,r}$ = Duración del tiempo de verde en la r -ésima fase de la intersección h , es un elemento del vector de tiempos de verde.

$\psi(C, \theta, \phi)$ = Vector de parámetros de la señalización (longitud del ciclo, desfase, longitud de fase).

$q^*(\psi)$ = Vector de flujos de equilibrio de usuario dado un conjunto de parámetros de señal ψ . En donde el vector de flujo del Equilibrio de Usuario $q^*(\psi)$, para una ψ dada, está dado por la desigualdad variacional (Smith, 1979):

$$t(\psi, q^*) \cdot (q - q^*) \geq 0 \quad \forall q \in Y \quad (10)$$

siendo t y q los vectores de funciones de tiempo de viaje y flujo del arco respectivamente y Y es el espacio factible del vector del flujo del arco.

³ Se considera como desfase a la diferencia entre el comienzo del verde entre dos intersecciones sucesivas.

$t_a[\psi, q^*(\psi)]$ = Tiempo de viaje en el arco a

S_h = Número total de fases en la intersección h .

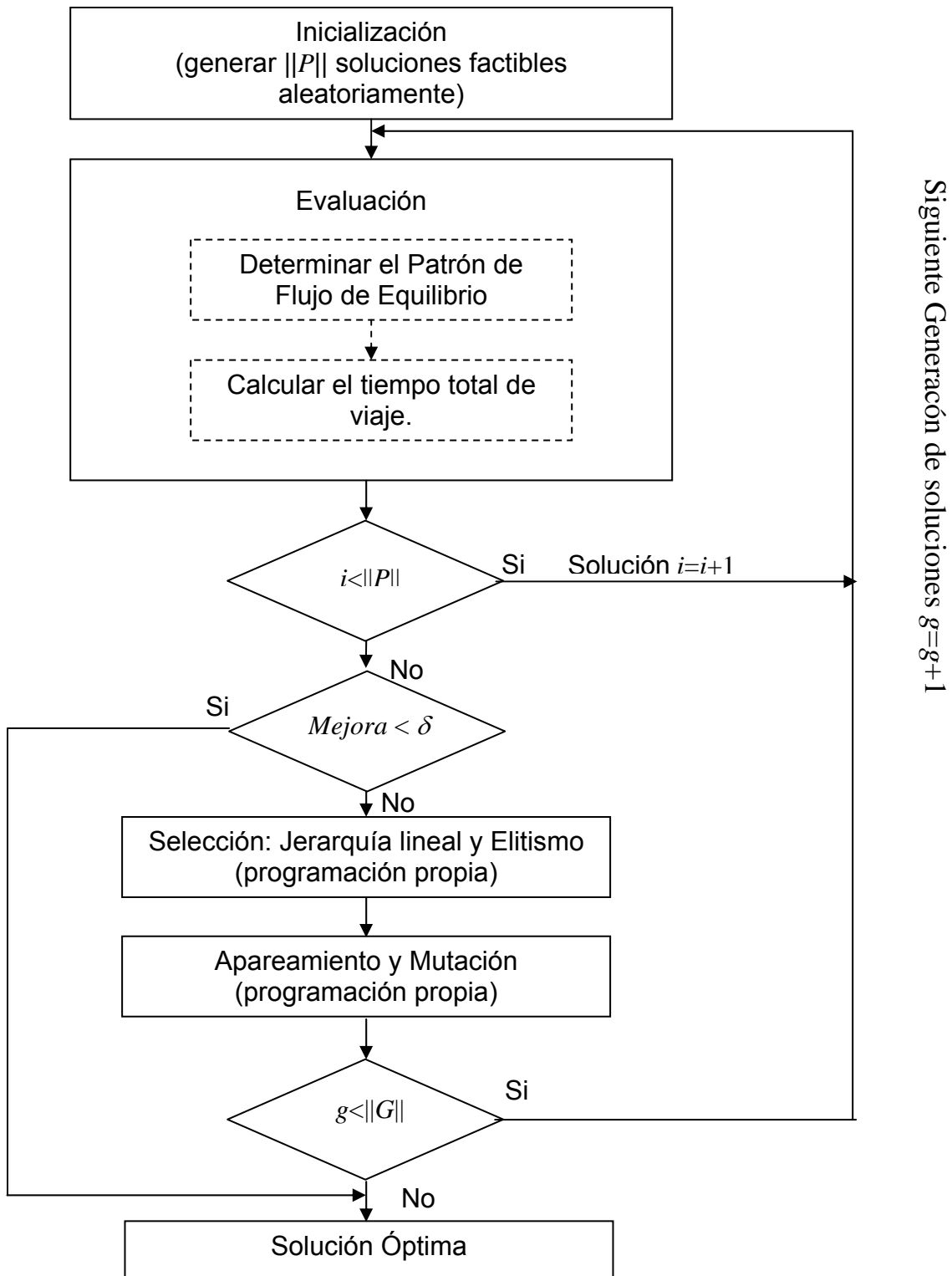
$I_{h,r}$ = Es el tiempo entre fases, para la intersección h , se define como el tiempo que transcurre entre el final de un tiempo de verde de la etapa r y el inicio del siguiente verde, el cual se considera igual a 4 segundos.

Proceso

Para resolver el problema, se utiliza el módulo de Asignación de Tránsito del software denominado TransCad aplicando el método de Equilibrio de Usuario (Wardrop). Resultado de este módulo, se obtiene el equilibrio del usuario y el tiempo total de viaje bajo equilibrio.

Los Algoritmos Genéticos generan un conjunto de parámetros de señal, inicialmente de manera aleatoria; cada uno de los elementos de ese conjunto, llamado generación, son tomados por el Módulo de Asignación de Tránsito el cual da los flujos de equilibrio y el tiempo total de viaje. El método de Algoritmos Genéticos, desarrollado en el módulo de programación de TransCad, identifica los elementos de la generación que tienen mejor comportamiento con respecto al tiempo total de viaje y con estos elementos, forma una nueva generación de parámetros de señal, repitiendo de manera iterativa el proceso hasta que la diferencia en el valor de la función objetivo entre dos iteraciones sucesivas sea menor a un determinado valor de convergencia. El proceso a seguir se resume en la Figura 3.6.

Figura 3.6. Diagrama de Flujo del Procedimiento de Solución.



En donde g indica el número de iteración, G el número máximo de iteraciones, P el tamaños de la generación y δ es el valor de convergencia.

Factibilidad de las restricciones.

Para el problema formulado, se explica la factibilidad técnica de las variables de diseño, la que se garantiza a partir de las ecuaciones (3)-(6).

Con el fin de mantener la coordinación a través de los ciclos, es necesario que cada intersección opere con un tiempo de ciclo común (IHT, 1997). El ciclo de tiempo común de la red, C , está restringido entre C_{min} y C_{max} , en donde C_{min} es determinado identificando el nodo que necesita mayor duración para acomodar el tiempo inter-verde y el mínimo número de tiempos de verde, esto es:

$$C_{min} = \text{Max} \left\{ \left(\sum_{r=1}^{S_h} \phi_{h,r \min} + \sum_{r=1}^{S_h} I_{h,r} \right) : h = 1, 2, \dots, N \right\} \quad (11)$$

C_{max} es restringida a 120 segundos. Para cualquier intersección h , el desfase puede variar únicamente entre cero y la duración de ciclo menos uno. En este caso, se considera 7 segundos como tiempo mínimo de verde, cifra considerada como adecuada en la práctica de ingeniería de tránsito. El tiempo máximo de verde en cada etapa $\phi_{h,r \max}$ se calcula suponiendo que todas las otras etapas en la intersección sólo necesitan el tiempo mínimo de

$$\text{verde: } \phi_{h,r \max} = C - \sum_{r=1}^{S_h} I_{h,r} - \sum_{\substack{y=1 \\ y \neq r}}^{S_h} \phi_{h,y \min}$$

En esta ecuación se asegura que los tiempos de verde en una intersección más los tiempos de desfase en ese ciclo igualan el ciclo total.

Metodología de Solución.

Para minimizar TT con respecto a los parámetros de las señales de tránsito, se plantea una metodología de optimización heurística basada en Algoritmos Genéticos, considerando que los flujos se mantienen en equilibrio. Así, cada vez que la función objetivo es evaluada en un determinado valor del vector de parámetros de la señal, se determina el valor del flujo de la red de tal manera que se establezca el equilibrio.

La idea básica es la siguiente: en cada iteración la función objetivo es evaluada en una elección particular de parámetros de señal, con éstos parámetros se determinan los flujos de equilibrio con el modulo de Asignación de Tránsito de TransCad. A pesar de que la idea básica es sencilla, existen puntos importantes en relación a la definición precisa de los parámetros y operadores que intervienen en el Algoritmo Genético a aplicar.

