

7. Planteamiento de objetivos e hipótesis del proyecto de investigación.

7.1 Introducción.

En este momento, los capítulos ya revisados anteriormente, nos serán de utilidad al tener presentes todos los antecedentes que existen alrededor del problema de investigación ya planteado (SPDP-sTW). A continuación los enumeramos con la intención preparar el terreno para enunciar el objetivo de nuestro proyecto.

1. El problema planteado es de naturaleza combinatoria y tiene una complejidad documentalmente probada a través de proyectos de investigación en el mundo al estar catalogado como NP-Hard, lo cual lo hace matemáticamente muy difícil de resolver.
2. El problema de investigación planteado posee ciertas características específicas que lo hacen diferenciable de proyectos de investigación previos y lo sitúan como una variante digna de exploración. El sustento de lo anterior puede ser verificado en el apartado 6.8.
3. En la práctica de la Logística de distribución de las empresas es común encontrar este tipo de problemas. Particularmente podemos referenciar a las empresas que se dedican a la entrega y recolección en ruta de productos que alternativamente requieren incluir en su proceso la *retornabilidad* del producto. Ejemplos de este tipo de negocios, van desde la industria embotelladora de refrescos y cerveza, pasando por las botanas y el pan, continuando con los productos lácteos y los jugos y terminando con los dulces y los cigarros.
4. Las instancias del problema de investigación, comúnmente encontradas en la práctica de las empresas, oscilan en el rango de los 60 a los 100 clientes (nodos) por ruta al día.
5. En la operación diaria y tomando en consideración el volumen de rutas de distribución a tener que administrar, normalmente las empresas disponen de poco tiempo para tomar decisiones respecto a cómo dar solución al problema en cuestión. Según la empírica del postulante, se

estima que 5 minutos pueda ser un tiempo que resulte práctico y útil aplicar como límite superior para el procesamiento correspondiente a la ruta de un vehículo.

6. Los algoritmos de solución exacta para solucionar problemas como el nuestro, consumen una cantidad de tiempo computacional que excede por mucho los 5 minutos que se acaban de definir en el punto anterior. Por tanto, lo anterior limita a los algoritmos de solución exacta en su aplicación en la práctica de las empresas debido a que estas últimas requieren soluciones en tiempo real. El sustento de lo anterior puede ser revisado en el apartado 6.1.
7. Lo anterior hace requerir soluciones que aunque no sean matemáticamente óptimas, si sean razonablemente buenas ($> 90\%$ del óptimo) y que puedan ser obtenidas en poco tiempo computacional (≤ 5 minutos).
8. Un porcentaje bajo (7 de 55) de las empresas muestreadas del área metropolitana de la ciudad de Monterrey (AMM) han intentado resolver este tipo de problema a través de la implementación de sistemas APS's. Las empresas del AMM no han tenido éxito con este tipo de sistemas. El sustento de lo anterior puede ser revisado en el estudio de campo desarrollado en el apartado 3.8.
9. Las empresas del AMM, en términos generales administran el proceso de la logística de manera manual y en base a la experiencia de la gente, lo cual limita la posibilidad de obtener mejores soluciones que logren generar beneficios en términos de disminución de costos. El sustento de ello también puede ser corroborado en el apartado 3.8.
10. En las empresas del AMM, es muy poco frecuente (menos del 6%) que los profesionistas lleguen a aplicar las matemáticas en la toma de decisiones en general ni tampoco para solucionar el problema de la logística de distribución en lo particular. El sustento puede ser revisado en el estudio de campo desarrollado en el apartado 2.5. De lo anterior, el postulante opina que dicha falta de aplicación de las matemáticas por parte de los profesionistas, inhabilita a su vez de modo indudable, el que haya la posibilidad de una mejora en la generación de soluciones para atender la problemática de la logística en la práctica.

7.2 Planteamiento del objetivo central del proyecto de investigación.

De todo lo antes expuesto, entonces surge la pregunta:

¿Cómo podemos encontrar soluciones razonablemente buenas y en un tiempo computacional práctico, para resolver el problema de Ruteo de distribución para entrega y recolección de producto con restricciones de ventana de horario negociables (SPDP-sTW) ?

De lo anterior, enunciamos ahora el Objetivo Central del Proyecto:

Determinar mediante un experimento de laboratorio si el Algoritmo propuesto e instrumentado a través de la aplicación de “Algoritmos Genéticos”, resulta ser viable para encontrar soluciones razonablemente buenas (arriba del 90% optimalidad) y en poco tiempo computacional (menos de 5 minutos) para resolver el problema de Ruteo de distribución para entrega y recolección de producto, con restricciones de ventana de horario negociables (SPDP-sTW). El postulante estima que ambos parámetros (arriba del 90% de optimalidad y en menos de 5 minutos) pueden ser de utilidad práctica para ser aplicados en la logística.

7.3 Planteamiento de los objetivos específicos y entregables del proyecto de investigación.

A continuación se enumeran los objetivos específicos del proyecto de investigación.

1. Desarrollar un generador computacional de *instancias aleatorias uniformemente distribuidas* con la finalidad de aplicarlas para la *instrumentación de la prueba experimental*.
2. Generar las *instancias numéricas de nuestro problema SPDP-sTW* que vayan a requerirse para la *prueba experimental* (revisar el apartado 8.11).
3. Modelar matemáticamente hablando el problema del SPDP-sTW a través de una formulación basada en técnicas de programación mixta entera (MIP).
4. Desarrollar e implementar computacionalmente un algoritmo de solución exacta basado en las estrategias de “Ramificación y Corte” (BC) para aplicarlo a las instancias generadas en el punto No. 2 y obtener así el cálculo para la cota inferior para el modelo generado en el punto No. 3.
5. Desarrollar e implementar computacionalmente un algoritmo metaheurístico basado en las estrategias de “Algoritmos Genéticos” (AG) para aplicarlo a las mismas instancias generadas en el punto No. 2 y resueltas hasta la optimalidad en el punto No. 4.
6. Desarrollar un experimento cuya *finalidad sea comparar las soluciones obtenidas con el Algoritmo Genético en el punto No. 6 versus la cota inferior obtenida en el punto No. 5.*

Es importante mencionar que fue en el apartado 6.4, en donde se estableció la pertinencia de aplicar el algoritmo de “Ramificación y Corte” (BC) para calcular la cota inferior. Es importante mencionar que aunque técnicamente el algoritmo de solución exacta basado en la técnica de “Ramificación y Corte” (BC) esta orientado en precisamente *obtener la solución óptima para el problema*, el objetivo de nuestro proyecto de investigación no está enfocado en demostrar la optimalidad de dicha solución. Como ya se definió antes, el propósito de lo anterior es que la solución obtenida mediante este método nos sirva como cota inferior para el análisis comparativo versus el Algoritmo Genético. A esta cota inferior, en términos estadísticos se le denomina “Grupo Control”.

7.4 Delimitación del proyecto de investigación.

El modelo matemático así como la propuesta algorítmica basada en la aplicación de un algoritmo genético, tienen el objetivo de reflejar los elementos que ocurren en el problema denominado “*Ruteo de distribución para un vehículo con entrega y recolección de producto con restricciones de ventana de horario negociables (SPDP-STW)*”. Se busca integrar las consideraciones y supuestos más relevantes, matemáticamente hablando, acerca de cómo el fenómeno se presenta en la realidad. De esta manera el proyecto doctoral a la vez que experimental será explicativo, ya que documentará en el modelo matemático, la forma mediante la cual las variables endógenas guardan una relación causal con las exógenas.

Además de lo anterior, el modelo experimental tiene como finalidad encontrar una solución cercana al óptimo para el problema a resolver. De este modo establecemos que el modelo experimental no tan solo busca simular o predecir la realidad del fenómeno, sino explorar el espacio solución de una manera eficiente de tal forma que sea posible aproximar la solución óptima para el problema. Cuando nos referimos al calificativo de eficiencia nos estamos refiriendo a que se reúnan las siguientes propiedades:

1. Dimensionalidad (*cantidad de nodos en cada instancia*): $100 \leq v \leq 120$
2. Tiempos de ejecución computacional ≤ 5 minutos
3. Grado de calidad : *porcentaje de optimalidad* $> 90\%$

La evidencia de “Viabilidad” del modelo propuesto será a través de la vía de la cuantificación matemática de su efectividad a ser evaluada durante el desarrollo del análisis experimental. Así pues la “viabilidad” será abordada de dos maneras:

1. Que el modelo logre abstraer matemáticamente el fenómeno que se presenta en el problema.
2. Que el modelo logre resolver eficientemente el problema.

En lo concerniente al producto entregable, éste será presentado en 3 modalidades:

1. En Lenguaje Natural: con la finalidad de explicarlo.
2. En Lenguaje Matemático: para brindar formalidad.
3. En Lenguaje Computacional: para implementarlo y probarlo.

Es importante precisar que la definición de las variables independientes y dependientes en el modelo propuesto no depende de pruebas de correlación estadística sino más bien de acuerdo a la empírica del

postulante y apoyado por la evidencia documental bibliográfica perteneciente a otras investigaciones que se hayan realizando en proyectos similares al problema del SPDP-sTW.

Aunque ya se había precisado en la exposición de los objetivos específicos en el apartado 7.3, que el presente proyecto de investigación no está enfocado en probar la optimalidad de las soluciones arrojadas por los algoritmos propuestos, es importante mencionar que de haber sido ese el objetivo hubiera existido una dificultad probada para poder llevar a cabo un análisis comparativo entre la propuesta del proyecto de investigación versus las propuestas provenientes de investigaciones previas en el tema. La razón de lo anterior obedece a los siguientes argumentos:

1. La mayor parte de los proyectos de investigación provienen de una propuesta de aplicación para problemas específicos y muy particulares. Las restricciones impuestas en el problema suelen ser diferentes, los objetivos también pueden variar dependiendo de la empresa en donde se va a resolver el problema y más aún el significado de algunos conceptos, tales como el nivel de servicio al cliente, pueden llegar a ser diferentes en cada negocio. Como evidencia de lo anterior, se puede revisar la tabla 6.2, en la cual se muestra en forma comparativa, los estudios de investigación que fueron encontrados en la revisión bibliográfica.
2. Es fácil entonces comprender que la cantidad de problemas específicos resulta inmenso y es poco probable que al menos dos proyectos de investigación ataquen exactamente la misma variante del problema.

Finalmente en lo que se refiere a la problemática propia de la implementación en la práctica del modelo, este aspecto está limitado por la participación de los recursos económicos y humanos de las empresas, por lo que en este proyecto de investigación, la etapa de la implementación queda fuera del alcance comprometido. No obstante, las instancias numéricas que serán utilizadas para demostrar la viabilidad de los modelos, serán tomadas a partir de condiciones típicas de la realidad.

7.5 Justificación científica del problema de investigación.

De acuerdo a la revisión bibliográfica realizada, hemos procurado establecer las similitudes al igual que las diferencias que el planteamiento de nuestro problema guarda respecto a otros proyectos de investigación desarrollados previamente. Lo anterior no es trivial ya que precisamente lo anterior contribuirá a dar sustento y justificación a la investigación desde el punto de vista científico. Recordemos que en el apartado 6.8, la delimitación del problema de investigación contribuyó a evidenciar algunas características específicas que hacen diferenciable y a la vez justificable la variante de nuestro problema de investigación. No obstante creemos importante enumerar los siguientes argumentos adicionales para reforzar la justificación del proyecto:

1. El problema del PDP-TW apunta hacia una de las más ricas clases de problemas combinatorios en el área del OR/MS. El PDP-TW es un problema NP-Hard¹²⁴.
2. En la revisión bibliográfica del PDP-TW, hemos identificado algunos “huecos”; entre ellos descubrimos que la variante para la consideración de ventanas de horario negociables ha sido un tema de investigación muy poco explorado en términos relativos¹²⁵.
3. Adicionalmente, el problema del PDP-TW, en relación a sus áreas de aplicación, la variante menos explorada ha sido la relacionada al tema de la distribución física para entrega y recolección de producto y bienes de consumo¹²⁶.
4. La variante de nuestro problema de investigación SPDP-sTW posee como parte de su estructura, ventanas de horario relativamente anchas (arriba del 50% de apertura). Esta propiedad ocasiona por consecuencia que el problema tenga un espacio solución más grande.
5. El PDP-TW usualmente en la práctica, se trabaja con instancias de gran escala (instancias arriba de 70 clientes). Lo anterior hace prever la necesidad de investigar algoritmos de solución que puedan ofrecer soluciones razonablemente buenas en tiempos de ejecución

¹²⁴ Supra: revisar apartado 6.1

¹²⁵ Supra: revisar apartado 6.7

¹²⁶ Mitrovic, Snezana. (1998), *Pickup and Delivery Problem with Time Windows*, “Technical Report SFU CMPT TR 1998-12”, Canada, pag 38-39.

computacional cortos ante la falta de tiempo para aplicar métodos de enumeración exhaustiva.

6. Se anticipa el requerimiento de un algoritmo que sea lo suficientemente “veloz” como para ofrecer soluciones aceptables en la práctica en un ambiente de toma de decisiones en línea.
7. Se requiere un algoritmo que aunque no ofrezca una solución matemáticamente óptima a costa de un tiempo computacional fuera del alcance práctico, si al menos ofrezca soluciones que sean razonables tanto en calidad como en el tiempo computacional requerido.
8. De acuerdo a la investigación bibliográfica que el postulante ha realizado, proveniente de proyectos de investigación previos en el mundo, es relevante mencionar que aunque los proyectos de investigación en el tema de Algoritmos Genéticos es amplísimo, no obstante es la primera ocasión en la que se va a experimentar con un algoritmo genético como propuesta de solución para el problema del SPDP-sTW en particular.
9. Los modelos y algoritmos propuestos por el postulante, buscan aprovechar las similitudes que nuestro problema tiene respecto a investigaciones hechas previamente. Las experiencias revisadas provenientes de otros trabajos de investigación nos hacen prever un comportamiento positivo para las iniciativas que serán propuestas.

No hay duda de que la familia de problemas del tipo PDP-TW son más difícil de resolver aún que el propio VRP-TW. La razón de lo anterior es puesto que el primero trata de una generalización del segundo ¹²⁷. Es decir, la familia de problemas pertenecientes al VRP, provienen de aplicaciones a problemas de naturaleza estática, mientras que los problemas de la familia del PDP son de naturaleza dinámica.

El razonamiento anterior justifica el dedicar el objetivo de la investigación no a probar la optimalidad de las soluciones ofrecidas por los algoritmos sino más bien enfocado en medir su efectividad sobre la base del costo-beneficio que resulta al ofrecer soluciones razonablemente aceptables para aplicaciones a ser implementadas en tiempo real donde hay poca disponibilidad de tiempo para ejercer la toma de

¹²⁷ Palmgren, Myrna. (2001), *A Column Generation Algorithm for the Log Truck Scheduling Problem*, Department of Science and Technology (ITN), Linköping University, Norrköping Sweden, pag 3.

decisiones. Esto último es el fundamento que precisamente caracteriza a los ambientes de distribución dinámicos.