Como se ha dicho anteriormente, los Algoritmos Genéticos, AG, en general, están inspirados en la teoría de la evolución, en la que los individuos más fuertes o mejor adaptados son los que sobreviven y los más débiles o menos adaptados son desechados. En este contexto, Inicialmente, se genera una población inicial de cromosomas, en cada uno de los cuales hay soluciones potenciales. El algoritmo genético evalúa la función objetivo en cada cromosoma y, a través de un proceso de selección probabilístico, se eligen algunos cromosomas para formar una *población intermedia*. Imitando el proceso de la lucha evolución de los seres vivos, los cromosomas que mejor se comporten, conforme su resultado en la función objetivo, son los que más probabilidad tienen de ser elegidos para formar la siguiente generación. Los cromosomas de la población intermedia son aleatoriamente cruzados para intercambiar el material genético y producir una descendencia en el proceso de entrecruzamiento. Por último, en el proceso de mutación, los genes de algunos cromosomas seleccionados de manera aleatoria, se mutan y generan la siguiente población. El proceso de evolución, va mejorando su comportamiento en relación a la función objetivo a través de las generaciones subsecuentes.

Los AG crean y mejoran un conjunto de parámetros de control de señales, para las cuales, el modelo de Asignación de Tránsito determina el patrón de flujo en equilibrio y el tiempo total de viaje respectivo, que es utilizado para evaluar el conjunto de parámetros de control de señales, y determinar la siguiente generación de parámetros que será mejor que la anterior. Esto se hace de manera repetitiva hasta que la solución converge o bien hasta que un determinado número de iteraciones es alcanzado.

Diseño del Cromosoma.

En el problema de Diseño de Red, el número total de variables de decisión, ε , está dado por el tiempo de ciclo común C , N desfases (uno para cada intersección) y S_m tiempos de verde en cada intersección, dando un total de:

$$\varepsilon = 1 + N + \sum_{m=1}^N S_m$$

En el algoritmo genético, cada una de estas variables es representada por una cadena de 8 bits, cada bit es tomado del conjunto binario $\{0,1\}$. Por ejemplo, si se considera un red vial con sólo dos intersecciones señalizadas ($N=2$), cada una de las cuales tiene dos etapas ($S_m=2 ; m=1,2$), el número de cadenas, ε , es: $\varepsilon = 1 + 2 + \sum_{m=1}^2 2 = 7$. Estas cadenas son combinadas para formar un cromosoma representando un vector de parámetros de señal factible en la red. En el ejemplo anterior, el cromosoma correspondiente, las cadenas, y las variables representadas se dan de la siguiente manera:

Ξ_1	Ξ_2	Ξ_3	Ξ_4	Ξ_5	Ξ_6	Ξ_7
00101010	10101010	00101011	11010100	00010100	00100111	11111110
C	θ_1	θ_2	$\phi_{1,1}$	$\phi_{1,2}$	$\phi_{2,1}$	$\phi_{2,2}$

La forma general de dicho cromosoma, se muestra a continuación para una red vial de N intersecciones señalizadas.

$$|C| \theta_1, \theta_2, \dots, \theta_N | \phi_{1,1}, \phi_{1,2} \dots \phi_{1,S_1} | \dots | \phi_{N,1}, \phi_{N,2} \dots \phi_{N,S_N} |$$

Codificación de Cromosomas y representación de restricciones

El desciframiento de los cromosomas está basado en Ceylan y Bell (2004)

- a) Duración de ciclo (C). Es la proporción de la diferencia $C_{max} - C_{min}$ más C_{min} :

$$C = C_{min} + \frac{\zeta[\Xi_m]}{2^8 - 1} (C_{max} - C_{min}) \quad m = 1$$

donde $\zeta[\Xi_m]$ es el equivalente en base 10 de Ξ_m ; $C_{max}=120$ segundos y C_{min} es calculado con la ecuación (8).

- b) Desfase (θ_h): Para una etapa de la intersección h , es la proporción de la longitud de ciclo definida por la m -ésima cadena del cromosoma:

$$\theta_h = \frac{\zeta[\Xi_m]}{2^8 - 1} (C - 1) \quad m = 2, 3, \dots, N + 1$$

donde $h=m-1$

- c) Tiempos de verde ($\phi_{h,r}$): Están definidos como la suma de la longitud mínima de fase y la proporción del tiempo restante de tiempo de verde $\phi_{h,r,max} - \phi_{h,r,min}$ como sigue:

$$\phi_{h,r} = \phi_{h,r,min} + (\phi_{h,r,max} - \phi_{h,r,min}) \frac{\zeta[\Xi_m]}{\sum_{m=x+1}^{x+S_h} \zeta[\Xi_m]} \quad m = N + 1, N + 2, \dots, \mathcal{E}$$

donde $x = N + 1 + \sum_{y=1}^{h-1} S_y$ es constante en la intersección h , $r=m - x$ es una fase en la intersección h y S_h es el número total de fases en la intersección h , $h=1, 2, \dots, N..$

⁴ Para un número B en base 2, de 8 dígitos, $(b_8 b_7 b_6 b_5 b_4 b_3 b_2 b_1)_2$ donde $b_k \in \{0, 1\}$, el equivalente en base 10 esta dado por

$$\zeta B = \sum_{k=1}^8 b_k 2^{k-1}$$

Inicialización.

La población inicial de cromosomas, de tamaño $||P||$, es aleatoriamente generada, con base tanto en la información específica de la red como en el número de intersecciones señalizadas N , y el número de fases en cada intersección (S_h).

Evaluación-Comportamiento.

Cada cromosoma es decodificado y enviado al modelo de asignación para obtener el correspondiente flujo bajo equilibrio, y el tiempo total de viaje asociado, TT . El software de optimización heurística desarrollado utiliza posteriormente a TT como la función de comportamiento para el proceso de selección.

Selección.

En la etapa de selección, el programa de optimización heurística lleva al cabo un muestreo estocástico con reemplazo (Goldberg, 1989) mediante una rueda de ruleta que usa una probabilidad p_k , basada en el resultado de evaluar el k -ésimo cromosoma en la función objetivo, esto es:

$$p_k = \frac{TT(\psi_k, q^*(\psi_k))}{\sum_{k=1}^P TT(\psi_k, q^*(\psi_k))}$$

Con base en la probabilidad definida se determinan los cromosomas que pasan a la siguiente generación sin ninguna alteración, conocida como generación élite, y que formarán, a partir del proceso de entrecruzamiento una nueva generación.

Entrecruzamiento.

Durante esta etapa, los cromosomas élite son presentados en el conjunto de descendencia para compartir sus mejores genes. Las selecciones son aleatoriamente cruzadas utilizando entrecruzamiento uniforme (Syswerda, 1989). Para realizar el entrecruzamiento, una máscara de cromosoma con una secuencia aleatoria de 0's y 1's se genera para cada par. El valor de la máscara del cromosoma determina cuál cromosoma del padre suple el bit en una posición determinada en el cromosoma de sus dos ascendientes. Los cromosomas no seleccionados pasan directamente al proceso de mutación.

Mutación

La probabilidad de mutación, P_m , determina la posibilidad de que ocurra una mutación en un determinado cromosoma. Para éstos, la mutación se genera seleccionando un punto (bit) de manera aleatoria en la longitud del cromosoma descendiente y cambiando el valor de ese bit, 0 si era 1 o viceversa.

Pseudocódigo del Procedimiento

Definición de variables

** *Variables de la red vial.*

- N (integer): número de intersecciones.
- Sm (integer): número de etapas en la intersección m-ésima.

** *Parámetros de AG's.*

- P (integer): tamaño de la población para los AG's.
- $Elite$ (integer): número de cromosomas que forman la siguiente generación sin alterarse.
- G (integer): Número máximo de iteraciones.
- Pm (real): Probabilidad de mutación de un cromosoma.
- δ (real): Criterio de convergencia.

** *Variables internas de AG's*

- Vect_Sm[N] (integer): Guarda el número de etapas de cada intersección.
- Mat_Aux1 [$P,7$]: Cuyas entradas son matrices o números en donde se manejan, para una generación, los cromosomas, sus comportamientos, y otros datos necesarios para el proceso.

Mat_φ [N,Max](real) :Matriz de tiempos de verde de cada etapa de cada intersección a transferir a TransCad.

Vect_θ [N](real):Vector de desfases a transferir a TransCad.

C(real): Duración del ciclo a transferir a TransCad.

Ξ [ε,8] (bool): Objeto tipo matriz que representa un cromosoma.

Padre[ε,8] (bool): Objeto tipo matriz que representa a uno de los cromosomas padres en el proceso de entrecruzamiento.

Madre[ε,8] (bool): Objeto tipo matriz que representa a uno de los cromosomas padres en el proceso de entrecruzamiento.

H1[ε,8] (bool): Objeto tipo matriz que representa a uno de los cromosomas hijo en el proceso de entrecruzamiento.

H2 [ε,8] (bool): Objeto tipo matriz que representa a uno de los cromosomas hijo en el proceso de entrecruzamiento.

** Variables auxiliares y contadores.

ε (integer): Tamaño del cromosoma.

g (integer): Contador .

m (integer): Contador .

i (integer): Contador.

j(integer): Contador.

k(integer): Contador.

f(integer): Contador.

r(integer): Contador.

w(integer): Contador.

z(integer): Contador.

x(integer): Variable auxiliar para determina la posición del cromosma que marca el inicio de cada desfase.

φmax(integer): Duración máxima de verde.

Max (integer): Máximo de etapas en las intersecciones de la red.

Cmin(real): Tiempo común mínimo.

Conv(bool): Variable auxiliar que indica si existe convergencia.

Aux (real): Variable auxiliar ocupada en diferentes etapas del proceso.

Aux1(real): Variable auxiliar ocupada en diferentes etapas del proceso.

Aux2(real): Variable auxiliar ocupada en diferentes etapas del proceso.

Sum (integer) : Variable auxiliar.

Mut(integer): Identificado el bit a mutar en un gen del cromosoma

Vect_Aux1 [(P-Elite)/2,2] (integer): Vector que indica el renglón de la matriz de los cromosomas a formar una nueva generación.

Vect_Aux2 [P]: Vector cuyas entradas son objetos tipo matrices en donde se maneja de forma temporal la nueva generación.

Vect_TT[P]: Vector en donde se ordenan los tiempos de traslado totales de cada cromosoma.

Mat_Opt[G,7](real): Matriz de los resultados óptimos de cada generación.

*****Funciones y procedimientos requeridos en el proceso*****

**** Función *numero_aleatorio* () que generadora aleatorios.**

(Programa que genera números aleatorios entre 0 y 1 con resultado en variable que se llama *numero_aleatorio(real)*).

numero_aleatorio=rand()

**** Función $\zeta(\Xi_m)$ que convierte un gen del cromosoma a base 10**

(Recibe el número de renglón del cromosoma Ξ [i ,j] en donde se encuentra el gen y da como resultado el equivalente en base 10)

Sum = 0

For j=1 to 8

Sum=Sum+ Ξ [k ,9-j]*2^(j-1)

Next j

$\zeta(B)$ =Sum

**** Procedimiento *Generador_de_Cromosmas***

(Genera de manera aleatoria cromosomas y los almacena en un vector).

For i=1 to ε

For j= 1 to 8

k=round(numero_aleatorio(),0)

Ξ [i ,j] =k

Next j

Next i

**** Función *Descifra_C***

(Calcula la duración del ciclo con base en el primer gen de cromosoma y da como resultado un real).

Descifra_C=Cmin + $\zeta(1)/(2^8-1)*(120-Cmin)$

**** Procedimiento *Descifra_Desfases***

(Descifra los desfases de cada intersección en un ciclo y los pone en el vector *Vect_θ*).

For i=1 to *N*

Vect_θ[i]=(C-1)* $\zeta(i+1)/(2^8-1)$

Next i

**** Procedimiento *Descifra_Verdes***

(Descifra la duración del tiempo de verde en cada etapa de cada intersección y las pone en el vector *Vect_φ*).

**** Determina la duración del verde máxima de cada etapa en cada intersección**

For i=1 to *N*

ϕ_{max} =C-11**Vect_Sm*[i]+7

**** Determina el rango del vector de cromosomas en donde se encuentran las φ's de la intersección que se maneja**

If i=1 then

x=*N*+1

```

Else
  Sum = 0
  For k=1 to i-1
    Sum=Sum+Vect_Sm[k]
  Next k
  x=N+1+sum
End if
** Determina la suma de los genes  $\phi$ 's de la intersección i
Sum=0
For k= x+1 to (x+Vect_Sm[i])
  Sum=Sum+ $\zeta$ [k]
Next k
For j=N+1 to  $\varepsilon$ 
  r=j-x
  Vect_ $\phi$  [i,j]=7+( $\phi$ max-7)* $\zeta$ (j)/Sum
Next j
Next i

```

**** Función *Tiempo_Total*** (C, Vect_ θ ,Vect_ ϕ)
 (Envía a TransCad los parámetros (C, Vect_ θ ,Vect_ ϕ) y arroja como resultado el tiempo total de viaje(real)).

**** Función *Flujo_Equilibrio*** (C, Vect_ θ ,Vect_ ϕ)
 (Envía a TransCad los parámetros (C, Vect_ θ ,Vect_ ϕ) y recibe el Flujo de Equilibrio(vector)).

**** Función *Rango*** (Mat_Aux1[i,4])
 (Recibe un TT y determina su rango en función a los demás TT de la generación)

```

For w=1 to P-1
  Mayor=Mat_Aux1[w,4];
  Posmayor=i;
  For z=w+1 to P
    If Mat_Aux1[z,4] < Mayor then
      Mayor=Mat_Aux1[z,4];
      Posmayor=z;
    End If
  Next z
  Vect_TT[w]=Mat_Aux1[Posmayor,4];
  Vect_TT[Posmayor]= Mat_Aux1[w,4];
Next w
For w=1 to P
  For z=1 to P
    If Vect_TT[w]=Mat_aux1[z,4] then
      Mat_aux1[z,7]=w
      z=P
    End If
  Next z
Next w

```

*****Programa Principal*****

**** Almacena el número de fases de cada una de las intersecciones y calcula el tamaño del cromosoma a partir de las Variables de Red y de los Parámetros de AG's.**

Sum=0

For i=1 to N

 Sum=Sum+Sm

 Vect_Sm[i]=Sm

Next i

$\varepsilon=1+N+Sum$

**** Determina Max**

Max=Vect_Sn[1]

For i=1 to N

 If Vect_Sn[i]>Max then

 Max=Vect_Sn[i]

 End If

Next i

**** Calcula Cmin**

Cmin=11*Max

**** Calcula la primera generación de cromosomas, los descifra**

For i=1 to P

Generador_de_cromosomas

Next i

**** Proceso iterativo para cada generación.**

For g=1 to G

 For i=1 to P

 If g=1 then

Generador_de_cromosomas

 End If

 C=Descifra_C

Descifra_Desfases

Descifra_Verdes

**** TransCad lee los parámetros C, Vect_θ, Vect_φ y calcula TT y FE con la matriz original de flujos sobre la red reemplazando ésta última con la nueva matriz de Flujos de Equilibrio.**

**** Guarda cada cromosoma de la generación en la matriz auxiliar**

 Mat_aux1[i,1]=i

 Mat_aux1[i,2]= Ξ

 Mat_aux1[i,3]=FE

```
        Mat_aux1[i,4]=TT
Next i
```

**** Asigna probabilidad y la probabilidad acumulada de selección a cada cromosoma**

```
Sum=0
For i=1 to P
    Sum=Sum+Mat_Aux1[i,4]
Next i
For i=1 to P
    Pr=Mat_Aux1[i,4]/Sum
    Mat_Aux1[i,5]=Pr
    If i=i then
        Mat_Aux1[i,6]=Pr
    Else
        Mat_Aux1[i,6]= Mat_Aux1[i-1,6]+Pr
    End If
Next i
```

**** Determina el rango cada cromosoma de acuerdo a su comportamiento**
Rango

**** Guarda en un vector el renglón de la matriz Mat_Aux1 correspondiente al óptimo.**

```
For i=1 to ε
    If Mat_Aux1[i,7]=1 then
        For j=1 to 7
            Mat_Opt[g,j]=Mat_Aux1[i,j]
        Next j
    End If
Next i
```

**** Identifica Criterio de Convergencia.**

```
If g>1 then
    If Mat_Opt[g-1,4]- Mat_Opt[g,4]<δ
        Conv=1
        g=G
    Else
        Conv=0
    End If
Else
    Conv=0
End If
```

**** Se reproduce mientras no exista convergencia.**

```
If Conv=0 then
```

**** Determina población intermedia.**

(Se generan (*P-Elite*) números aleatorios entre el 0 y 1 y para cada número aleatorio se determina el intervalo de Probabilidad Acumulada en el que se encuentra)

**** Escoge al padre.**

For i=1 to (*P-Elite*)/2

Aux=numero_aleatorio();

**** Selecciona el cromosoma correspondiente al padre en función al intervalo de probabilidad acumulada.**

k=2

If Aux<Mat_Aux1[k,6] then

r=1

Else

If Aux>Mat_Aux1[k-1,6] and Aux<=Mat_Aux1[k,6] then

r=k

Else

k=k+1

End if

End if

Vect_Aux1[i,1]=r

Next i

**** Escoge a la madre.**

For i=1 to (*P-Elite*)/2

Aux=numero_aleatorio();