Más allá de los argumentos ya antes revisados, es innegable que desde el punto de vista científico este tipo de problema es fascinante por su complejidad matemática. Investigadores en los campos de las matemáticas, las ciencias computacionales y por supuesto del OR/MS han dedicado décadas en tratar de desarrollar algoritmos que resuelvan eficientemente el problema en cuestión¹²⁸.

¹²⁸ Lenstra, K. (1990), *A Variable Depth Approach for the Single-Vehicle Pickup and Delivery Problem with Time Windows*, "COSOR No. 90-48", Eindhoven University of Technology, Alemania, pag 35-90.

7.6 Relevancia práctica del problema de investigación.

En el capítulo 3 se evidenció que uno de los problemas más comunes en la práctica de la logística operativa de las empresas corresponde al problema de distribución de ruta. De acuerdo a la empírica del postulante, se estima que la mayor parte de las operaciones de entrega y recolección de producto en ruta en la práctica real de las empresas, están sujetas a restricciones de horario en el cliente y a restricciones de capacidad de carga. El proyecto de investigación se enfoca precisamente a este tipo de problemas (SPDP-sTW). A continuación se exponen algunas argumentaciones, las cuales tienen como objetivo dar a conocer la relevancia práctica del problema de investigación:

1. No es trivial que en las empresas que enfrentan los problemas de logística de ruteo, actualmente requieren de sistemas (o software) especializados con la finalidad de obtener rutas económicas para la distribución de sus productos. No obstante, la oferta computacional (sistemas APS) para resolver este tipo de problemas es cara y poco rigurosa.
2. Las empresas tales como las muestreadas en el antecedente del capítulo 3, descansan este tipo de decisiones sobre una base dependiente de la experiencia de la gente que viene ejecutando esta operación de manera empírica. Lo anterior limita la posibilidad de obtener mejores soluciones que logren generar beneficios en términos de disminución de costos.
3. Resultan evidentes los costos asociados en la solución óptima del problema para los cuales apenas se consiga un ahorro marginal en su ejecución, pueden llegar a resultar en ahorros sustanciales en términos absolutos para la operación del negocio ¹²⁹. Por tanto, se justifica el interés para dedicar investigación en este problema y conseguir estos ahorros potenciales en el ámbito de la distribución física de productos.
4. Aunque la optimización de los costos en la distribución física es un objetivo frecuentemente buscado en las empresas, no obstante existen otros beneficios que no deben ser subestimados. Por ejemplo, en el planteamiento del problema se puede apreciar que la necesidad de tener que considerar una ventana de horario para la atención de cada uno de los clientes, representa una fuerte necesidad y compromiso con el nivel de servicio que la empresa esté dispuesta a ofrecer.

¹²⁹ Savelsbergh, M. (1998), *Local Search in Physical Distribution Management*, Eindhoven University of Technology, Rotterdam Holanda, pag 5.

7.7 Propuesta metodológica para la solución del problema de investigación.

El presente apartado tiene el objetivo de exponer los elementos desarrollados en la propuesta de solución para el problema de investigación. Dicha propuesta de solución está compuesta por 4 rutinas de pre-procesamiento y otra de post-procesamiento en referencia a donde aparece la rutina del Algoritmo Genético Generacional. A continuación se enumeran las 6 fases:

1. Fase de descomposición basado en el método SPP (problema de la ruta más corta o Shortest path problem en inglés) para el análisis y explotación topológica de la versión original de la red logística.
2. Fase de compresión vía estrategia de clusterización mediante criterios de maximización de afinidad.
3. Fase de compresión discriminante de arcos basada en la heurística de los “k” nodos vecinos más cercanos para obtener la versión compacta de la red.
4. Fase de generación agresiva de cortes aplicada a la versión compacta de la red.
5. Fase evolutiva generacional para explotar el “pool” de cortes generados en el paso 4.
6. Fase de descompresión para desagregar la ruta propuesta para la versión original de la red.

Es importante precisar que los parámetros utilizados para el funcionamiento de los 6 algoritmos, han sido aplicados en función a la empírica del investigador. Esta fuera del alcance del objetivo del proyecto de investigación, el determinar como los valores de los parámetros empleados en estos algoritmos puede llegar a modificar el comportamiento en el desempeño de los resultados obtenidos. El objetivo del proyecto de investigación es medir y contrastar el efecto producido en el % de optimalidad obtenido mediante esta metodología con respecto al grupo control al hacer variar el recurso computacional aplicado. A continuación se presentará de manera general los aspectos técnicos que describen a los algoritmos implementados en las seis (6) fases de la propuesta de solución.

Fase # 1: Fase de descomposición basada en el método SPP para el análisis y explotación topológica de la versión original de la red logística:

No hay duda que la complejidad topológica de los problemas de ruteo logístico está presente en los casos de la realidad. En ese sentido, el generador de instancias aleatorias desarrollado como parte del proyecto de investigación procuró simular dicha característica con la finalidad de apegar las instancias de investigación a los casos de la realidad. Dicha realidad en lo tocante a los aspectos topológicos del tránsito de una ciudad, tiene que ver con lo siguiente:

- a. Proporción de calles (arcos topológicos) con tránsito en ambos sentidos.
- b. Proporción de calles de un solo sentido (oriente, poniente, norte y sur).
- c. Proporción de calles sin salida.

Las instancias aleatorias serán generadas sobre la base de un 30% de incidencia aplicado al inciso “a”. Otro 60% para el inciso “b”, del cual 15% de incidencia corresponde a cada una de las 4 orientaciones cardinales. Y finalmente tenemos un 10% de incidencia aplicado para las calles sin salida.

Con lo anterior expuesto, es fácil darse cuenta que cada una de las esquinas incluidas en una zona a ser analizada, corresponde a un nodo topológico (virtual) a ser considerado como parte en la red. Se dice entonces que dicho nodo resulta ser virtual debido a que no es un nodo al cual se requiera llegar por razones de entrega y/o de recolección de los pedidos a ser atendidos en los clientes tangibles de una ruta. Luego entonces, la única razón por la cual dichos nodos virtuales requieren ser considerados como parte de la red topológica, es debido a que estos son requeridos como puntos de conexión para poder realizar un ruteo que resulte factible de acuerdo a las restricciones de tránsito en la zona a ser analizada.

En base a la empírica del investigador, podemos establecer que para un caso típico, una zona lo suficientemente amplia como para cubrir un total de al menos 100 clientes requeriría al menos un total de 50 manzanas, es decir 2 clientes por manzana. Luego entonces, si la definición topológica de cada manzana ocupa típicamente 4 nodos, entonces podemos definir que para un problema típico requerimos considerar alrededor de 200 nodos virtuales más adicionalmente otros 100 nodos que corresponden a los clientes reales a ser atendidos, es decir 300 nodos en total.

Con todo lo antes expuesto, el algoritmo requerido para esta primera fase de pre-procesamiento, requiere cubrir los siguientes 2 objetivos:

- a. Identificar la secuencia de movimientos requeridos en términos cardinales (norte, sur, oriente, y poniente) para trasladarse óptimamente entre cada uno de los clientes (nodos reales).
- b. Calcular el costo incurrido en dicho movimiento.

Así entonces, el investigador hará uso de un algoritmo basado en el SPP con la finalidad de cubrir ambos objetivos arriba mencionados. Dicho algoritmo SPP es aplicado de manera combinatoria sobre los 100 clientes (o nodos reales) que estén siendo considerados en cada instancia aleatoria a ser revisada en el experimento. El resultado de dicho procedimiento combinatorio es entonces vaciado en dos matrices independientes de magnitud (100 x 100). Los 10 mil elementos de cada matriz hacen referencia a la logística requerida para ir del cliente "i" al "j". La primera matriz está dedicada al cálculo del costo y la segunda matriz se enfoca en identificar la secuencia de movimientos requeridos.

En cuanto a la complejidad matemática del SPP, es probado que dicho problema NO es NP-Hard, ya que su tiempo de ejecución computacional crece de manera polinomial respecto al tamaño del problema. La experiencia computacional del investigador al aplicar el algoritmo del SPP para las instancias antes mencionadas (300 nodos), es que en menos de 70 segundos de ejecución computacional se tiene completamente resuelta la totalidad de los 10 mil elementos de las dos matrices ya antes mencionadas.

No hay duda que la contribución final de esta fase de pre-procesamiento de la red, es la de simplificar el tratamiento de la versión original de la red. Dicho de otro modo, este pre-procesamiento permite pasar de un total de 300 nodos a tan solo 100 nodos. Lo anterior reduce sensiblemente la cantidad de nodos y a la vez el espacio solución para el problema en cuestión lo cual será crucial para las siguientes fases del método propuesto.

En lo referente a la implementación computacional del presente algoritmo, su codificación puede ser revisada en el Anexo A del documento de tesis.

Fase # 2: Fase de compresión vía estrategia de clusterización mediante criterios de maximización de afinidad:

Los métodos de compresión de redes tienen en general la finalidad de agrupar (o clusterizar) “ n ” nodos de una red para obtener como resultado una cantidad más reducida de “ m ” meta-nodos (*donde: $m < n$*). El objetivo de lo anterior es que los “ m ” meta-nodos resultantes, puedan contribuir a generar una red mucho más sencilla para su tratamiento durante las etapas posteriores de optimización.

Habitualmente una red simplificada que resulta al aplicar este tipo algoritmos, se rige por criterios únicamente de contigüidad geográfica. Es decir el algoritmo intenta agrupar a los clientes sobre la base de la vecindad. No obstante lo anterior, nuestro problema de investigación no puede ser atendido tan solo en función a este único criterio. Las estructuras de las ventanas de horario que están asociadas a cada uno de los “ n ” nodos obtenidos como resultado de la fase # 1, requieren ser compatibles entre si para poder agruparse. Dicho de otro modo, aunque un nodo resulte tener afinidad geográfica con respecto a otro, si acaso estos dos nodos no comparten ventanas de horario “similares”, entonces el algoritmo no los agrupará para formar un solo meta-nodo.

Ahora bien, la “similitud” tanto geográfica como en lo referente a las ventanas de horario requeridas para que dos o más nodos se agrupen, va a depender fuertemente del factor de compresión que el investigador decida ejercer sobre la red original. A mayor factor de compresión se aplique como parámetro en el algoritmo, menor similitud se requiere y viceversa.

En cuanto a la generación de los meta-nodos, existen diversos enfoques que pueden ser aplicados y todos ellos dependiendo del atributo que este siendo analizado. Por ejemplo en el atributo geográfico, podemos promediar las coordenadas geodésicas (x , y) de los nodos a ser agrupados. En lo referente a las estructuras de las ventanas de horario, es posible que el meta-nodo tenga una ventana de horario que sea suficientemente amplia como para incluir a todos los nodos que estén siendo agrupados.

La experiencia computacional del investigador en las instancias probadas, *verifica que reducir una red de 100 nodos con un factor de compresión del 50%* resulta ser convenientemente suficiente para esperar un beneficio en el tratamiento computacional de la red resultante para las fases posteriores de optimización, pero sin comprometer la “pureza” de la red en la conformación de los meta-nodos. La codificación computacional del algoritmo puede ser revisado en el Anexo B del documento de tesis.

Fase # 3: Fase de compresión discriminante de arcos basada en la heurística de los “k” nodos vecinos más cercanos para obtener la versión compacta de la red:

La aplicación del método discriminante de arcos tiene el objetivo de eliminar todos aquellos arcos entre los nodos que resulten poco probables de poder ser utilizados en la solución de la ruta. Así entonces, aplicar este tipo de heurística de manera previa (pre-proceso) a cualquier tipo de algoritmo combinatorio puede contribuir significativamente a disminuir el recurso computacional necesario para obtener la solución a un problema.

Una heurística para discriminar aquellos arcos poco probables a ser utilizados en la solución de un problema de redes, es a través de la aplicación de criterios de vecindad. Es decir, para cada nodo “X” en cuestión se prohíbe el uso de aquellos arcos que apunten a los nodos “Yi” que no logren estar lo suficientemente cerca del nodo “X”. Un método de prohibición que puede ser aplicado es a través del uso de valores suficientemente diferenciados (valores grandes) en los coeficientes de la función de costo. Lo anterior asegura que la función objetivo del modelo matemático incorpore suficientes elementos informativos para penalizar el espacio solución del problema. De lo explicado anteriormente, es fácil comprender que a pesar de que la función objetivo se robustece, no obstante el total del espacio solución no logra disminuirse, ya que la cantidad de arcos válidos en la red, y por consecuencia de variables binarias, sigue estando inalterado. Por el motivo anterior, más que una estrategia de penalización lo que el investigador aquí propone, es un método de exclusión.

En el caso nuestro, el método que el investigador empleó está basado en la heurística de los “k” nodos vecinos más cercanos, donde “k” representa ser un subconjunto de los “n” nodos reales de la red. Como es previsible tenemos que típicamente el valor de “k” debe ser un valor muy inferior al valor de “n”, ($k \ll n$). Entre menor sea el valor de “k” logramos que el procedimiento combinatorio sea menos demandante de recurso computacional, pero en contraposición, se incrementa la probabilidad de dejar fuera del espacio solución la solución óptima para el problema en cuestión.

En base a la experiencia computacional del investigador, aplicar un 20% como factor discriminante resulta ser suficientemente robusto como para asegurar que el espacio solución del problema modificado no esté comprometido. De esta manera la cantidad combinatoria de arcos válidos se reduce drásticamente, por ejemplo a un 80% de la cantidad de arcos contenidos en el problema original. La codificación del algoritmo puede ser revisado en el Anexo C.

Fase # 4: Fase de generación agresiva de cortes aplicada a la versión compacta de la red:

Esta fase del método algorítmico resulta ser crucial, ya que es la que tiene el objetivo de encontrar lo más rápidamente posible una solución factible (no necesariamente la óptima) al problema que esté siendo tratado. Como ya se trató ampliamente a lo largo del capítulo 6 del documento, la principal intratabilidad que nuestro problema SPDP-sTW posee es en lo referente a las restricciones de las ventanas de horario de los nodos así como en la capacidad de carga del vehículo. Ambas restricciones se ven confrontadas por la actividad de entrega y recolección de producto a lo largo de la ruta.

Así entonces, encontrar para el problema una solución al menos factible lo antes posible en términos del tiempo del recurso computacional resultará ser una estrategia útil para ser explotada en el siguiente paso correspondiente a la fase evolutiva. Se habla de una generación agresiva de cortes, debido a que el algoritmo va agregando cortes al espacio solución del problema tomando en consideración lo siguiente:

1. Las estructuras de las ventanas de horario de los nodos y
2. La capacidad de carga disponible en el camión en cada punto de entrega y/o recolección.