**** Selecciona el cromosoma correspondiente a la madre en función al intervalo de probabilidad acumulada.**

k=2

If Aux<Mat_Aux1[k,6] then

r=1

Else

If Aux>Mat_Aux1[k-1,6] and Aux<=Mat_Aux1[k,6] then

r=k

Else

k=k+1

End if

End if

Vect_Aux1[i,1]=r

Next i

**** Genera a los hijos a partir de los cromosomas elegidos.**

**** Determina probabilísticamente los genes a cruzar.**

For j=1 to (*P-Elite*)/2

For i=1 to *P*

**** Identifica a los padres para cruzarlos.**

If Vect_Aux1[j,1]=Mat_aux1[i,1]

Padre=Mat_Aux1[i,2]

Else

```

        If Vect_Aux1[j,2]=Mat_Aux1[i,2]
            Madre=Mat_aux1[i,2]
        End If
    End if
Next i

```

**** Genera dos descendientes probabilísticamente a partir de los padres y los muta.**

```

For k=1 to ε
    For f = 1 to 8
        Aux1=numero_aleatorio()
        Aux2=numero_aleatorio()
        If Aux1<Aux2 then
            H1[k,f]=Padre[k,f]
            H2[k,f]=Madre[k,f]
        Else
            H1[k,f]=Madre[k,f]
            H2[k,f]=Padre[k,f]
        End If
    Next f
    ** Genera el proceso de mutación de H1 y H2
    Aux1=numero_aleatorio()
    If Aux1<Pm then
        Aux2=numero_aleatorio()
        Mut= round(Aux2*10,0)-round(round(Aux2*10)/8,0)*8+1
        If H1[k,Mut]=0 then
            H1[k,mut]=1
        Else
            H1[k,mut]=0
        End if
    End if
    Aux1=numero_aleatorio()
    If Aux1<Pm then
        Aux2=numero_aleatorio()
        Mut= round(Aux2*10,0)-round(round(Aux2*10)/8,0)*8+1
        If H1[k,Mut]=0 then
            H2[k,mut]=1
        Else
            H2[k,mut]=0
        End if
    End if
End if
Next k

```

**** Los pone en un vector auxiliar en donde se forma la nueva generación.**

```

Vect_Aux2[j]=H1
Vect_Aux2[(P-Elite)/2+j]=H2
Next j

```

```

** Integra a la generación los cromosomas que pasan por élite.
  For k=1 to Elite
    For i=1 to P
      If Mat_Aux1[i,7]=k
        Vect_Aux2[P-Elite+k]=Mat_Aux1[i,7]
      End If
    Next i
  Next k

** Transfiere la nueva generación a la matriz Mat_Aux1.
  For j=1 to P
    Mat_aux1[2,j]=Vect_aux2[j]
  Next j
End If
Next g

```


IV. RESULTADOS Y DISCUSIÓN

Características de la red vial.

Los vectores de información de la subred de Querétaro en la que se aplica el modelo constan de los siguientes campos:

length (longitud en kilómetros)

dir (sentidos de la vía): 0 si es de doble sentido, -1 o 1 de un solo sentido;

lclas (la clasificación de las vialidades):

- 1 vías rápidas,
- 2 vías colectoras,
- 3 vías secundarias,
- 4 vías locales,
- 5 vías regionales, (carreteras que salen o entran a la zona conurbada y ya no llevan tránsito local);

ncarr: número de carriles;

nombre: nombre de la vialidad.

Se considera, como caso práctico, la aplicación del modelo en la red vial formada por el área acotada por Av. Constituyentes Oriente, Boulevard Bernardo Quintana, Autopista México-Querétaro y Av. Luis Pasteur. Los semáforos a optimizar son las que se encuentran en el tramo de Avenida Constituyentes Oriente entre Blvd. Bernardo Quintana y su intersección con Av. Luis Pasteur. En la figura 4.1 a) y b) se presenta el mapa de la red vial de Querétaro y de la subred considerada.

Figura 4.1a) Mapa de la Red Vial de Querétaro

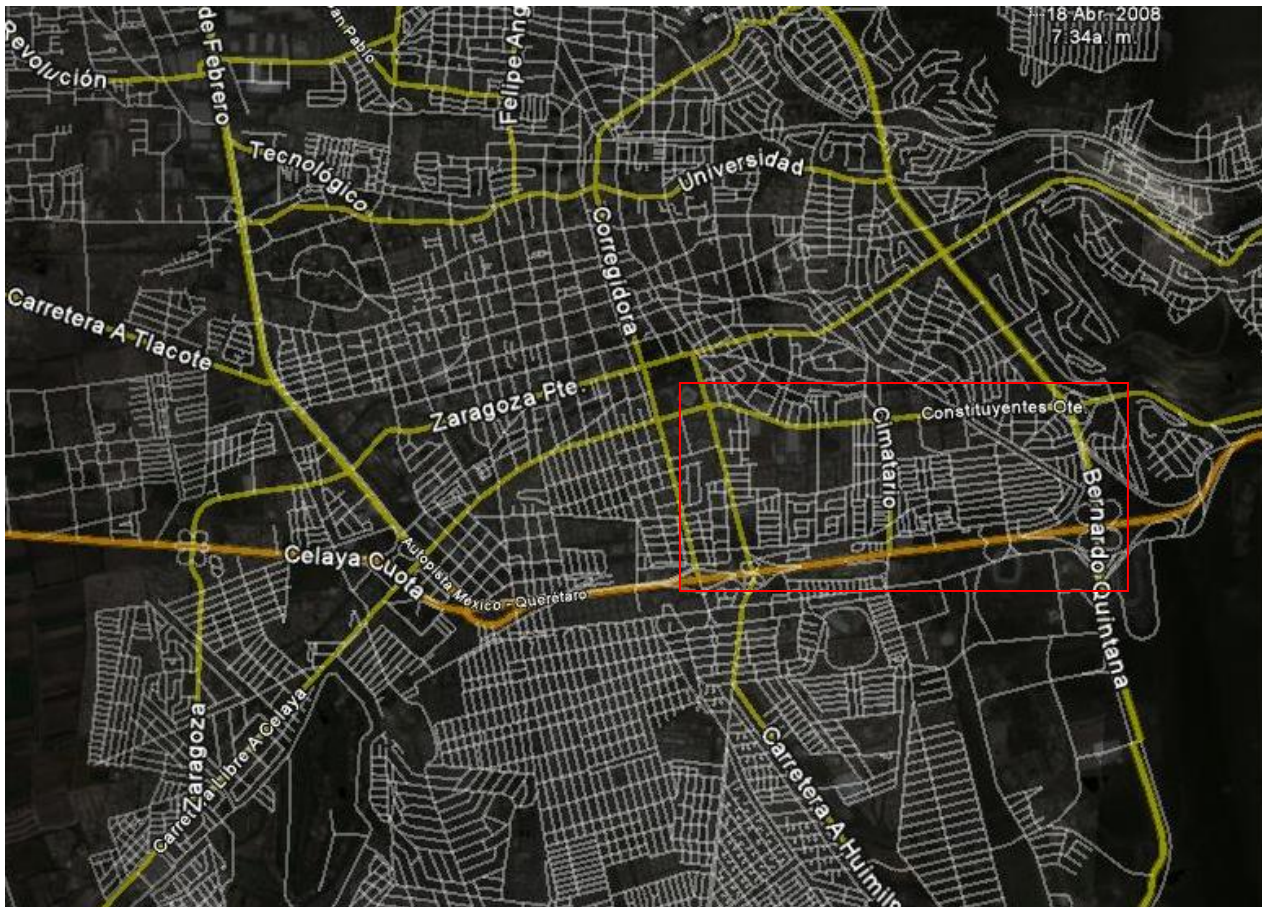
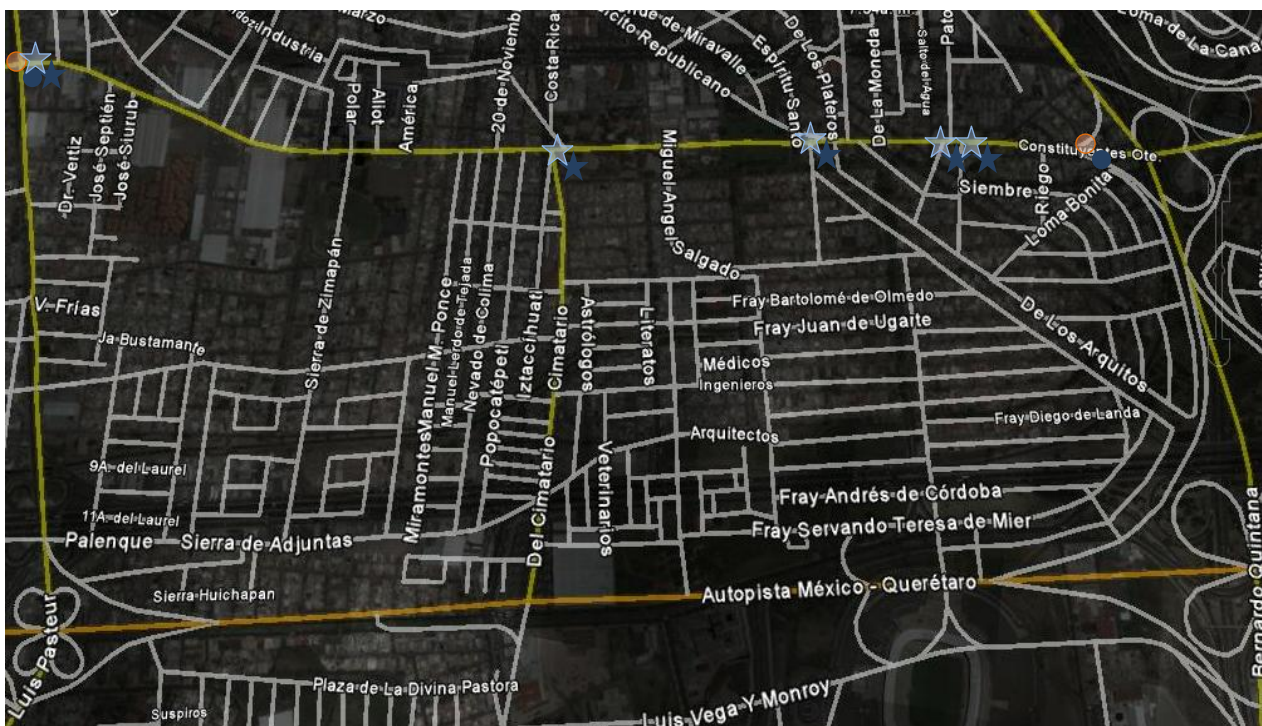