Ambas consideraciones son utilizadas para ir agregando cortes en la función de desempeño del algoritmo genético (función fitness). En un sentido topológico, estos cortes tienen el objetivo de delimitar en forma deliberada las secuencias de visita hacia y desde cada uno de los nodos que van apareciendo en las soluciones tentativas de ruteo para el problema. Así entonces, en aquellas soluciones tentativas en las que se verifique que se esté incumpliendo con al menos una de las dos restricciones mencionadas en el párrafo anterior, es entonces que se van a ir agregando los cortes necesarios para que vuelvan a seguir apareciendo en iteraciones posteriores.

Se dice que la generación de cortes es agresiva, debido a que la primera solución factible que se obtiene al ir aplicando de manera iterativa estos cortes, no necesariamente es la óptima. Dicho de otro modo, los cortes agregados al espacio solución del problema en las primeras iteraciones, pueden luego estar en contraposición a los cortes agregados en iteraciones posteriores. Esto puede causar que algunos cortes que hayan sido agregados al “pool” eventualmente estén en el mejor de los casos como redundantes y en el peor de los casos como innecesarios. Mantener en el “pool” del sistema de restricciones un corte innecesariamente equivale a estar en riesgo de dejar fuera del espacio solución algunas soluciones potenciales y mejores que la que actualmente se tiene. El algoritmo concluye una vez que se obtiene la primera solución factible para el problema.

Como ya se mencionó antes, en la presente fase algorítmica se agregan dos tipos de cortes. El primer corte tiene que ver con las estructuras de las ventanas de horario y el segundo con las restricciones referentes a la capacidad de carga del camión.

La lógica que se utiliza para generar los cortes que dan tratamiento a las ventanas de horario, inicia al identificar el nodo que tiene la mayor desviación respecto a su estructura particular de horario. Dicho de otro modo, una vez que se tiene un tour completo (ó factible), se verifica para cada nodo si éste cumple o no, con la estructura de horario impuesto. Si el nodo cubre la consideración entonces se prosigue a verificar el siguiente nodo incluido en el tour. En caso de que el nodo en cuestión, esté siendo arribado en el tour en un horario que esté fuera de la ventana de tiempo especificado, entonces quiere decir que el nodo no está siendo atendido correctamente. En tal caso se calcula la desviación en la estructura de horario para el nodo. Dicha desviación se calcula de la siguiente manera:

1. Si el tiempo en el cual se arriba al nodo es antes de la hora de apertura del nodo, entonces la desviación es igual a la diferencia entre el tiempo de apertura del nodo menos el tiempo en el cual se arriba al nodo.
2. En caso de que el tiempo en el cual se arriba al nodo es después de la hora de cierre del nodo, entonces la desviación es igual a la diferencia entre el tiempo en el cual se arriba al nodo menos el tiempo de cierre del nodo.

El procedimiento anterior se aplica para cada uno de los nodos del tour. Hecho lo anterior, se procede a identificar el nodo con la mayor desviación en el tour respecto a la ventana de horario, dicho nodo lo llamaremos “nodo pivote”. A continuación se verifican uno a uno todos los nodos del tour en los cuales su tiempo de arribo coincide con la estructura de horario del nodo pivote, a estos nodos se les denominarán “nodos afines”. Finalmente el algoritmo genera el corte a través de agregar una restricción que asegure que el “nodo pivote” se conecte a uno de los “nodos afines” a la estructura de horario. La generación del corte se representa matemáticamente de la siguiente manera:

$$\begin{aligned}
 & \text{sea } I = \{1..n\} \text{ (nodos de la red)} \\
 & k \in I \text{ (nodo pivote)} \\
 & j \subseteq I \{1..m\} \text{ (nodos afines a } k) \\
 & \sum_{j=1}^m (x_{jk} + x_{kj}) \geq 1 \quad \forall k \subseteq I
 \end{aligned}$$

La restricción anterior, asegura que el nodo pivote “ k ” utilice al menos uno de los arcos que lo conectan a cualquiera de los nodos afines “ j ”. Los cortes generados mediante esta estrategia agresiva, se guardan convenientemente en un pool de restricciones separados del resto de las restricciones del sistema global. El procedimiento antes descrito se aplica de manera iterativa hasta que el recurso computacional encuentra la primera solución factible al problema. Dicha solución factible tendría que ser con un solo tour completo y respetando las estructuras de las ventanas de horario.

A partir de que se obtiene la solución factible anterior, entonces ya solo quedan por considerar las restricciones de capacidad de carga. Para esto se vuelve a aplicar el mismo procedimiento de cortes agresivos ya antes expuesto. Las diferencias en esta otra ocasión serán:

1. El criterio de desviación para identificar el nodo pivote ya no será el de las ventanas de horario, sino será aquel nodo que mantenga el máximo exceso de carga requerida a lo largo del tour con respecto a la capacidad de carga del vehículo.
2. Los nodos afines para el nodo pivote, igualmente se identifican a partir de revisar aquellos nodos que pueden ser conectados hacia el nodo pivote sin que se genere alguna violación a la restricción de capacidad de carga del vehículo.

Una vez que el recurso computacional encuentra la primera solución factible para el problema global, entonces el algoritmo finaliza el procesamiento computacional de la fase de generación de cortes. Dicha solución factible global es aquella que corresponde a:

1. Un tour completo que cubra el total de los nodos de la red.
2. Que el tour asegure que a cada nodo se llegue dentro de la ventana de horario estipulada.
3. Que las operaciones de carga y descarga aplicadas al arribar a cada uno de los nodos que conformar el tour, no violen la capacidad de carga del vehículo.

La codificación computacional del algoritmo que se ha venido explicando a lo largo del presente apartado, puede ser revisada en el Anexo D.

Fase # 5: Fase evolutiva generacional para explotar el “pool” de cortes generados:

La fase evolutiva de la metodología es en la que se busca aproximar la solución óptima para la versión compacta de la red logística definida en la fase # 3. Como es de esperarse, la cantidad de cortes que llegan a acumularse en el “pool” del sistema de restricciones es significativo. La experiencia computacional del investigador indica que la cantidad de cortes que tienen que ser tratados en la presente fase evolutiva llega a estar en el orden de los 15 a los 40 cortes dependiendo de la instancia numérica que esté siendo analizada. Entre más intrincada y/o compleja sea la estructura de las restricciones del problema, mayor es la cantidad de cortes necesarios para lograr encontrar la primera solución factible para un problema.

Es fácil comprender que el proceso de aproximación a la solución óptima verdadera de la versión compacta de la red se logra al identificar cuales de los cortes incluidos en el “pool” son convenientemente necesarios eliminar del “pool”. Dicho proceso de eliminación de cortes no puede ser visto de manera individual en cada corte, ya que la presencia y/o la eliminación de un corte puede estar comprometiendo simultáneamente la presencia y/o la eliminación de otro o de otros cortes a la vez. Así entonces, identificar cuáles son los cortes que conviene eliminar en el pool, debe ser visto cómo un sub-problema de naturaleza combinatoria.

La circunstancia antes descrita, hace que el problema sea apropiado para ser atendido mediante el uso de estrategias de recombinación evolutiva a través del uso de codificación binaria. La codificación binaria permite naturalmente representar dos posibilidades, las cuales en nuestro caso significan la eliminación (0) o la presencia (1) del corte en el “pool” del sistema de restricciones. Si acaso ejemplificáramos una instancia con 30 cortes congregados en el “pool”, entonces estamos hablando de que se requieren de 30 genes para conformar la estructura genética del organismo que representa la solución potencial para el sub-problema del “pool” de cortes. La cantidad de combinaciones resultante de este sub-problema es igual a $30!$ (30 factorial, es decir $2.65252859812191 \times 10^{+32}$).

Aunque el espacio solución del sub-problema anterior es enorme, no obstante resulta mucho menor si lo comparamos con respecto al dimensionamiento del problema correspondiente a la versión compacta de la red. Para efectos de ejemplificar el dimensionamiento de una instancia típica para este último caso pudiéramos estipular los siguientes parámetros:

- a. Cantidad de nodos requeridos a tratar en la versión original de la red = 100 nodos
- b. Cantidad de nodos resultantes de la fase de compresión vía clusterización = 50 nodos
- c. Cantidad de arcos resultantes de la fase de compresión discriminante (obtención de la versión compacta de la red) = 50 nodos x 20 vecinos = 1000 arcos
- d. Aplicando la propiedad de simetría de la red tendríamos entonces un total = $1000 / 2 = 500$ arcos diferentes.

Luego entonces, si aplicáramos equivalentemente una codificación binaria para resolver este problema combinatorio, tendríamos entonces un total de combinaciones igual a $(500)!$. Nótese que el número más grande al cual se le puede calcular su función factorial en el paquete computacional de Microsoft Excel es $170!$ (170 Factorial).

Como puede comprobarse, resolver nuestro sub-problema basado en la solución combinatoria del "pool" de cortes resulta ser extremadamente menos difícil en comparación a resolver el problema referido a la versión compacta de la red, ya no digamos si acaso se trata de solucionar el problema correspondiente a la versión original de la red.

En cuanto a la implementación del algoritmo evolutivo, la codificación binaria de los cortes se complementa con operadores de selección tipo torneo y con un factor de cruzamiento al 50%. Es decir, el material genético del organismo padre y del organismo madre es al 50% cada uno. El método de reproducción es a través de dos puntos de cruce generados aleatoriamente a lo largo de la longitud de la cadena cromosómica.

Por otro lado, el factor de mutación es inicializado a un 7.5% en la primera generación, pero se va auto-ajustando en cada generación subsecuente de acuerdo al cálculo porcentual de individuos que sean idénticos genéticamente hablando en la población. Dicho de otro modo, a medida que la población aumenta su nivel de degeneración, en una medida proporcional se aplica una curva exponencial de crecimiento en el porcentaje de mutación con un límite asintótico al 50%. Esto quiere decir que en el mayor de los casos, por cada dos nuevos organismos que nacen en la población, uno tendrá mutado su contenido genético en cierta posición cromosómica elegida de manera aleatoria. Finalmente el factor de elitismo está limitado a un 15% sobre el total de la población total.

La codificación del algoritmo aquí descrito puede ser revisada en el Anexo E.

Fase # 6: Fase de descompresión de la ruta propuesta para la versión original de la red:

La última fase de la metodología corresponde a la descompresión de la ruta a ser propuesta como solución para la versión original de la red. Aunque el proceso de optimización evolutivo se lleva a cabo durante la ejecución computacional de la fase # 5, no obstante, la solución obtenida corresponde para la versión compacta de la red y por tanto no resulta ser suficiente para dar respuesta al problema original en cuestión. De ahí que sea necesario incorporar en la metodología una última fase de post-procesamiento con la finalidad de ofrecer una solución que satisfaga los requerimientos iniciales.

Nuestra fase de post-procesamiento tiene el objetivo de propagar la solución de la red compacta en una solución que sea topológicamente equivalente a la versión original y a la vez asegure que las restricciones impuestas en la versión original del problema sean cubiertas. A las restricciones a las que nos estamos refiriendo son naturalmente a:

1. Las estructuras de las ventanas de horario de los nodos y
2. La capacidad de carga disponible en el camión en cada punto de entrega y/o recolección.

La presente fase contiene dos rutinas principales las cuales son:

1. Rutina de desagregación para convertir los meta-nodos formados en la fase # 2, en los nodos correspondientes a la versión original de la red. Por ejemplo para desagregar los 50 meta-nodos optimizados durante la fase # 5 evolutiva y convertirlos en los 100 nodos reales (clientes). Esta rutina de desagregación está enfocada en determinar la secuencia óptima en la cual los nodos que integran a los meta-nodos se deben desagregar. Por ejemplo, si acaso un meta-nodo “x” estuviera compuesto por 3 nodos y luego otro meta-nodo contiguo “y” tuviera 2 nodos, entonces el algoritmo debe establecer la secuencia óptima en la cual conviene ordenarse los 5 nodos entre sí dada la contigüidad de los dos meta-nodos “x, y”. Desde luego el problema no es tan sencillo, ya que por ejemplo en un caso como el nuestro, se tendrían que ordenar 50 meta-nodos considerando la secuencia obtenida en la fase # 5 evolutiva. No obstante lo anterior, la experiencia computacional del investigador muestra que el problema puede ser resuelto hasta la optimalidad en un tiempo computacional que no excede de los 15 segundos.

2. Rutina de ruteo topológico para convertir la red de nodos reales (clientes) en la versión original de la red. Debemos recordar que la versión original de la red puede llegar a contener típicamente 300 nodos, a continuación se ejemplifica lo anterior:
 - a. Cien (100) nodos reales (clientes).
 - b. Doscientos (200) nodos topológicos: a razón de 2 clientes por manzana, 100 clientes nos requeriría 50 manzanas. Si cada manzana ocupa típicamente 4 esquinas (nodos topológicos), entonces tenemos 200 nodos virtuales.

Básicamente lo que el algoritmo de ruteo topológico realiza, es establecer la secuencia de movimientos requeridos en términos cardinales (norte, sur, oriente y poniente) para trasladarse óptimamente entre cada uno de los clientes (nodos reales). Este proceso, debe tomar rigurosamente en cuenta la secuencia de visita establecida en la rutina de desagregación en donde los meta-nodos son convertidos de vuelta en los nodos reales (clientes) de la red.

Este proceso de reconversión descrito, hace uso de la información generada en la fase # 1. De esta manera, la rutina de ruteo topológico simplemente va sustituyendo las secuencias de visita entre los nodos reales (clientes) por las cadenas correspondientes a los movimientos cardinales que son necesarios para trasladarse de nodo en nodo. Dichas cadenas describen los movimientos en términos de dirección (oriente, poniente, norte y sur) que son óptimamente necesarios realizar en términos de los nodos topológicos (esquinas de las manzanas) que se requieren arribar para conectar a cada uno de los nodos reales (clientes) en la red. Igualmente, cada cadena de movimientos esta asociada al costo incurrido para hacerlo lo cual permite calcular implícitamente el costo de la ruta formada.

Con todo lo antes descrito, finalmente se tiene la solución propuesta para el ruteo de la versión original de la red. La codificación del algoritmo aquí expuesto puede ser revisada en el Anexo F del documento de tesis.

A continuación en la siguiente figura 7.1, se expone de manera esquemática el proceso algorítmico mediante el cual se resumen las 6 fases que fueron expuestas en el presente apartado y con lo cual el investigador da por concluida la explicación de la propuesta metodológica para el planteamiento del problema.

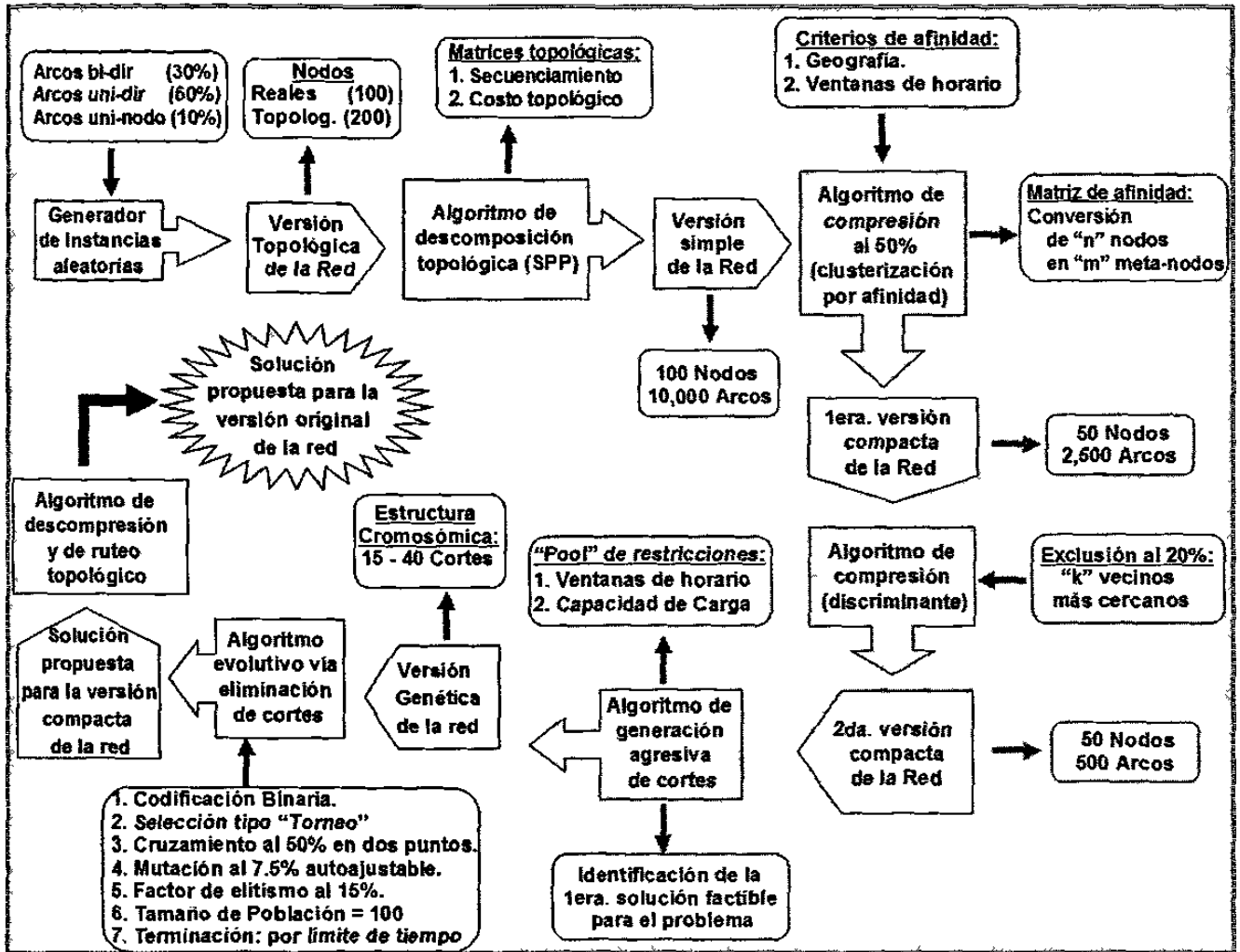


Figura 7.1 Diagrama de flujo para la propuesta de solución del SPDP-sTW.

7.8 Declaración de la hipótesis de trabajo del proyecto de investigación.

Con lo expuesto hasta este momento y tomando en cuenta la falta de evidencia teórica respecto a la aplicación exitosa de los algoritmos genéticos para la solución del problema SPDP-sTW, podemos establecer la siguiente hipótesis nula:

H₀: La metaheurística propuesta basada en la aplicación de un algoritmo genético para solucionar el problema de un vehículo único con entrega y recolección de producto con restricciones de capacidad de carga y con ventanas de horario negociables (SPDP-sTW), no logra mejorar el 90% de optimalidad en tiempos de ejecución computacional menores a los 5 minutos para instancias de tamaño $100 \leq w \leq 120$, (done w = cantidad de nodos en la red).

A continuación establecemos nuestra hipótesis de investigación:

H_a: La metaheurística propuesta basada en la aplicación de un algoritmo genético para solucionar el problema de un vehículo único con entrega y recolección de producto con restricciones de capacidad de carga y con ventanas de horario negociables (SPDP-sTW), antes de 5 minutos de esfuerzo computacional, logra mejorar el 90% de optimalidad respecto a la solución registrada en el grupo control y resuelto mediante un algoritmo exacto de Ramificación y Corte (BC) sin límite de tiempo, para instancias de tamaño $100 \leq w \leq 120$, (done w = cantidad de nodos en la red).

Resumiendo lo anterior, tenemos que:

$$*H_0*: \mu_{\text{Grupo experimental}} \leq 90\% (\mu_{\text{Grupo Control}})$$

$$*H_a*: \mu_{\text{Grupo experimental}} > 90\% (\mu_{\text{Grupo Control}})$$

Donde μ es la media poblacional correspondiente a los porcentajes de optimalidad alcanzados en la medición del 5to. minuto de esfuerzo computacional para el problema SPDP-sTW y para instancias de tamaño $100 \leq w \leq 120$, (done w = cantidad de nodos en la red).

7.9 Comentarios finales: taxonomía general del proyecto de investigación.

Resumimos al proyecto de investigación en término de las siguientes propiedades:

1. Tipo de Investigación: explicativa y experimental.
2. Tipo de Hipótesis estadística: análisis de varianza mediante estadística no paramétrica.
3. Instrumento de Investigación:
 - i. Algoritmo de solución exacta basado en generación de constructiva del hiperplanos o cortes (BC).
 - ii. Algoritmo meta-heurístico de aproximación basado en un Algoritmo Genético de estado estable (AG).
4. Diseño de la Investigación: Experimental / Experimento Puro
5. Tipo de la Investigación:
 - i. Es descriptivo: ya que busca identificar los "que's".
 - ii. Es explicativo: ya que procura exponer los "como's".
 - iii. Es predictivo: ya que reacciona de manera predecible a las "causas y a los efectos" que caracterizan al fenómeno en cuestión dando a conocer los "Porque's".
 - iv. Y sobre todo es prescriptivo: ya que finalmente, el objetivo que busca alcanzar es el hallazgo de la solución al problema, es decir busca su condición de operación óptima.

Así pues, en este último sentido y siendo el problema de investigación del tipo "prescriptivo", entonces es claramente legítimo que la metodología propuesta para el "análisis de superficie" utilizado para el estudio analítico del espacio solución de un problema, no será realizado a través de la aplicación de las técnicas de regresión múltiple sino más bien a través de la aplicación de modelos de optimización matemática.

Concluido el planteamiento de los objetivos e hipótesis de la investigación, pasaremos en el siguiente capítulo a exponer el diseño experimental que será propuesto para dar respuesta a la hipótesis presentada.

8. Diseño experimental.

8.1 Introducción: objetivos del esquema general del experimento.

El objetivo principal de nuestro diseño experimental es el de evaluar el comportamiento del método propuesto versus el grupo control. Lo anterior se logra mediante un análisis comparativo que permita determinar que tan lejos o cerca la solución del Algoritmo Genético está de la solución óptima verificable a través del grupo control. Así pues, el análisis comparativo busca dar respuesta a lo siguiente:

1. ¿Cómo se comporta el modelo propuesto durante el proceso de búsqueda computacional ?
2. ¿Qué tan robusto es el modelo cuando éste está expuesto a consideraciones variadas ?
3. ¿Cuál es la probabilidad de que el modelo consistentemente encuentre o aproxime la solución óptima de un problema ?

Debemos partir de que el problema de investigación, será abordado en el experimento a través de dos instrumentos algorítmicos:

- a. Un algoritmo de solución exacta basado en el método de Ramificación y Corte (BC) vía la generación de hiperplanos. Este instrumento nos servirá para el desarrollo del grupo control del experimento.
- b. Un algoritmo de aproximación meta-heurístico basado en un Algoritmo Genético. Este instrumento será confrontado versus el grupo control en lo referente a diversos criterios de efectividad que serán tratados más adelante.

En los problemas de programación lineal (PL), sucede habitualmente que la aplicación del método "Simplex" puede ser visto como una "caja negra". Es decir, no se requiere por parte del investigador hacer ningún tipo de manipulación en los parámetros de funcionamiento del algoritmo. Por otro lado en los modelos de programación mixta entera (MIP), siempre será necesario la intervención del investigador que contribuya a la manipulación de los parámetros del algoritmo que esté siendo utilizado para la solución del problema¹³⁰.

¹³⁰ Bixby, Robert. (1999), *MIP: theory and practice closing the gap*, Ilog Cplex Division, Department of Computational and Applied Mathematics Rice University, Houston EUA, pag 3.

Dada la naturaleza del problema de investigación (NP-Hard), es oportuno destacar que en ocasiones será muy difícil determinar la solución óptima exacta para una instancia particular del problema. Lo anterior es natural, debido a que nuestro problema de investigación tiene un tiempo de solución computacional que crece no en forma polinomial sino en forma exponencial. Con la explicación anterior, toma mayor relevancia aún el algoritmo de Ramificación y Corte (BC) ya que será éste el que defina al grupo control contra el cual se podrá medir el desempeño y la efectividad del Algoritmo Genético.

Así pues, el objetivo del algoritmo BC será el identificar la solución óptima verificable para cada instancia a ser aplicada como unidad experimental para el problema de investigación. Dicho de otro modo, la solución obtenida mediante el algoritmo BC será nuestra cota inferior para la prueba experimental a ser aplicada al Algoritmo Genético. Las cotas inferiores calculadas mediante la aplicación del algoritmo BC, aseguran una adecuada estimación del error experimental. Dicha estimación resulta ser una estrategia experimental bastante agresiva ya que garantiza obtener una medición precisa respecto al grado o efectividad con que el Algoritmo Genético está igualando el valor alcanzado por el grupo control identificado mediante el algoritmo de solución exacta basado en el BC. Lo anterior resulta favorable para el propósito de la validez del experimento del laboratorio.

El éxito del experimento no sólo depende de la forma en que se instrumente el Algoritmo Genético. El postulante está convencido de que la definición del grupo control jugará un papel crítico para la prueba de la hipótesis doctoral. Por tal razón, es plausible precisar que la contribución científica del experimento recae en la implementación de ambos instrumentos algorítmicos, es decir, tanto el Algoritmo de Ramificación y Corte (BC) así como el Algoritmo Genético.

El primer paso se da a partir de la implementación computacional del modelo matemático de programación mixta entera (MIP). Para el grupo control, éste se resolverá a través del algoritmo de planos de corte basado en la generación de hiperplanos vía el método BC. Para el desarrollo del algoritmo basado en el BC, nuestra propuesta será a través del estudio del poliedro que conforma el espacio solución para nuestro problema de investigación. Así entonces, el algoritmo propuesto, dinámicamente irá agregando hiperplanos (o cortes) a un sistema de restricciones del tipo $Ax \leq b$. La generación de los hiperplanos a ser manipulados en el experimento provienen naturalmente de la explotación de la estructura del problema asociado al polígono del espacio solución.

En general, las ideas desarrolladas en el campo científico en lo referente al método BC, han sido encaminadas hacia propósitos de investigación experimental. Hasta hace apenas 5 años, poco se había hecho en lo referente a su aplicación en el software de optimización para uso comercial. Como ejemplo de lo anterior podemos nombrar a LINGO©, OSL©, XPRESS-MP© y CPLEX© (marcas registradas), como aquellos que ya vienen incorporando mucho de estas capacidades y teorías acumuladas en el ámbito experimental ¹³¹.

Sobre la base de lo anterior, el algoritmo propuesto de investigación para la generación de los hiperplanos estará basado en un esquema de BC y se toma la decisión de implementarlo en el código computacional de XPRESS-MP© Versión (13.26). El postulante considera que el esquema de desarrollo algorítmico es robusto. A continuación en el apartado 8.2, se exponen los fundamentos del algoritmo basado en el BC para la conformación del Grupo Control.

¹³¹ Bixby, Robert. (1999), *MIP: theory and practice closing the gap*, Ilog Cplex Division, Department of Computational and Applied Mathematics Rice University, Houston EUA, pag 2.

8.2 Aplicación del procedimiento algorítmico de ramificación y corte (BC) como grupo control para la prueba experimental.

Un método trivial para resolver un problema de optimización combinatoria sería a través de una enumeración exhaustiva del espacio solución y seleccionando la mejor solución factible. No obstante el método anterior como ya se revisó antes puede consumir siglos o milenios en términos de recursos computacionales.

El método de “Ramificación y Corte” (o en inglés Branch & Cut, BC), fue desarrollado en 1960 por Land y Doig ¹³². Se trata de un algoritmo de enumeración parcial, el cual consiste en delimitar el espacio de búsqueda para la solución de un problema. Así pues, BC genera cortes (o restricciones) sobre la base de que cada uno de ellos delimite el espacio de búsqueda y por consecuencia logre incrementar la probabilidad de encontrar la solución óptima para el problema. Esta estrategia puede ser vista como la analogía al “divide y vencerás” en el que de una manera sistemática se van adicionando restricciones al problema original.

El algoritmo BC inicia resolviendo la versión relajada del problema original a atender¹³³. Lo anterior se logra simplemente mediante programación lineal. Si acaso la solución obtenida para el problema de programación lineal antes descrito logra satisfacer intrínsecamente las restricciones de integralidad, pues entonces decimos que el problema original ha sido resuelto. De otra manera quiere decir que es necesario agregar más restricciones (o cortes) que logren que el valor de las variables obtenidas mediante programación lineal no violen las restricciones de integralidad.

La condición anterior es lo que origina el proceso de ramificación. Es decir, a partir de cualquier variable x_j que no haya resultado entera (fraccional) se procede a ramificar dos sub-problemas descendientes del problema original. La primera ramificación buscará explorar el espacio solución correspondiente a un valor igual a cero, para esto se incluye la restricción (o corte) $x_j = 0$. La segunda ramificación buscará probar con el corte opuesto haciendo $x_j = 1$.

¹³² Land, A; Doig, H. (1960), *An Automatic Method for Solving Discrete Programming Problems*, Econometrica, EUA, pag 97-120.

¹³³ Supra: revisar pag 155.

A través de este procedimiento recursivo, se van desarrollando los cortes que van acotando el espacio solución obtenido mediante la relajación del problema original. Podemos resumir lo anterior mediante la siguiente Figura 8.1.

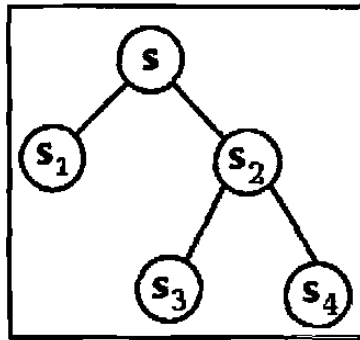


Figura 8.1 Proceso de bifurcación en el procedimiento de ramificación y corte.