Figura 4.1 b) Subred considerada.



En la figura 4.1 b) se muestran con botones en forma de estrella los semáforos cuya operación es manipulada de tal manera de minimizar el tiempo total de recorrido desde la intersección de Boulevard Bernardo Quintana hasta la intersección de Av. Constituyentes con Av. Luis Pasteur, las que tienen las siguientes características:

Semáforo en la intersección Av. Constituyentes con calle de Los Plateros (figura 4.2), el cual consta de cinco fases:

Fase 1: con verde sobre Av. Constituyentes en ambas direcciones y rojo sobre Calle de Los Plateros en ambas direcciones.

Fase 2: con verde sobre Av. Constituyentes dirección Blvd. Bernardo Quintana- Av. Luis Pasteur, permitiendo vuelta a la izquierda a Calle de los Plateros y vuelta en “U”, y rojo sobre Av. Constituyentes en dirección Av. Luis Pasteur - Blvd. Bernardo Quintana y sobre calle de Los Plateros.

Fase 3: con verde sobre calle de Los Plateros en ambas direcciones y roja sobre Av. Constituyentes en ambas direcciones.

Fase 4: con verde sobre calle de Los Plateros en dirección Autopista México-Querétaro – Av. Constituyentes permitiendo vuelta hacia la izquierda y rojo en calle de los Plateros dirección Av. Constituyentes – Autopista México- Querétaro y sobre Av. Constituyentes en ambas direcciones.

Fase 5: con verde sobre calle de Los plateros en dirección Av. Constituyentes - Autopista México-Querétaro permitiendo vuelta a la izquierda y rojo en calle de los Plateros dirección Autopista México-Querétaro - Av. Constituyentes y sobre Av. Constituyentes en ambas direcciones.

Semáforo en la intersección Av. Constituyentes con calle Ejército Republicano (figura 4.2), el cual consta de dos fases:

Fase 1: con verde sobre Av. Constituyentes en ambas direcciones.

Fase 2: con verde sobre Av. Constituyentes dirección Av. Luis Pasteur - Blvd. Bernardo Quintana, permitiendo vuelta a la izquierda a calle Ejercito Republicano y

vuelta en “U”, y rojo sobre Av. Constituyentes en dirección Blvd. Bernardo Quintana- Av. Luis Pasteur.

Figura 4.2. Mapa de Intersecciones Av. Constituyentes con calle de Los Plateros y con calle Ejército Republicano.



(fuente: Google Earth)

Semáforo en la intersección Av. Constituyentes con Av. Cimatario (figura 4.3), el cual consta de tres fases:

Fase 1: con verde sobre Av. Constituyentes en ambas direcciones y rojo sobre Av. Cimatario en ambas direcciones.

Fase 2: con verde sobre Av. Constituyentes dirección Blvd. Bernardo Quintana- Av. Luis Pasteur, permitiendo vuelta a la izquierda sobre Av. Del Cimatario y vuelta en “U”, y rojo sobre Av. Constituyentes en dirección Av. Luis Pasteur - Blvd. Bernardo Quintana y sobre Av. Del Cimatario.

Fase 3: con verde sobre Av. Cimatario en ambas direcciones.

Semáforo en la intersección Av. Constituyentes con calle 20 de Noviembre al norte y calle Nevado de Toluca al Sur (figura 4.3), el cual consta de cuatro fases:

Fase 1: con verde sobre Av. Constituyentes en ambas direcciones y rojo sobre las calles de 20 de Noviembre y Nevado de Toluca.

Fase 2: con verde sobre Av. Constituyentes dirección Av. Luis Pasteur - Blvd. Bernardo Quintana, permitiendo vuelta a la izquierda a calle 20 de Noviembre y vuelta en "U", y rojo sobre Av. Constituyentes en dirección Blvd. Bernardo Quintana - Av. Luis Pasteur y sobre las calles de 20 de Noviembre y Nevado de Toluca.

Fase 3: con verde sobre calle de 20 de Noviembre y roja sobre Av. Constituyentes en ambas direcciones y sobre la calle Nevado de Toluca.

Fase 4: con verde sobre calle de Nevado de Toluca y roja sobre Av. Constituyentes en ambas direcciones y sobre la calle 20 de Noviembre.

Figura 4.3. Mapa de Intersecciones Av. Constituyentes con Av. Cimatario y con calle 20 de Noviembre al norte y calle Nevado de Toluca al Sur.



(Fuente: Google Earth)

Semáforo en la intersección Av. Constituyentes con Av. Luis Pasteur (figura 4.4), el cual consta de cuatro fases:

Fase 1: con verde sobre Av. Constituyentes en ambas direcciones y rojo Av. Luis Pasteur en ambas direcciones.

Fase 2: con verde sobre Av. Constituyentes dirección Blvd. Bernardo Quintana- Av. Luis Pasteur, permitiendo vuelta a la izquierda sobre Av. Luis Pasteur y vuelta en “U”, y rojo sobre Av. Constituyentes en dirección Av. Luis Pasteur - Blvd. Bernardo Quintana y sobre Av. Luis Pasteur en ambos sentidos.

Fase 3: con verde sobre Av. Constituyentes dirección Av. Luis Pasteur - Blvd. Bernardo Quintana, permitiendo vuelta a la izquierda sobre Av. Luis Pasteur y vuelta en “U”, y rojo sobre Av. Constituyentes en dirección Blvd. Bernardo Quintana - Av. Luis Pasteur y sobre Av. Luis Pasteur en ambos sentidos.

Fase 4: con verde sobre Av. Luis Pasteur en ambos sentidos y roja sobre Av. Constituyentes en ambas direcciones.

Figura 4.4. Mapa de Intersecciones Av. Constituyentes con Av. Luis Pasteur.



(Fuente: Google Earth).

En la figura 4.1 b) se muestran con botones redondos los puntos que se consideran como puntos de origen y destino. El flujo de demanda considerado es de 12,000 automóviles por hora de la intersección de Av. Constituyentes con Av. Luis Pasteur hasta Av. Constituyentes intersección con Boulevard Bernardo Quintana y viceversa.