Como puede ser revisado en la Figura 8.1, las bifurcaciones son generadas a partir de la adición de nuevas restricciones o cortes a partir del problema original S . Así entonces podemos definir lo siguiente:

$$S = S^1 \cup S^2; \text{ donde :}$$

$$S^1 = S \cap \{x \in R^+ : x \leq x^*\} \quad y$$

$$S^2 = S \cap \{x \in R^+ : x \geq x^*\}$$

Así pues, llevando al extremo el procedimiento anterior, esta sub-división de problemas pudiera parecer como un método de enumeración exhaustiva al cubrir todos los elementos en S . No obstante, si al procedimiento anterior le agregamos una rutina que permita revisar cuando ya no resulta necesario seguir bifurcando dentro de S_k , pues entonces lo que resulta es un procedimiento de enumeración parcial lo cual permite obtener ahorros en términos de recursos computacionales.

La explicación anterior resulta ser muy simple y general. Para ser un poco más formal definamos a P como el problema a tener que ser resuelto. Así entonces diremos que P tiene un conjunto de posibles soluciones factibles definidas por S . Entonces a lo largo de cualquier punto durante el proceso de búsqueda de la solución óptima para un problema, tendremos que S estará compuesta por un subconjunto S_0 de soluciones ya antes consideradas y como complemento otro subconjunto de soluciones aún por explorar S_1, S_2, \dots, S_n y así hasta agotar el total de soluciones factibles en S .

Matemáticamente hablando tenemos entonces que:

$$S=S_0 \cup S_1 \cup \dots \cup S_n$$

De acuerdo al procedimiento de búsqueda, cada uno de los S_k son mantenidos y actualizados en una lista de subproblemas candidatos P_k . A su vez para cada uno de los P_k se determina si acaso es necesario profundizar el proceso de búsqueda dentro del S_k correspondiente. Si la respuesta es negativa, entonces quiere decir que ya no hace falta profundizar dentro del sub-espacio solución representado por S_k y por tanto se procede a remover de la lista de problemas candidatos a P_k y se define a S_{k+1} como el siguiente espacio de solución a explorar. Si por el contrario, la respuesta es afirmativa entonces se aplica el proceso de enumeración al sub-espacio solución de S_k , lo cual se logra subdividiendo el sub-problema P_k en nuevos sub-problemas candidatos los cuales resultan al aplicar a S_k nuevos cortes y restricciones. El proceso antes descrito opera de una manera iterativa y recursiva hasta no tener ningún otro sub-problema que revisar de la lista de candidatos en cuyo caso se logra entonces determinar cual es la solución óptima.

El proceso mediante el cual se determina cuando un sub-problema P_k requiere subdividirse para profundizar la exploración de S_k se le denomina bifurcación. La bifurcación se desarrolla a partir de determinar para cada sub-problema P_k su límite superior U_k , su límite inferior L_k y su solución óptima definida por Z_k ; donde $L_k \leq Z_k \leq U_k$. En general, el proceso de bifurcación para un problema P_k no es necesario continuar cuando se haya verificado que el problema P_k no tiene soluciones factibles o cuando la solución óptima del sub-problema no resulte mejor que la última solución mejor encontrada hasta el momento a lo largo del proceso de búsqueda. A esta solución mejor encontrada se le denominada "titular" (o "incumbent" en inglés) y se le puede representar por I . Así pues, cuando $I \leq Z_k$ se identifica que ya no es necesario bifurcar en P_k .

Desde luego, si queremos aplicar lo anterior para un problema determinado, se requerirán dos cosas adicionales:

1. Un método adecuado para identificar en qué orden se irán revisando cada uno de los sub-problemas P_k . Lo anterior no es trivial, ya que entre más rápido se accedan los sub-problemas P_k que arrojen mejores valores óptimos Z_k , más rápido podrán ser desechados otros sub-

problemas P_k , lo que a su vez evitará tener que profundizar en sus sub-espacios S_k redundando en una acotación del espacio solución.

2. Un método adecuado para evaluar el límite superior U_k y el límite inferior L_k para cada sub-problema P_k . En general se busca lograr que los límites sean lo más acotados posibles, es decir que U_k y L_k estén lo más cerca entre sí. Entre más parecidos sean entre sí U_k y L_k , entonces se reduda en que la búsqueda de Z_k sea más rápida para converger. Tenemos entonces que el proceso de búsqueda de la solución óptima termina cuando:

$$U_k = L_k = Z_k$$

Los métodos utilizados para reducir el valor de U_k , han sido mediante heurísticas que logren encontrar lo antes posible las mejores soluciones enteras para el problema. Por su parte, los métodos utilizados para aumentar el valor de L_k , han sido mediante el desarrollo de algoritmos para la generación de planos de corte (o “cutting planes” en inglés) que logren limitar el espacio solución. A continuación resumimos el procedimiento BC a través del siguiente pseudo-código:

- 1: Seleccionar un sub-problema S_k y definirlo como activo.
- 2: Si S_k resulta infactible entonces
 - Borra S_k
 - Se vuelve al paso 1.
 - De otro modo
 - Calcula el valor óptimo (límite inferior) para S_k definido como L_k .
- 3: Si $L_k \geq I_k$ (*incumbent*) entonces
 - Borra S_k (No es necesario profundizar)
 - Se vuelve al paso 1.
 - De otro modo
 - Revisar si L_k es una solución factible para el problema original S .
 - Si es factible entonces
 - Se define $I_k = L_k$ (nueva solución titular o incumbent)
 - De otro modo
 - Se bifurca el problema S_k en S_{k1} y S_{k2} agregándose a la lista de sub-problemas.
4. Se vuelve al paso 1.

El procedimiento BC fue aplicado para el problema del TSP por George Dantzig, Ray Fulkerson y Selmer Johnson en 1964. El concepto aplicado por ellos fue explotar la estructura matemática del problema mediante un procedimiento recursivo a partir de la relajación del problema original y luego resuelto por aproximaciones sucesivas mediante programación lineal y un generador de cortes aplicados de manera incremental en cada iteración. Así el problema original del TSP definido por:

$$\text{Minimizar } c^T x \text{ sujeto a } x \in S.$$

Podría entonces en su versión relajada mediante programación lineal re-expresarse mediante:

$$\text{Minimizar } c^T x \text{ sujeto a } Ax \leq b.$$

La clave del planteamiento anterior enfrentado por Dantzig, Fulkerson y Johnson fue simplemente descubrir cual es el conjunto de restricciones o cortes a incluir en el sistema $Ax \leq b$. que logran que las variables “ x ” satisfagan el requerimiento de integralidad impuesto por “ S ” en el problema original. Efectivamente, la gran contribución de Dantzig, Fulkerson y Johnson fue identificar que al resolver la versión relajada del problema mediante un procedimiento de programación lineal basado en aproximaciones sucesivas, no solamente servía esto para ir progresivamente calculando el límite inferior del problema original. A la vez contribuyó a identificar el sistema de restricciones o cortes que lograban delimitar aquellas soluciones temporales de x^* que no cubrían los requerimientos impuestos por S . Así pues, los espacios solución pertenecientes a soluciones invalidas x^* podrían ser aisladas del espacio S a través de la inclusión progresiva de desigualdades dentro del sistema $Ax \leq b$. A estas desigualdades o cortes se les denomina hiperplanos. La generación de estos hiperplanos eventualmente logran delimitar y ajustar la relajación del problema original.

Así pues, el reto consiste en identificar aquel sistema de desigualdades lineales que satisfagan el conjunto de soluciones validas incluidas dentro del espacio S pero que a la vez violen la solución óptima x^* . Lo anterior logra que las restricción de integralidad de las variables discretas del problema original sean resueltas mediante la identificación de los hiperplanos lineales que logren “simular” el espacio solución del problema original. El reto anterior normalmente inicia con la aplicación del sentido común al requerirse explotar las propiedades combinatorias del problema y reflejarlas en su espacio solución S a través del uso de desigualdades lineales ¹³⁴.

¹³⁴ Applegate, David. (1995), *Finding Cuts in the TSP*, Princeton University, EUA, pag 5.

Lo que se busca es sustituir el sentido común antes mencionado por algún algoritmo que logre generar de manera automática las restricciones o cortes que se requieren para conformar el sistema de desigualdades requerido. Esta necesidad anterior es lo que da origen y sustento al campo de estudio en el área de las matemáticas discretas o combinatorias. Aunque el tratamiento de este último tema esta fuera del alcance del problema de investigación, no obstante se puede revisar la obra de Ralph Gomory quién intentó dar respuesta a esta pregunta a través de sus algoritmos para generación automática de cortes desarrollados en la década de los sesentas ¹³⁵.

La investigación bibliográfica realizada no deja duda acerca de la estrecha relación que existe en las estrategias de corte de Gomory como parte de la estructura algorítmica del método BC. Lo anterior resulta particularmente relevante sobre todo cuando se trata de dar solución a problemas de optimización combinatorios de gran escala ¹³⁶.

Sobre la base de lo antes expuesto, resulta necesario para el investigador, el apoyarse con alguna herramienta (o algoritmo) que le facilite la tarea de generar los hiperplanos que hagan falta durante el proceso de búsqueda de la cota inferior para cada instancia en particular. Así, tal y como se expuso en el apartado 8.1, será un esquema de BC implementado en el código computacional de XPRESS-MP© Versión (13.26), el cual será utilizado por el postulante para dicho propósito. Es precisamente a través de la utilización de dicha infraestructura (software) computacional, mediante la cual podemos obtener provecho del trabajo de investigación de Ralph Gomory ya antes expuesto en el párrafo anterior.

¹³⁵ Gomory, R. (1963), *An algorithm for integer solutions to linear programs*, Recent Advances in Mathematical Programming (R. L. Graves and P. Wolfe, eds.), McGraw-Hill, New York, pag 269-302.

¹³⁶ Padberg, M; Rinaldi, G. (1991), *A Branch and cut algorithm for the resolution of large-scale symmetric traveling salesman problem*, "SIAM Review No. 33", EUA, pag 60-95.

8.3 Técnica estadística para la prueba de la hipótesis de investigación.

La técnica que será aplicada para la prueba de la hipótesis central del proyecto será a través de un “Diseño de Experimentos”. La teoría relacionada al tema de “Diseño de Experimentos” es amplia y en términos de su teoría podemos citar al matemático inglés Ronald Aylmer Fisher como uno de sus principales precursores en el siglo XX. Los modelos de “Diseño de experimentos” son modelos estadísticos clásicos cuyo objetivo es averiguar si determinado(s) factor(es) influyen en la variable de interés. El objetivo del diseño de experimentos es determinar si acaso al utilizar un determinado tratamiento, se produce alguna mejora o no en el proceso¹³⁷. Para ello se debe experimentar utilizando un tratamiento y luego otro.

Es comprensible que si se repite un experimento, en condiciones de igualdad donde la mayoría de las causas de variabilidad están controladas (por ejemplo en el laboratorio), los resultados previsiblemente resultarán con una variabilidad pequeña¹³⁸. Uno de los principales objetivos del diseño de experimentos, es precisamente controlar la variabilidad de un proceso estocástico. En general, un experimento está sometido a tres tipos de variabilidad que a continuación se mencionan:

1. Variabilidad sistemática y planificada: esta variabilidad viene originada por la dispersión de los resultados debido a diferencias sistemáticas entre las distintas condiciones experimentales impuestas en el diseño deliberadamente por el investigador. Este es precisamente el tipo de variabilidad que se intenta identificar con el diseño estadístico. Cuando este tipo de variabilidad está presente y es cuantificada, entonces es posible lograr obtener inferencias acerca del proceso en cuestión.
2. Variabilidad no planificada del experimento: esta es la variabilidad no planificada debida a causas desconocidas (ruido aleatorio). Es una variabilidad impredecible. Una manera simple de explicar esta variabilidad es identificándola como la causante de que en un laboratorio, al tomarse medidas repetidas de un mismo experimento, ocurra que en algunos casos, la segunda medida no sea igual a la primera y, más aún, existe la posibilidad del error en la predicción del valor de la tercera. La presencia de esta variabilidad supone la principal causa de conclusiones erróneas. Sin embargo, bajo el aparente caos, existe un patrón regular de comportamiento en esas medidas, todas ellas tenderán a fluctuar en torno a un valor central y siguiendo un modelo

¹³⁷ Fisher, R. (1971), *The Design of Experiments*, Hafner Press & Macmillan Publishers, London, pag 25-50.

¹³⁸ Lorenzen, T. (1993), *Design of Experiments*, Marcel Dekker Inc, EUA, pag 67.

de probabilidad¹³⁹. Si el experimento es bien planificado, entonces será posible estimar dicha variabilidad, lo que permitirá obtener conclusiones y poder hacer inferencias.

3. La tercera fuente de variabilidad proviene del error humano al desarrollar el experimento de una manera desigual al momento de aplicar las mediciones en cada tratamiento.

El diseño de cualquier experimento se desarrolla con la finalidad de obtener mayor conocimiento acerca del comportamiento del proceso de interés. La experimentación forma parte natural de las investigaciones científicas, en muchas de las cuales, los resultados del proceso de interés se ven afectados por la presencia de distintos factores, cuya influencia puede estar oculta por la variabilidad de los resultados muestrales. Así entonces, es fundamental conocer los factores que influyen realmente y así entonces poder estimar esta influencia.

Para conseguir lo anterior, es necesario experimentar. Experimentar significa hacer variar planeadamente las condiciones que afectan a las unidades experimentales y así observar la respuesta en la variable de interés. Del análisis y estudio de la información recogida se obtienen las conclusiones. La forma tradicional que se utiliza en la experimentación, se basa en estudiar las variables controladas (factores) uno a uno, esto significa, hacer variar los niveles de un factor permaneciendo fijos los demás. En cambio, el diseño de experimentos busca sistemáticamente estudiar el efecto de todos los factores.

Aunque la metodología del diseño de experimentos se basa en la experimentación, para que la metodología sea eficaz es necesario que el experimento esté bien diseñado. Así entonces, la metodología del "Diseño de Experimentos" se concentra en establecer las condiciones de realización de un proceso para aumentar la probabilidad de detectar cambios significativos en la respuesta de la variable de interés¹⁴⁰. Dicha metodología se expondrá en el apartado 8.5. Pero antes que exponamos la metodología, definiremos en el siguiente apartado 8.4, la terminología que estaremos empleando a fin de empezar a relacionar el significado de dichos términos respecto a los elementos que estaremos aplicando para el caso de nuestro experimento.

¹³⁹ Hinkelmann, K. (1994), *Design and Analysis of Experiments*, McMillan, EUA, pag 65-70.

¹⁴⁰ Hicks, C. (1994), *Fundamental Concepts in the Design of Experiments*, Holt and Winston Ed, New York, pag 70-96.

8.4 Terminología empleada en el diseño de experimentos.

A continuación enumeramos la terminología que estaremos empleando para explicar el diseño experimental que aplicaremos para nuestro problema de investigación:

1. **Unidad experimental:** son los objetos, o instancias numéricas en nuestro caso, sobre los que se van a aplicar los experimentos.
2. **Variable de interés o respuesta:** es la variable que se desea estudiar. En nuestro caso se referirá al % de optimalidad alcanzado.
3. **Factor:** son las variables independientes (o variables controladas) que se presume pueden influir en la variabilidad de la variable de interés. Los factores son los elementos sobre los cuales el investigador puede hacer las manipulaciones en el experimento.
4. **Factor tratamiento:** es un factor, del cual el investigador, se interesa en conocer su influencia en la respuesta de la variable de interés. En nuestro caso, existen dos factores para los cuales tenemos el interés en el experimento. El primero se refiere al tipo de algoritmo aplicado y el segundo es el tiempo computacional ejercido en cada tratamiento.
5. **Factor bloque:** es un factor en el que no se está interesado en conocer su influencia en la respuesta pero se supone que ésta existe y se quiere controlar para disminuir la variabilidad residual. En nuestro caso hablaríamos acerca de los parámetros técnicos de los algoritmos genéticos que estarían siendo fijados como constantes durante todo el experimento.
6. **Niveles:** estos representan los valores posibles que pueden existir para cada uno de los factores involucrados. Estos niveles, como más adelante se explicará, pueden ser seleccionados por el investigador o elegidos aleatoriamente según sea conveniente en cada caso.
7. **Tratamiento:** es una combinación específica en la cual se tiene valorizado el nivel correspondiente a cada uno de los factores que están siendo estudiados en el experimento. En un diseño con un factor único, estos son los distintos niveles del factor. Por otro lado, en un diseño con varios factores son las distintas combinaciones de niveles de cada uno de los factores entre sí. Por lo

tanto, un tratamiento es en particular una condición experimental que se desea comparar. En nuestro caso, un tratamiento sería la aplicación de una versión particular de los algoritmos genéticos que vayan a ser confrontados (1er factor), sobre la base de un límite de tiempo computacional también previamente especificado (2do factor).

8. Observación experimental: es cada medición obtenida como respuesta de la variable de interés.
9. Tamaño del Experimento: es el número total de observaciones recogidas en el diseño.
10. Interacción de factores: existe interacción entre dos factores $F1$ y $F2$, si el efecto de algún nivel de $F1$ se modifica al cambiar el nivel en $F2$.
11. Ortogonalidad de factores: dos factores $F1$ y $F2$ con I y J niveles respectivamente, son ortogonales si en cada nivel I de $F1$ el número de observaciones de los J niveles de $F2$ están en las mismas proporciones¹⁴¹. En nuestro caso los dos factores, tipo de algoritmo y tiempo computacional, son ortogonales, ya que como se revisará más adelante, existe el mismo número de niveles para cada uno de los factores, es decir, 4 tipos de algoritmos para $F1$ y 4 intervalos de tiempo computacional para el factor $F2$.
12. Diseño equilibrado o balanceado: es aquél diseño en el que todos los tratamientos son asignados a un número igual de unidades experimentales. En nuestro caso, tomando en cuenta que tenemos presentes dos factores ortogonales con 4 niveles cada uno, entonces hablamos de que existirán un total de 16 tratamientos a ser aplicados en el experimento. Luego entonces, a cada uno de los 16 tratamientos le aplicaremos un total de 40 instancias de prueba. Por tanto, podemos decir que nuestro experimento posee un diseño equilibrado.

¹⁴¹ Toutengburg, H. (1995), *Experimental Design and Model Choice*, Physica-Verlag, EUA, pag 125-143.

8.5 Metodología empleada para la planificación del experimento.

A continuación se enumeran las etapas que fueron aplicadas para la planificación de nuestro experimento. Dicha metodología es extraída y resumida de la obra de Weber¹⁴².

1. Definir los objetivos del experimento: debemos definir cuáles son concretamente las preguntas a las que debe dar respuesta en el experimento. Tratar de abordar problemas colaterales puede complicar innecesariamente el experimento. Desde este paso, resulta útil esquematizar el tipo de conclusiones que se espera obtener para poder ir anticipando el modelo o matriz del diseño experimental. Los objetivos más comunes por los cuales se busca hacer un experimento son:
 - a. Determinar las principales causas de variación en la respuesta de la variable de interés.
 - b. Encontrar las condiciones experimentales con las que se consigue un valor extremo en la variable de interés.
 - c. Establecer un modelo estadístico que permita hacer predicciones de respuestas futuras sobre la base de la variable de interés a partir de los parámetros manipulados en el experimento.
 - d. Comparar las respuestas obtenidas en la variable de interés al hacer variar los parámetros (variables controladas) aplicados en cada tratamiento. Este último caso es lo que buscamos cubrir en nuestro proyecto de investigación. En nuestro caso, lo que buscamos determinar es la relación que existe entre el porcentaje de optimalidad alcanzado en los experimentos al aplicar diferentes tipos de instrumentos (algoritmos a ser confrontados) así como también al hacer variar la cantidad de tiempo computacional permitido en cada instrumento. Dicho de otro modo, buscamos conocer la influencia del algoritmo utilizado así como del tiempo de ejecución computacional respecto a la variable de interés (porcentaje de optimalidad).

2. Identificar todas las posibles fuentes de variación:
 - a. Variables controladas (o factores) de tratamiento: un factor es aquella variable fuente cuyo efecto sobre la variable respuesta es de particular interés para el investigador. En nuestro caso existen 2 factores: el tipo de algoritmo experimentado y el tiempo computacional aplicado en cada experimento.
 - b. Niveles de los factores: son los tipos o grados específicos del factor que se tendrán en cuenta en la realización del experimento. Los niveles de estos factores, son básicamente los

¹⁴² Weber, D. (2000), *A First Course in the Design of Experiments*, CRC Press, EUA, pag 20-78.

grados o valores para los cuales se planea experimentar su efecto en la variable de interés. En nuestro caso, podemos hablar de 4 niveles para el factor del algoritmo computacional y también de 4 niveles para el factor del tiempo computacional. Nótese que, en el diseño de experimentos, es común que los factores cuantitativos sean tratados como cualitativos. Por ejemplo en nuestro caso, los 4 niveles aplicados para el factor del tiempo computacional están codificados de acuerdo a los 4 intervalos de medición previamente diseñados por el investigador. Por lo general, un factor no suele tener más de cuatro niveles¹⁴³.

- c. Factores de ruido: son aquellas fuentes de variabilidad que no son del interés directo para el investigador, pero que se contemplan en el diseño experimental a fin de reducir la variabilidad no planificada. Lo anterior aplicado a nuestro experimento, podemos mencionar que los parámetros técnicos de los algoritmos genéticos, no resultan ser del interés directo para ser experimentados en nuestro proyecto de investigación. Esto significa en nuestro caso que por ejemplo, el tamaño de la población, el porcentaje de cruzamiento, el de mutación y el de elitismo entre otros, son fijados a un mismo valor en todos los algoritmos que van a ser confrontados con la finalidad de eliminar cualquier fuente de variabilidad proveniente de estos factores que no son del interés en este caso.
- d. Características descriptivas de las unidades experimentales: este aspecto aplicado a nuestro caso de investigación, se revisará más adelante en el apartado 8.10.

3. Definir una regla para asignar las unidades experimentales a los tratamientos: la regla de asignación especifica cuáles unidades experimentales se observarán bajo cada tratamiento. Las unidades experimentales son el “material” que sirve para evaluar la respuesta de la variable de interés y al que se le aplican los distintos niveles de los factores que estén siendo experimentados. Resulta crucial que las unidades experimentales sean representativas de la población sobre la que se han fijado los objetivos del estudio. Ahora bien, para el caso de nuestro experimento decíamos en el punto anterior que estamos trabajando con 2 factores. Así entonces, un “tratamiento” corresponde a cada una de las combinaciones de los “niveles” que estarán siendo combinados para cada uno de los 2 factores. Luego entonces una “observación” corresponde al resultado obtenido en la variable de interés de acuerdo a una condición específica y determinada por cada uno de los tratamientos que estén siendo aplicados. Las técnicas de asignación se revisarán a detalle en el apartado 8.6.

¹⁴³ Hunter, W. (1986), *Statistics for Experimenters*, John Wiley & Sons, EUA, pag 82.

4. Definir la variable de interés (ó respuesta): es importante precisar de antemano cuál es la variable respuesta y en qué unidades se medirá. Naturalmente, la respuesta está condicionada por los objetivos del experimento. Así por ejemplo en nuestro caso, la variable respuesta estará determinada en una métrica convenientemente establecida en minutos de tiempo computacional.

5. Diseñar el modelo o matriz del experimento: la matriz experimental que se defina, debe indicar la relación que se supone existe entre la respuesta de la variable de interés y las fuentes de variación (factores). En dicha matriz, la respuesta viene dada por una combinación de los factores que representan las principales fuentes de variación planificada. Los modelos de diseño de experimentos se pueden clasificar básicamente en dos tipos:
 - a. Modelo de efectos fijos: es aquel en el cual los niveles de los factores se encuentran predeterminados deliberadamente por el investigador. Este caso resulta apropiado cuando el interés se centra en comparar el efecto de la variable de interés, únicamente en niveles específicos en los factores. Este es precisamente nuestro caso de investigación.
 - b. Modelo de efectos aleatorios: este caso es aquél en el cual para un factor en particular, sólo se incluyen en el experimento una muestra aleatoria simple de todos los posibles niveles que puedan existir para dicho factor. Evidentemente se utilizan estos modelos cuando se tiene un número muy grande de niveles y no es razonable o posible trabajar con todos ellos.

6. Análisis estadístico de la información: en general el análisis de la información dependerá de:
 - a. Los objetivos del experimento determinados en el paso 1. Es necesario delimitar los objetivos del experimento antes de comenzar con el análisis estadístico. Si no se hace así, tratar que el experimento responda a otras cuestiones a posteriori puede ser imposible e incorrecto.
 - b. Del diseño de asignación en las unidades experimentales en los tratamientos visto en el paso 3.
 - c. Del modelo experimental seleccionado en el paso 5.

7. Determinar el tamaño muestral de las unidades a ser aplicadas en el experimento: esto consiste en determinar el número mínimo de observaciones que se deberán aplicar para alcanzar los objetivos del experimento. Existen varios métodos para hacer lo anterior, sin embargo lo que es común en todos estos métodos, es que se requiere el conocimiento del tamaño de la variabilidad no planificada (error experimental). Estimar lo anterior a priori no es fácil. Habitualmente se estima dicho elemento a partir de experimentos pilotos y/o en base a las experiencias previas en trabajos de investigación semejantes.

8.6 Estrategia para la asignación de las unidades experimentales en el experimento.

Para asignar las unidades experimentales a los tratamientos existen tres estrategias básicas las cuales son:

1. Aleatorización: consiste en aleatorizar en las unidades experimentales, todos los factores no controlados por el investigador que pueden influir en los resultados. En este caso, el investigador asigna las unidades experimentales a los tratamientos al azar. La única consideración es el número de observaciones que deberán ser tomadas en cuenta para cada tratamiento.
2. Técnica de bloques: consiste en dividir o particionar las unidades experimentales en grupos llamados bloques, de modo que las observaciones realizadas en cada bloque se realicen bajo condiciones experimentales similares. En este diseño, el investigador agrupa las unidades experimentales en bloques, a continuación determina la distribución de los tratamientos en cada bloque y, por último, asigna al azar las unidades experimentales a los tratamientos dentro de cada bloque. A esta técnica, también se le conoce como diseño “fila-columna”¹⁴⁴, y se caracteriza porque existen unidades experimentales en todas las celdas (intersecciones de fila y columna) en la matriz experimental.
3. Diseño factorial: se trata de una estrategia experimental que consiste en cruzar los niveles de todos los factores en todas las combinaciones posibles. Esta técnica permite detectar la existencia de interacción entre los diferentes factores. El diseño factorial resulta muy común encontrarlo en el estudio sobre mejora de procesos industriales (control de calidad). En estos casos es usual trabajar en problemas en los que hay muchos factores que pueden influir en la variable de interés. Los diseños factoriales 2^k son diseños en los que se trabaja con k factores, todos ellos con únicamente dos niveles (binarios). El número de observaciones que necesita un diseño factorial 2^k es muy grande ($n = 2^k$). Precisamente, debido al motivo anterior, la utilización de éste tipo de técnica tiene el gran inconveniente de necesitar un número elevado de observaciones.

Aunque en nuestro experimento no estamos interesados en determinar la existencia de interacción entre los diferentes factores, emplearemos ésta última técnica en virtud de que las 40 unidades experimentales (instancias aleatorias) serán aplicadas a cada uno de los 16 tratamientos que resultan al combinar exhaustivamente los dos factores que estamos considerando en nuestro caso.

¹⁴⁴ Montgomery, D. (1991), *Diseño y análisis de experimentos*, Grupo Editorial Ibero-Américo, México, pag 82-125.

8.7 Posibles errores en el diseño de experimentos.

La utilización de los modelos de diseño de experimentos se basa en la experimentación y en el análisis de los resultados que se obtienen. Por tal motivo resulta crucial la calidad de los datos. Lo anterior resulta relevante, sobre todo cuando los datos no provienen de un experimento controlado (del laboratorio), sino son extraídos de una fuente de opinión social¹⁴⁵. De lo anterior, resulta importante entonces tomar en cuenta que en el análisis estadístico de los datos pueden aparecer diferentes errores. Los más comunes son:

1. Inconsistencia de los datos: las condiciones en que se aplica cada tratamiento no son equiparables, lo cual ejerce una influencia no deseada sobre los datos recogidos en el experimento. Si derivamos este concepto para nuestro caso de investigación, sería como si a cada algoritmo que va a ser experimentado, lo ejecutáramos a través del uso de procesadores computacionales de velocidad (megahertz) diferentes.
2. Presencia de variables compuestas: puede ocurrir que al experimentar con un proceso existan dos o más variables las cuales en el diseño son establecidas como una sola. Lo anterior puede llevar a situaciones confusas. Por ejemplo, en un proceso existen dos variables $X1$ y $X2$ las cuales influyen en diferente grado en la respuesta de la variable de interés Y , pero si en el diseño del experimento se define solo una variable $X3$ que “intenta” agrupar el efecto de las otras dos, entonces no es posible distinguir si la influencia es debida a una u otra o a ambas variables.
3. Falta de detalle en los datos: se presenta si solo se dispone de datos de una variable (por ejemplo de $X1$ y no de $X2$), lo que puede llevar a pensar que la variable influyente es la $X1$ cuando, en realidad, la variable influyente es la $X2$ (variable oculta).
4. El rango de experimentación de los parámetros es limitado: si el rango de una de las variables controladas en el proceso es insuficiente, entonces no es posible saber su influencia fuera de ese rango y puede quedar oculta, al menos parcialmente, su relación con la variable de interés.