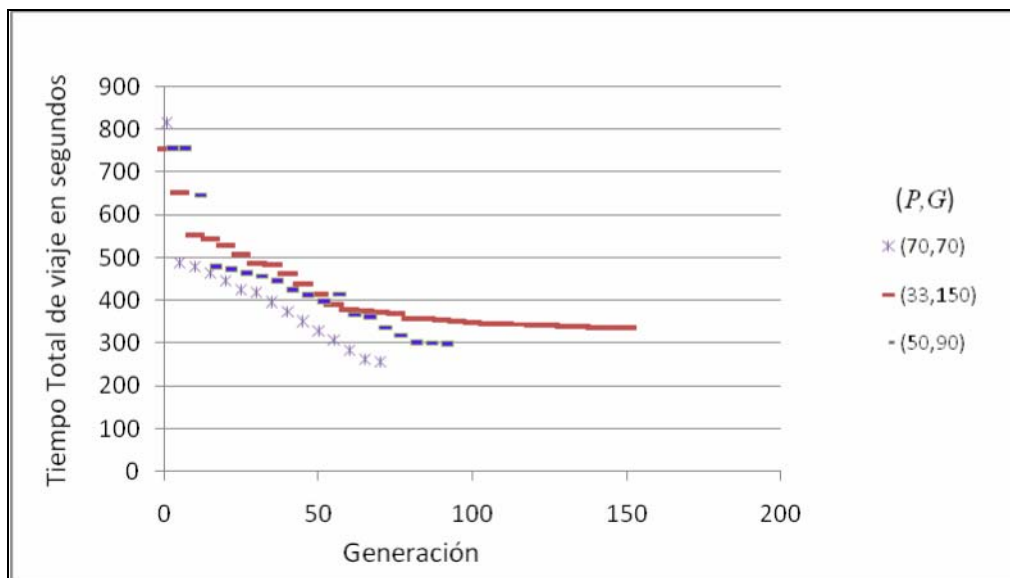
Resultados.

De acuerdo con Withley (1989) la diversidad de población y el sesgo de selección son dos importantes parámetros que influyen en la búsqueda genética. Conforme el sesgo de selección se incrementa – aumentando el parámetro *élite* en el proceso de selección de parámetros- la búsqueda se enfoca en los individuos mejores explotando sus mejores comportamientos, lo cual permite una convergencia rápida, sin embargo puede resultar en una convergencia prematura. Con un parámetro de selección bajo, por otro lado, el proceso se enfoca en la diversidad y el comportamiento de la búsqueda exploratoria. En el programa desarrollado los resultados pueden verse alterados por altos valores del tamaño de población P , y de la probabilidad de mutación P_m . Asimismo, un valor alto en el número de iteraciones máximo G puede incrementar el número de oportunidades para que el algoritmo genético encuentre la solución. Muchos investigadores en el campo de la optimización evolucionaria han tratado de investigar los efectos de los parámetros de los algoritmos genéticos de tal manera de poder definir los valores óptimos de los parámetros (Goldberg, 2002). Desafortunadamente, los resultados más avanzados en los ajustes óptimos de los parámetros de los algoritmos genéticos están limitados a problemas muy simples.

En ese trabajo se considera el método de Algoritmos Genéticos analizando el comportamiento de la evolución de la mejor solución y de la diversidad poblacional. Las figuras 4.5 y 4.6 presentan una comparación de los diferentes valores de población y número de generación en el proceso de búsqueda. El número total de cromosomas evaluados ha sido fijado alrededor de 5,000 de tal manera de identificar si es mejor determinar inicialmente el tamaño de población P , y posteriormente comprometer el número de generaciones G , o viceversa. Los otros parámetros fueron fijados de la siguiente manera: $P_m = 0.15$ y $élite = 2$. En la figura 4.5 se presenta, en el eje vertical, el tiempo total de viaje por automóvil resultado en la mejor solución de cada generación, contra el número de generación de dónde proviene

esta solución en el eje horizontal. Esto indica la velocidad de convergencia. Como se puede observar en esta figura, cuando el tamaño de la población P es menor, la velocidad de convergencia disminuye, mientras que un número grande de generaciones, representadas por el parámetro G , puede favorecer la salida del proceso de óptimos locales continuando la exploración; este es el caso de la gráfica correspondiente a $(P,G) = (33, 150)$ en donde se había encontrado, alrededor de la generación cuarenta, un óptimo local; por medio de los procesos de mutación se diversificó la búsqueda mejorando el resultado a partir de la generación 60. Con base en la gráfica presentada en esta figura, se concluye que para los parámetros $(P,G) = (33, 150)$ y $(P,G) = (50, 90)$, los tiempos mínimos de recorrido representan óptimos locales, esto debido a que no existe un tamaño de población adecuado, lo cual trae consigo una falta de diversidad genética, dificultando la exploración diversificada. Al aumentar el tamaño poblacional ($P=70$), la solución óptima tiene una mejora de 77.94 segundos y de 40.62 segundos respectivamente.

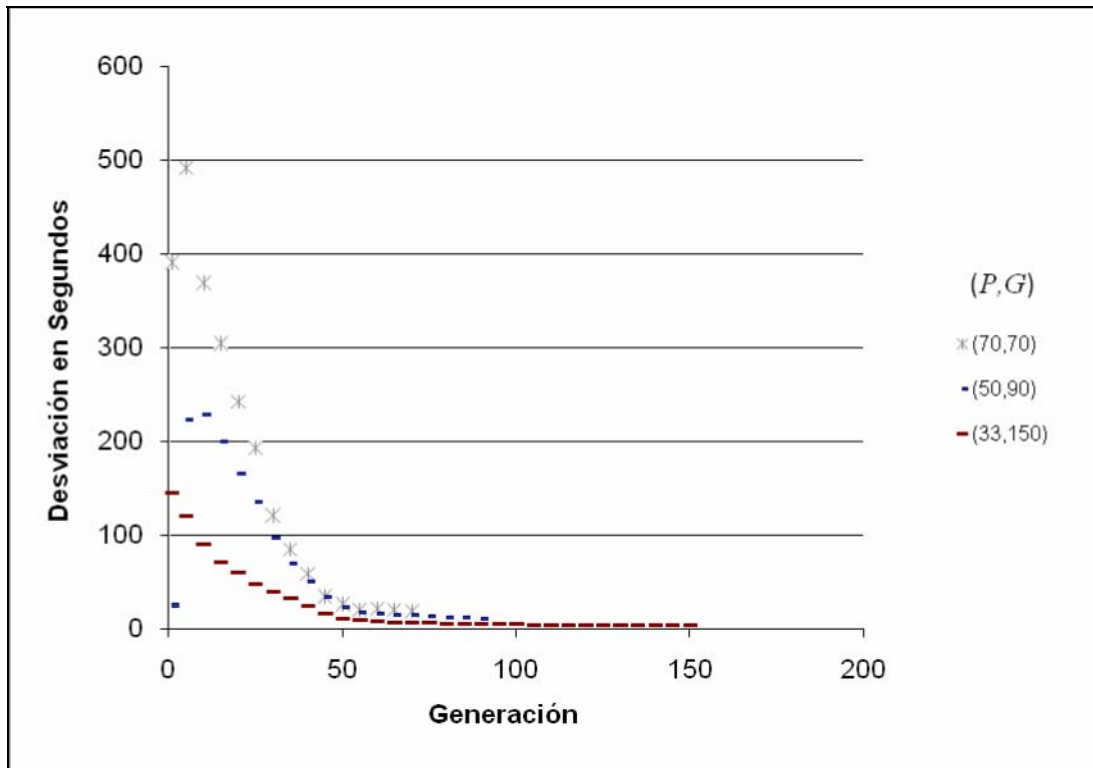
Figura 4.5. Evolución de la Mejor Solución.



En la figura 4.6 se presenta la gráfica de las desviaciones del Tiempo Total de Viaje (TT) en las diferentes soluciones de cada generación, suavizada con el método de promedios móviles con 20 generaciones. De ésta se concluye la importancia de mantener un tamaño de población grande, lo cual garantiza la diversidad genética potenciando la búsqueda exploratoria del modelo.

De ambas gráficas se concluye que existe una mayor posibilidad de mejora de la mejor solución en cada generación cuando se mantienen diversificada la población; conforme las iteraciones del proceso crecen, esta diversificación disminuye lo cual lleva a la convergencia del proceso hacia resultados óptimos.

Figura 4.6. Desviación del Tiempo Total de Viaje



El efecto de la probabilidad de mutación P_m es presentado en las figuras 4.7 y 4.8; considerando probabilidades de 0.15, 0.25 y 0.35. Los valores de P , G y $\acute{e}lite$ son fijados en 70, 70, 2 respectivamente. La mejor solución es encontrada para $P_m = 0.25$, sin embargo, no se puede explicar a partir de las gráficas la influencia que tiene el parámetro P_m en el tiempo total óptimo de cada generación ni en la diversidad de soluciones de cada generación. En primera instancia, se podría pensar que una alta probabilidad de mutación debería mantener una alta diversidad en los cromosomas de una generación facilitando la búsqueda exploratoria. Gracias al parámetro $\acute{e}lite$, podemos asegurar que la convergencia de los resultados está garantizada ya que al menos la mejor solución de cada generación siempre se mantiene en la siguiente generación.