¹⁴⁵ Cochran, W. (1977), *Experimental Design*, John Wiley & Sons, New York, pag 56.

8.8 Definición de los parámetros a ser manipulados en el experimento.

Los parámetros de un experimento, resultan ser los elementos que están bajo el control del investigador. Naturalmente todos estos parámetros están incluidos como parte componente de cada una de las instancias que serán generadas aleatoriamente. En términos generales un parámetro a ser manipulado, necesariamente corresponde a cada una de las variables exógenas del modelo matemático ya presentado en el apartado 6.3. A continuación se enumeran los parámetros que serán manipulados en el experimento. La siguiente lista no pretende volver a enunciar los elementos ya cubiertos en la formulación matemática, pero si al menos complementar la definición de aquellos parámetros que contribuyan a precisar su significado.

1. Cliente: es toda aquella persona que requiere de un servicio de entrega o de recolección de un producto en particular.
2. Sitio o Localización: es todo aquel punto en el cual un cliente requiere que le sea entregado o recogido el producto.
3. Ventana de horario: son los intervalos de tiempo fijados por el cliente durante los cuales es permisible la atención de la entrega o la recolección del producto. Se define entonces que las ventanas de horario se cumplen si y solo si se llega después de la hora de apertura del cliente y antes de la hora de cierre.
4. Demanda: es la cantidad de producto que el cliente requiere le sea entregado o también la cantidad de producto que el cliente requiere le sea recogido.
5. Programa de horarios (“Time Schedule”, en inglés): es la lista de horarios a los cuales se tiene determinado que el vehículo debe llegar en cada parada.
6. Costo de traslado: es el costo, distancia o métrica de esfuerzo incurrido para ir de un cliente a otro dentro de la red.

8.9 Medición de las variables para la prueba experimental.

Las variables son los elementos para los cuales el investigador no tiene acceso a manipular su contenido durante el experimento sino tan sólo a estudiar su comportamiento. Las dos variables principales que se buscan medir en el experimento son:

1. El desempeño o velocidad computacional con la cual el Algoritmo Genético converge en una solución factible. La variable velocidad aparecerá factorizada en el matriz experimental la cual se expondrá en el apartado 8.13.
2. La calidad de la solución respecto a la solución registrada por el grupo control mediante el algoritmo de Ramificación y Corte basado en el BC. A esta medida de calidad le denominamos “porcentaje de optimalidad relativa”. El porcentaje de optimalidad relativa será la variable en donde se aplique la prueba estadística para la hipótesis (revisar apartado 8.13).

Es importante aquí precisar, que el presente experimento, no tiene el objetivo de incluir a “ w ” (*cantidad de nodos en cada instancia*) como parte de las variables a ser medidas en el experimento. Dicho de otro modo, no se busca establecer en el experimento, el como el valor de “ w ” (*donde: $100 \leq w \leq 120$*), influye en la variable “velocidad” y/o en la variable “calidad” para las instancias que sean aplicadas.

Así, mientras en el Algoritmo Genético se busca medir con qué rapidez se mejora una solución, con el Algoritmo de Ramificación y Corte basado en el BC se obtendrá el valor de referencia (o grupo control) para poder medir el desempeño del Algoritmo Genético propuesto. Lo anterior ratifica la importancia que la aplicación del Algoritmo de Ramificación y Corte basado en el BC tiene para nuestro experimento, ya que de otra manera no se tendría disponible una referencia cuantitativa para la prueba de la hipótesis de investigación.

De lo anterior, resulta interesante determinar la relación que guarda el tiempo de ejecución computacional versus el porcentaje de optimalidad alcanzado. Lo anterior servirá para establecer la relación costo-beneficio entre el tiempo que se tiene disponible para ejercer la toma de decisiones versus el nivel de calidad deseado. La comprensión de ésta relación resulta ser crítico cuando se trata de implementar algoritmos que tendrán que ofrecer soluciones en ambientes en tiempo real.

8.10 Estudio descriptivo de las instancias numéricas a ser utilizadas para la prueba experimental.

Aunque se expondrá este tema en detalle más adelante en el apartado 8.12, podemos en este momento establecer que será necesario generar instancias numéricas para cubrir el requerimiento de las unidades experimentales para poder desarrollar el experimento. Las instancias numéricas (o datasets), tendrán que ser aleatoriamente generadas y serán utilizadas como unidades experimentales para ser aplicadas tanto en el Algoritmo de Ramificación y Corte (BC) que servirá para la integración del grupo control, así como también en el método propuesto basado en un Algoritmo Genético.

Para lo anterior será necesario desarrollar un generador computacional que permita explotar la dimensionalidad cuantitativa de las instancias numéricas que serán utilizadas para el experimento. En la conformación de las instancias a ser obtenidas a través del generador aleatorio, es necesario que se consideren algunos aspectos que le darán validez al experimento. A continuación se enumeran los principales:

1. El cálculo aleatorio de las variables exógenas del problema, se asume que deban ser números positivos y enteros.
2. Se requiere que la desigualdad del triángulo se cumpla tanto para la matriz de costos C_{ij} , así como también para la matriz de tiempos T_{ij} , esto para cada uno de los nodos de la red. Esto matemáticamente se expresa como:

$$\{ : i,j,k \in V \ C_{ij} + C_{jk} \geq C_{ik} \text{ y también } T_{ij} + T_{jk} \geq T_{ik} \}$$

3. Lo antes expuesto para la conformación de la instancia, no debe ser visto como una limitante en el alcance práctico del planteamiento del problema. Después de todo resulta natural ya que para trasladarse de un nodo "i" a nodo "k" siempre resultará más económico hacerlo de manera directa que a través de la intermediación de un 3er. nodo "j" que solo sirva como medio de interconexión. De hecho esta propiedad aquí definida, será de utilidad ya que nos permitirá incluir en la variable T_{ij} , no solamente el tiempo que se ocupa para trasladarse del nodo "i" al nodo "j, sino también eventualmente el tiempo que se ocupa en el cliente "i" en ser atendido ya sea por (1) el proceso de entrega o, (2) el proceso de recolección.

4. Se requiere que las instancias generadas cubran una fase de pre-procesamiento aplicado con la finalidad de explotar la estructura del problema. El principal objetivo de la fase de pre-procesamiento es construir una formulación equivalente al problema original pero que logre conformar su espacio solución lo más ajustado posible.
5. Cuando nos referimos a tener un espacio solución ajustado, nos referimos a identificar aquellos cortes (o hiperplanos) que a su vez ofrezcan una cota inferior y superior al problema que logren acelerar el proceso de búsqueda de la solución óptima en el espacio solución. El reto de lo anterior consiste en que:
 - a. La formulación ajustada del problema debe asegurar no perder el acceso a la solución óptima del problema original a la vez que,
 - b. La solución óptima del problema ajustado debe mantenerse siendo también una solución óptima para el problema original.

En cuanto a la fase de pre-procesamiento a ser incluida en el generador de instancias aleatorias, podemos a su vez enumerar las siguientes estrategias:

1. Procedimiento para reducción de amplitud en las ventanas de horario (“Time Windows Reduction”, en inglés): este procedimiento tiene la finalidad de acotar o cerrar lo más posible la amplitud de las ventanas de horario definidas para cada cliente. Lo anterior se desarrolla de acuerdo a las restricciones de precedencia que resulten por consecuencia lógica de la revisión de las propias ventanas de horario. Este procedimiento resulta crucial ya que está probado científicamente que existe una dificultad intrínseca en relación al tamaño de la ventana de horario. Entre más amplia sea la ventana de horario, más grande es el espacio solución a tener que ser explorado y por tanto más difícil resulta el problema en ser resuelto¹⁴⁶.
2. Procedimiento para construcción de red (“Network Construction”, en inglés): busca obtener ventaja de los resultados obtenidos en el procedimiento anterior de reducción de las ventanas de horario. El objetivo es conformar una red que incluya solo aquellos arcos que resulten factibles de acuerdo a las ventanas de horario definidas en cada uno de los nodos, así como también de acuerdo a las relaciones de precedencia y a las restricciones de capacidad. La explotación de la estructura lógica anterior, tiene el objetivo de lograr formular una red mucho más esbelta, ya que logra identificar una gran cantidad de arcos que pueden ser omitidos del planteamiento del modelo matemático lo cual redundará en aspectos de eficiencia computacional.

¹⁴⁶ Desrosiers, J; Dumas, Y; Solomon, M; Soumis, F. (1995), *Time constrained routing and scheduling*, G.L. Nemhauser editor, *Handbooks in Operations Research and Management Science*, Holanda, pag 23-30.

8.11 Cálculo para el tamaño de la muestra.

El cálculo del tamaño de muestra “*n*” equivale en nuestro caso a la determinación de la cantidad de instancias aleatorias que serán utilizadas para el experimento. Si ya antes habíamos precisado que el rigor del experimento dependía de la calidad de los integrantes en el grupo control, entonces tenemos que precisar que la validez del experimento dependerá de que la cantidad de instancias que sean empleadas para el experimento, sea lo suficientemente representativo. En nuestro caso podemos afirmar que la cantidad de elementos en la población es infinito “ ∞ ”. Es comprensible entonces pensar que va a ser necesario determinar un tamaño de muestra suficiente para lograr alcanzar la validez del experimento.

Para el cálculo del tamaño de muestra existen diversos procedimientos. Uno que es muy simple de utilizar es a través de la aplicación de la siguiente formula:

$$n = \frac{z * p * q}{B^2}$$

De la formula anterior, a continuación explicamos en significado y el valor de cada variable:

z = Nivel de Confiabilidad = este valor esta asociado al área bajo la curva de distribución normal que se tiene a partir de cierta cantidad de desviaciones estándar “ σ ”.
Así por ejemplo si se quiere establecer en el experimento una confiabilidad del 95%, entonces quiere decir que solo nos vamos a permitir un 5% de probabilidad por arriba o abajo (prueba de dos colas) del criterio marcado.
Así pues, un 95% de confiabilidad equivale a 1.96 desviaciones estándar.

p = Probabilidad de éxito = este valor representa la probabilidad esperada de que un evento puntual del experimento resulte favorable. Este valor proviene de la empírica del investigador así como de las evidencias de éxito en experimentos provenientes de otras investigaciones.
Así entonces, si especificamos un 95%, quiere decir entonces que vamos a comprometer este valor para su verificación al final del experimento.

q = Probabilidad de fracaso = este valor representa la probabilidad estimada por el investigador de que un evento puntual en el experimento resulte fuera de lo deseado.

$B^2 = \text{Error experimental} =$ es el valor del error experimental que nos vamos a permitir.

Un error experimental menor al 5%, es una excelente medida.

Si entonces aplicamos los parámetros anteriores a la formula obtenemos entonces:

$$n = \frac{z * p * q}{B^2} = \frac{1.96 * (0.95) * (0.05)}{(0.05)^2} = \frac{0.0931}{0.0025} = 37.24 \approx 38$$

Como ya antes mencionábamos, existen otros métodos para calcular el tamaño de la muestra. Por ejemplo, si nosotros ya conociéramos, en base a un pre-muestreo, información adicional acerca de cómo se comportan ambas poblaciones (la del grupo control basada en el BC y la del grupo perteneciente al Algoritmo Genético), pues entonces podríamos aplicar una metodología diferente. A continuación en la Figura 8.2, se muestra el cálculo para el tamaño de la muestra basado en el conocimiento de la información de ambas muestras.

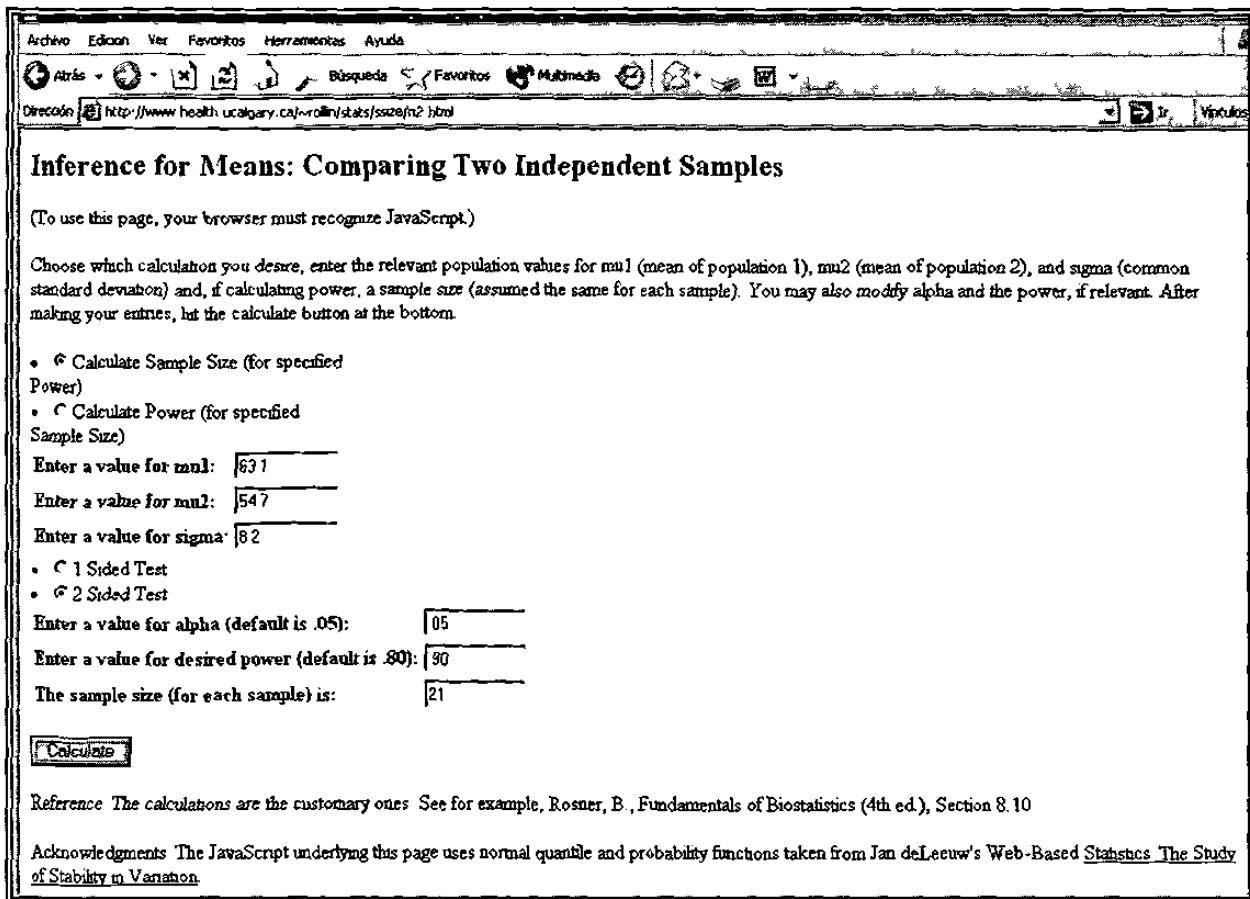


Figura 8.2 Cálculo del tamaño de la muestra a partir de información disponible de las muestras ¹⁴⁷.