Figura 4.7. Evolución de la Mejor Solución con respecto a P_m

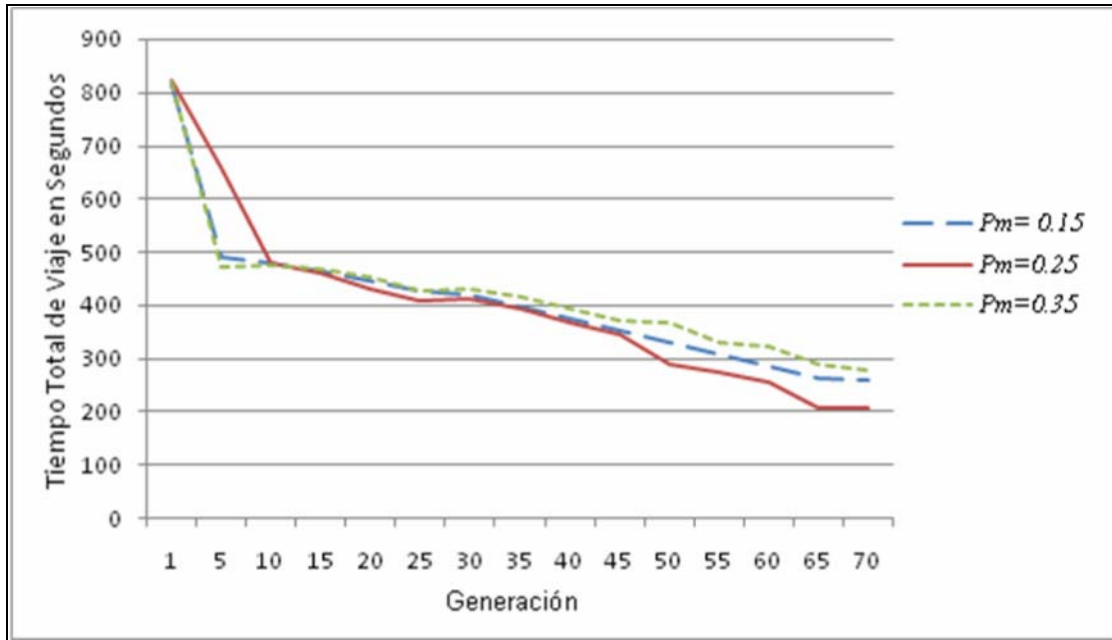
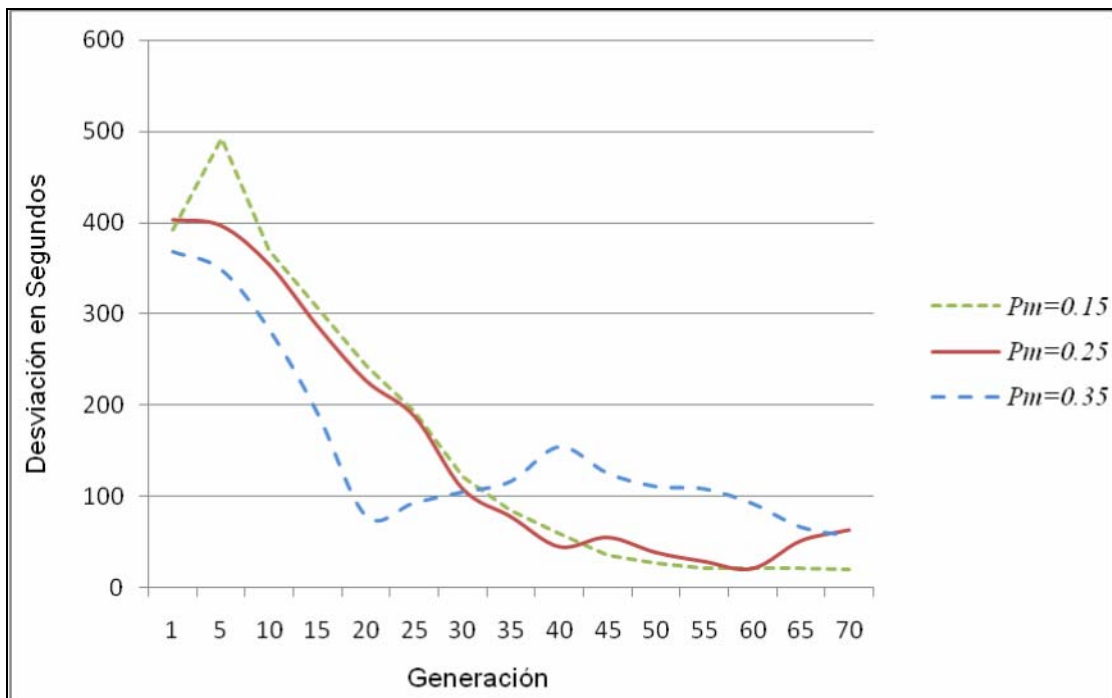


Figura 4.8 Desviación del Tiempo Total de Viaje con respecto a P_m



Conclusiones.

En ese trabajo se presentó la aplicación de un método de optimización de tiempo total de recorrido en una red vial considerando como variable de decisión la programación de los semáforos pertenecientes a la vía. El método aplicado integra la respuesta de los usuarios a la “impedancia” que representa la programación de los semáforos en la elección de sus rutas, adelantándose a los efectos que estos cambios generan en las distribuciones de los automóviles en las rutas dentro de la red.

Los Algoritmos Genéticos son aplicados considerando el problema de optimización como un problema binivel en donde la función a optimizar es el tiempo total de recorrido considerando como restricción el flujo de equilibrio que se da en la red conforme la Teoría Wardropiana. Los parámetros de decisión son, la duración de los ciclos de las señales semaforizadas en la vía, así como los lapsos de verde en cada una de las etapas de las mismas. El modelo arroja resultados promisorios en la localización de tiempos de señal óptimos con flujos estables (de equilibrio). Investigaciones futuras pueden considerar diferentes redes de tránsito con diferentes niveles de congestión.

El procedimiento puede ser corrido en muchas ocasiones utilizando diferentes semillas aleatorias para la generación de la primera generación de cromosomas con el fin de mejorar el comportamiento de la convergencia. La combinación del tamaño de población, número de iteraciones, probabilidad de mutación así como el diseño del cromosoma contribuye en la manera en cómo el Algoritmo Genético se desarrolla. La determinación de los valores óptimos para los parámetros de los algoritmos genéticos es actualmente un problema difícil que requiere investigación futura.

REFERENCIAS

- Allsop, R.E. (1974). Some Possibilities for Using Traffic Control to Influence Trip Distribution and Route Choice. In Sixth International Symposium on Transportation and Traffic Theory, pp. 345-374.
- Beckman, M, et. al. (1956). Studies in the Economics of Transportation. Yale University Press.
- Ben Ayed, O., Boyce, D.E., & Blair C.E. III (1988). A general bi-level linear programming formulation of the network design problem, Transportation Research, Part B, 23(4), 311-8.
- Bergendorff , P., Hearn, E.W., & Ramana, M.V. (1997). Congestional toll pricing of traffic network, in P.M. Pardalos, D.W. Hearn, and W.W. Hager (eds.). Lecture Notes in Economics and Mathematical Systems, 450: Network Optimization, Springer, Berlin, 51-71.
- BPR. 1964. Traffic Assignment Manual For Application with a Large, High Speed Computer. Washington,; U.S. Dept. of Commerce Bureau of Public Roads Office el Planning Urban Planning Division.
- Branston, D. (1976). Link Capacity Functions: A Review. Transportation Research 19, 223-236.
- Bullock, D. and Urbanik, T. (2000). Traffic Signal Systems.: Addressing Diverse Technologies and Complex User Needs. *Transportation in the New Millenium*. Transportation Research Board. (en línea). Disponible en:
<http://www.nationalacademies.org/trb/publications/millennium/00116.pdf>
- Cantarella, G.E., Improta, G. & Sortza, A. (1991). Iterative Procedure for Equilibrium Network Traffic Signal Settings. Transportation Research, 25A, 241-249
- CEPAL (2001). *Measures for Controlling Traffic Congestion*. Bulletin FAL 182. ISSN0256 9779.
- Ceylan, H. & Bell, M.G.H. (2004) Traffic signal traffic optimization based on genetic algorithm approach, including driver's routing. Transportation Research Part B, 38(4). 329-42
- Daganzo, C. & Sheffi, Y. (1977). On Stochastic Models of Traffic Assignment. Transportation Science, 16, 332-360
- Dantzig, G.B., Harvey, R.P., Landsdown, Z. F. Robinson, D.W. & Maier, S.F. (1979), Formulating and Solving the network design problem by decomposition. Transportation Research Part B, 13(1). 5-17.
- Davidson, K.B. (1966). A Flow Travel Time Relationship for Use in Transportation Planning. Australian Road Research Board 3, 183-194.

- Dickinson, T.J. (1981). A Note on Traffic Assignment and Signal Timings in a Signal Controlled Network. *Transportation Research* 15B, 267-271.
- Downs, Anthony (1992). *Stuck in Traffic: Coping with Peak-Hour Traffic Congestion*, The Brookings Institution: Washington, DC. ISBN-0-8157-1923-X
- Feather, Clive D.W. (2005). The Road Network Paradox.(en línea). Disponible en: <http://www.davros.org/science/roadparadox.html>.
- Foy, M., Benekohal, R.F. & Goldberg, D.E. (1992). Signal timing determination using Genetic Algorithms. *Transportation Research Record*, 1365. 108-15
- Frank, M. and Wolfe P. (1956). An algorithm for quadratic programming. *Naval Research Logistics Quarterly*, 3. 95-110.
- Friesz, T.L., Tobin, R.L., Cho, H.J. & Mehta, N.J. (1990). Sensitivity analysis based heuristic algorithms for mathematical programs with variational inequality constraints, *Mathematical Programming*, 48, 265-84.
- Friesz, T.L., Cho, H.J., Mehta, N.J., Tobin, R.L., & Anandalingam, G. (1992). A simulated annealing approach to the network design problem with variational inequality constraints, *Transportation Science* , 26. 18-26.
- Ghassan, A. & Benekohal R.F. (2000) Genetic algorithms for traffic signal control and queue management of oversaturated two-way arterials. *Transportation Research Record* 1727. Transportation Research Board. Washington, D.C., pp. 61-67
- Goldberg, D.E. (1989). *Genetic algorithms in search, Optimization and Machine Learning*. Addison Wesley Longman Inc., Reading., M.A.
- Goldberg, D.E. (2002). *The Design of Innovation. Lesson from and for Competent Genetic Algorithms*, Kluwer Academic Publishers, Boston.
- Harker, P.T. & Friesz, T.L. (1984) Bounding the Solution of the Continuous Equilibrium Network Design Problem. In *Proceedings of the Ninth International Symposium Of Transportation and Traffic Theory*. Delft:VNU Science Press.
- Hearn, D.W. & Ramana, M.V. (1998). Solving Congestion toll pricing models, in S. Nguyen and P. Marcotte, (eds.) *Equilibrium and Advanced Transportation Modeling*, Kluwer Academic Publishers, Boston, pp 109-24.
- Hillier, F.S. y Lieberman, G.J. (2006). *Introducción a la investigación de operaciones*. 8ª Edición. Cap. 13 "Metaheurística". McGraw-Hill Interamericana, México
- IHT Institution of Highways and Transportation (1997). *Transport in the Urban Enviroment*. Institution of Highways and transportation, London.
- IMT, Instituto Mexicano del Transporte (2007). *Los Principios de Wardrop: Matemáticas de la Congestión Vehicular y Base del Peaje para Reducir su Costo Social*. Publicación de Divulgación Interna. Num 344.

- Head, L., Mirchandani, P.B. and Shelby, S. (2005). The RHODES Prototype: A Description and Some Results. *Transportation Research Record*.
- Lawphongpanich, S. & Hearn, D.W., (2004). An MPEC approach to second-best toll pricing. *Mathematical Programming B*. 101(1). 33-55
- Le Blanc, L. J. (1973). *Mathematical Programming Algorithms for Large Scale Network Equilibrium and Network Design Problems*. In Department of Industrial Engineering, pp. 142. Evanston, IL: Northwestern University.
- Le Blanc, L. & Boyce, D. (1986). A bi-level programming for exact solution of the network design problem with user optimal flows. *Transportation Research, Part B*, 20(3), 259-65.
- Lee, C. (1998). *Combined traffic signal control assignments: algorithms, implementation and numerical results*. PhD. Dissertation, The University of Texas at Austin, Texas.
- Marcotte, P. (1981). An analysis of heuristic for a continuous network design problem, in V.F. Hurdle et. al. (eds.), *Proceedings of the 8th International Symposium of Transportation and Traffic Theory*, University of Toronto Press, Toronto, 27-34.
- Meng, Q., Yang, H. & Bell, M. G. H. (2001). An equivalent continuously differentiable model and locally convergent algorithm for the continuous network design problem. *Transportation Research Part B*, 35(1). 83-105.
- Montana, T. Hussain, and T. Saxena (2002). *Adaptive Reconfiguration of Data Networks Using Genetic Algorithms*, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*.
- Ortúzar, J. de D. & Willumsen, L. (1994). *Modelling Transport*, 2nd. Edition, John Wiley U.K.
- Park, B., Messer, C.J. & Urbanik, T. (2000). Enhanced genetic algorithm for signal timing optimization of oversaturated intersections. *Transportation Research Record*, 1727. 32-41.
- Park B. & Yun I. (2005). *Stochastic optimization method for coordinated actuated signal systems*. Final Report of ITS Center Project: Evaluation of Advancing traffic signal controllers. Center for Transportation Studies, University of Virginia, U.S.A.
- Rico Galeana, Oscar Armando (1989), *La Regulación del Flujo Vehicular en Sistemas Viales*, Instituto Politécnico Nacional, 46-49, 51.
- Robertson, D.I. (1969). TRANSYT method for area traffic control. *Traffic Engineering and Control*, 10. 273-81.
- Sheffi, Y. & Powell, W. (1982). An Algorithm for the Equilibrium Assignment Problem with Random Link Times. *Networks* 12, 191-207
- Sheffi, Y. (1985) *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Englewood Cliffs, N.J.: Prentice-Hall.

- Smith, M.J., (1979). The existence, uniqueness and stability of traffic equilibria, *Transportation Research Part B*, 13(4). 295-304.
- Steenbrick, P.A. (1974). *Optimization of Transport Networks*. John Wiley & Sons, London.
- Suwansirikul, C. Friesz, T.L. & Tobin, R.L. (1987). Equilibrium Decomposed Optimization: A heuristic for the continuous equilibrium network design problem, *Transportation Science*, 21(4), 254-63.
- Syswerda, G. (1989). Uniform crossover in genetic algorithms, in *Proceeding of the 3rd. International Conference of Genetic Algorithms*, Morgan Kaufmann, 2-9.
- Taale H., & Van Zuylen (2003). The effect of anticipatory traffic control for several small networks. Paper presented at 82nd Transportation Research Board Annual Meeting, Washington D.C., January 12-16, 2003.
- Teklu, F, Sumalee, A. & Watling, D. (2007). A genetic algorithm approach for optimizing traffic control signals considering routing. *Computer-Aided Civil and Infrastructure Engineering*. Vol. 22. pp. 31-43.
- Verhoef, E.T. (2002). Second best congestion pricing in several networks: Heuristic algorithms for finding second-best optimal toll levels and toll points. *Transportation Research Part B*. 36(8), 707-29.
- Wardrop, J. G., (1952). Some theoretical aspects of road traffic research, *Proceedings, Institute of Civil Engineers, PART II, Vol.1*, pp. 325-378
- Webster, F.V., (1958). *Traffic Signal Settings*. Road Research Technical Paper 39. London: Road Research Laboratory.
- Webster, F.V. & Cobbe, B.M. (1966) *Traffic Signals*, Road Research Technical Paper. HMSO, London 56
- Whitley, D. (1989). The GENITOR algorithm and selection pressure: Why rank based allocation of reproductive trial the best, in J. Schaffer, (eds.) *Proceedings of the 3rd. International Conference of Genetic Algorithms*. Morgan Kaufmann, 116-21.
- Wikipedia. (2006). Wikipedia, the free encyclopedia. *Genetic algorithm*. (en línea). Disponible en: http://en.wikipedia.org/wiki/Genetic_algorithm
- Wikipedia. (2007a). Wikipedia, the free encyclopedia. *Braess' paradox*. (en línea). Disponible en: http://en.wikipedia.org/wiki/Braess_paradox
- Wikipedia. (2007b). Wikipedia, the free encyclopedia. *Lewis-Mogridge Position*. (en línea). Disponible en: http://en.wikipedia.org/wiki/Lewis-Mogridge_Position
- Wikipedia. (2007c). Wikipedia, the free encyclopedia. *Traffic Light/History*. (en línea). Disponible en: http://en.wikipedia.org/wiki/Traffic_light#History
- Yang, H. (1997). Sensitivity analysis for elastic-demand network equilibrium problem with applications, *Transportation Research Part B*, 31(1), 55-70.

- Yin, Yafeng, (2000). Genetic-Algorithms-Based Approach for Bilevel Programming Models. Journal of Transportation Engineering, 115-120.
- Zangwill, W.L. (1969). Nonlinear Programming; A Unified Approach. Englewood Cliffs, N.J. : Prentice-Hall.