¹⁴⁷ Disponible en <http://www.health.ucalgary.ca/>

Como puede verificarse en la figura anterior, es fácil darse cuenta que el tener información a partir de un pre-muestreo resulta útil, ya que logra que el tamaño de la muestra pueda disminuirse en favor del costo y del tiempo requerido para la ejecución del experimento. No obstante se puede verificar que el tamaño de la muestra, al disminuir de 38 a solo 21 unidades experimentales, también pierde la propiedad de normalidad que se obtiene al manejar muestras por arriba de 30 elementos.

El postulante establece manejar el experimento a través de un tamaño de muestra suficientemente grande como para cubrir el requerimiento de ambos procedimientos. Más aún, se emplearán un total de 40 unidades experimentales, con la tranquilidad y soporte que nos brinda el haber revisado los dos procedimientos alternos. Recordemos que la dimensionalidad (cantidad de nodos a ser tratados) para cada una de las 40 instancias, ya fue definido en el apartado 7.4 ($100 \leq v \leq 120$).

8.12 Planteamiento de la prueba experimental.

Definiremos para propósitos de nuestro diseño experimental, que el desempeño de un algoritmo está determinado por la efectividad del mismo. Ahora bien, cuando nos referimos en nuestro experimento a la efectividad del algoritmo, estamos hablando de las siguientes características a ser cubiertas:

1. Eficiencia en el tiempo de ejecución computacional: $\leq 5 \text{ minutos}$
2. Calidad en las soluciones arrojadas: *porcentaje de optimalidad* $> 90\%$
3. Cantidad de nodos a ser considerados en la instancia: $100 \leq v \leq 120$

Así, el mecanismo mediante el cual podemos probar la efectividad de un algoritmo propuesto para solucionar un problema determinado, pues es precisamente el diseño de la prueba experimental. El diseño de la prueba experimental estará determinado por el desempeño que muestre la aplicación del Algoritmo Genético a través de la manipulación de los parámetros del proceso de optimización. En nuestro caso, el diseño experimental está constituido por dos pruebas a ser aplicadas a cada instancia:

1. Prueba de Desempeño Computacional: consiste en medir el valor de la función objetivo que resulta después de una cantidad de tiempo fija y predeterminada de esfuerzo computacional.
2. Prueba de Calidad para la Exploración del Espacio Solución: consiste en medir que tan cerca se termina respecto a la solución óptima para un problema determinado después de una cantidad de tiempo fija de esfuerzo computacional. La medición de esta prueba se lleva a cabo mediante el cálculo de un valor denominado “Porcentaje de Optimalidad Relativa”, el cual se calcula de la siguiente manera:

$$\text{porcentaje de optimalidad} = 1 - \frac{(\text{solución alcanzada} - \text{cota inferior})}{\text{cota inferior}} \times 100\%$$

relativa

Por ejemplo, supongamos que para una instancia de prueba, la cota inferior obtenida mediante la aplicación del algoritmo BC resultase 529.12 (costo de la ruta). Luego supongamos que la solución alcanzada mediante la aplicación del algoritmo genético propuesto fuese 544.89, por tanto aplicando la fórmula del “porcentaje de optimalidad relativa” obtendríamos un valor de 97.02%. Es comprensible darse cuenta que, para el valor del “porcentaje de optimalidad relativa”, entre más se acerque al 100%, será indicativo de una mayor cercanía entre la solución alcanzada y la cota inferior (solución óptima). También

es fácil darse cuenta que este “porcentaje de optimalidad relativa”, nunca podrá ser mayor al 100%, ya que esto significaría que el algoritmo genético propuesto estaría encontrando un valor “mejor” que el referenciado en la cota inferior, lo cual equivale a invalidar al grupo control.

Ambas pruebas experimentales contribuyen a lograr los siguientes objetivos:

1. Calibrar los parámetros del Algoritmo Genético: es decir, establecer aquellos parámetros, o combinación de ellos, que hagan que el algoritmo alcance su máximo desempeño.
2. Contrastar la solución obtenida mediante el Algoritmo Genético con respecto a la solución del grupo control obtenida a través del algoritmo basado en el BC.

La metodología general para la medición de las variables a ser monitoreadas durante la prueba experimental es como sigue:

1. Se generan “n” instancias Aleatorias con valores uniformemente distribuidos. Estas instancias vendrán a conformar el conjunto de las unidades experimentales. El valor de “n” es precisamente el “Tamaño de la Muestra” experimental ya antes tratado.
2. Para cada una de las “n” instancias obtenidas en el punto 1, se ejecuta computacionalmente el algoritmo basado en el BC. Para cada ejecución computacional se aplicarán 4 mediciones en intervalos de tiempo predeterminados los cuales se definirán más adelante.
3. Para cada una de las 4 mediciones a realizar en el punto 2, se mide el valor de la función objetivo que haya sido alcanzada hasta este momento y ésta se registra como la mejor solución encontrada hasta cada momento correspondiente.
4. Si después de haber transcurrido la 4ta medición aún no finaliza el algoritmo BC su proceso de exploración para la solución óptima, entonces se deja continuar el algoritmo hasta su finalización con un límite de tiempo a 5 horas. Una vez que el algoritmo BC concluye, el valor reportado en la función objetivo, se establece como la cota inferior aplicada para cada unidad experimental que esté siendo tratada.
5. Las cotas inferiores registradas en el punto 4 integrarán el grupo control del experimento.

6. Para cada una de las “n” instancias generadas en el punto 1, de manera equivalente se ejecuta computacionalmente el Algoritmo Genético. Al igual que en el tratamiento del grupo control, para cada ejecución computacional se aplicarán 4 mediciones en los mismos intervalos de tiempo predeterminados.

7. El resultado de cada una de las corridas desarrolladas en el punto 6 vendrá a conformar el grupo de los tratamientos experimentales a ser comparadas con respecto al grupo control obtenido en el punto 5.

8.13 Implementación del experimento.

Conviene recordar que en el aparatado anterior se definieron dos pruebas para el diseño del experimento. Los resultados a ser obtenidos en la 1era prueba experimental, referente al desempeño computacional del algoritmo, serán tratados a través de una exposición descriptiva que permita visualizar el desempeño obtenido en cada una de las 4 mediciones a ser aplicadas para cada uno de los 4 algoritmos y a su vez para cada una de las unidades experimentales que vayan a ser probadas.

De este modo entonces, será la 2da. prueba experimental que ya antes denominamos como “Prueba de Calidad para la Exploración del Espacio Solución” la que será aplicada de manera específica para la prueba de la hipótesis de investigación. Recordaremos a su vez que la variable que en esta prueba vamos a medir es la que nombramos como “porcentaje de optimalidad relativa”. La medición de esta variable se hará mediante la aplicación de los 4 instrumentos experimentales. A continuación se enumera cada uno de estos:

1. Algoritmo de Ramificación y Corte (BC) como método de solución exacta. Este algoritmo es que el contribuirá a la conformación del grupo control que será tomado en cuenta en relación a los otros tres algoritmos.
2. Algoritmo Genético básico implementado mediante el motor (o software) de optimización “Evolver” © Versión 6.0 de Palisade Corporation.
3. Algoritmo Evolutivo básico implementado mediante el motor (o software) de optimización “Solver” © Versión 4.0 de Frontline System.
4. Algoritmo Evolutivo en versión avanzada (propuesta) e implementado mediante el motor (o software) de optimización “Solver” © Versión 4.0 de Frontline System.

El 1er. algoritmo (BC) correspondiente al grupo control, tendrá el propósito de establecer la referencia de optimalidad para cada una de las unidades experimentales (o instancias). El algoritmo BC se aplicará directamente sobre la base del problema original, es decir, en la estructura de la red sin compactar, a diferencia de cómo ocurrirá con el algoritmo propuesto #4. Los instrumentos #2, #3 y #4 serán entonces aplicados para verificar su desempeño en la búsqueda de la solución óptima que haya sido registrada mediante el algoritmo #1. El 4to. instrumento, corresponde precisamente a la propuesta algorítmica del proyecto de investigación. Esta última será entonces comparada contra las otras dos implementaciones básicas, la primera del algoritmo genético (2) y la segunda del algoritmo evolutivo (3).

En lo referente al algoritmo genético propuesto #4, su tiempo de ejecución computacional corresponde al tiempo total incurrido en las 6 fases expuestas ya anteriormente en el apartado 7.7 de la propuesta metodológica. Debemos recordar que en las 6 fases algorítmicas, tenemos 4 fases de pre-procesamiento una más de post-procesamiento, aparte por supuesto de la fase evolutiva.

Para cada uno de los 4 instrumentos antes descritos, la medición del “porcentaje de optimalidad relativa” será aplicado en 4 momentos sucesivos de tiempo durante la ejecución de cada experimento:

- a. Después de transcurrir el minuto #3 de ejecución computacional.
- b. Al minuto # 5.
- c. Al minuto # 8.
- d. Al minuto # 10.

El 1er. instrumento basado en el algoritmo BC, no está sujeto a un límite de tiempo, ya que su objetivo es obtener la solución óptima para cada unidad experimental. De lo anterior debemos suponer que las instancias que serán aplicadas deberán tener una complejidad matemática lo suficientemente intensa como para requerir el consumo de un mínimo de 10 minutos de esfuerzo computacional para alcanzar la solución óptima mediante el algoritmo BC.

A continuación se resumen las características dimensionales del diseño del experimento:

1. Cantidad de instrumentos a ser aplicados = 4 instrumentos / experimento
2. Cantidad de mediciones que serán aplicadas = 4 mediciones / instrumento
3. Cantidad de tiempo requerido en cada experimento = al menos 10 minutos / medición
4. Cantidad de unidades experimentales = 40 instancias

De lo anterior, tenemos entonces un experimento de 6400 minutos computacionales ($4 \times 4 \times 10 \times 40$). Lo anterior equivale a 107 horas (o casi 5 días naturales) de recurso computacional.

Concluidas las mediciones a ser aplicadas durante la prueba de calidad para la exploración del espacio solución ya antes descrito, se procede a calcular para cada unidad experimental el “porcentaje de optimalidad relativa”. Así, para cada una de las unidades experimentales (instancias) se tendrán entonces 16 valores cuantitativos correspondientes al cálculo de “porcentaje de optimalidad relativa”. La razón por la cual son 16, es debido a que estamos hablando de 4 instrumentos de medición (4 algoritmos) y además para cada uno de estos se aplicarán 4 mediciones en los intervalos de tiempo que ya antes explicamos. Sí

estamos hablando de 16 mediciones para cada una de las 40 unidades experimentales, entonces estamos hablando de un total de 640 mediciones, es decir 160 mediciones por cada instrumento algorítmico.

Así pues, para cada una de las 640 mediciones, la variable de “porcentaje de optimalidad relativa” será calculada a través del procedimiento que ya antes describimos. El siguiente paso del diseño del experimento consiste en estratificar las 640 mediciones a través de 4 grupos de 160 mediciones cada uno. Cada agrupación corresponde a las mediciones correspondientes a cada uno de los 4 instrumentos algorítmicos que están siendo comparados en el experimento. Luego a su vez, las 160 mediciones de cada uno de los 4 instrumentos, requiere volver a estratificarse en 4 grupos de 40 mediciones. En ésta ocasión, cada uno de los 4 grupos corresponde a las 4 mediciones que estamos aplicando para cada instrumento y a su vez a cada unidad experimental.

Así obtenemos 16 (4 x 4) grupos de 40 mediciones (unidades experimentales). Dadas las condiciones de normalidad que fueron procuradas al establecer el tamaño de la muestra en el apartado 8.11, el siguiente paso consiste en que para cada uno de los 16 grupos de 40 mediciones, se procede a calcular los siguientes parámetros estadísticos aplicados a la variable “porcentaje de optimalidad relativa”:

1. Media muestral (m).
2. Desviación estándar muestral (s).

La matriz del diseño del experimento, consiste en una tabla de 4 filas por 4 columnas. Cada una de las 4 filas de la matriz representa a cada una de las 4 mediciones que son hechas en cada unidad experimental. Por otro lado tenemos que cada una de las 4 columnas de la matriz representa a cada una de los 4 instrumentos que están siendo comparados. En la Tabla 8.1, se muestra lo anterior.

Tabla 8.1 Matriz resultante del diseño del experimento.

		Instrumentos de Medición a ser comparados			
		Algoritmo BB (Grupo Control)	Algoritmo Genético Básico (Evolver)	Algoritmo Evolutivo Básico (Frontline)	Algoritmo Evolutivo Avanzado (Frontline)
Mediciones a ser aplicadas después de cierto tiempo computacional	3er Minuto	($m11, s11$)	($m12, s12$)	($m13, s13$)	($m14, s14$)
	5to Minuto	($m21, s21$)	($m22, s22$)	($m23, s23$)	($m24, s24$)
	8avo Minuto	($m31, s31$)	($m32, s32$)	($m33, s33$)	($m34, s34$)
	10mo Minuto	($m41, s41$)	($m42, s42$)	($m43, s43$)	($m44, s44$)

A su vez, el significado de cada una de las celdas de la matriz corresponde al cálculo de la Media muestral (m_{ij}) así como a la Desviación estándar muestral (S_{ij}). Así por ejemplo (m_{ij}) sería la media muestral verificada para el instrumento algorítmico “j” en la medición del intervalo de tiempo computacional “i”.

La prueba estadística que utilizaremos para la comprobación de la hipótesis será a través de la aplicación de una Prueba “T”. Dicha Prueba “T” se aplicará de manera individual a cada uno de los 16 elementos de la matriz. La aplicación de la Prueba “T” en cada estadístico (m_{ij}, S_{ij}) calculará la probabilidad de que las mediciones obtenidas mediante la aplicación del instrumento algorítmico “j” y en el intervalo de tiempo computacional “i” hayan obtenido un “porcentaje de optimalidad relativa” mayor al 90%. Dicho de otro modo, buscamos calcular para cada uno de los 16 elementos de la matriz el valor de $P(x > 90\%)$, donde “P(x)” es el área bajo la curva para el estadístico de una distribución “T” de Student. De esta manera ahora el diseño de nuestra matriz se presenta a continuación:

Tabla 8.2 Matriz de probabilidades para la prueba de la hipótesis.

Prueba T de Student para la Variable "% de Optimalidad Relativa"

		Instrumentos de Medición a ser comparados			
		Algoritmo BB (Grupo Control)	Algoritmo Genético Básico (Evolver)	Algoritmo Evolutivo Básico (Frontline)	Algoritmo Evolutivo Avanzado (Frontline)
Mediciones a ser aplicadas después de cierto tiempo computacional	3er Minuto	$P(X11 > 90\%)$	$P(X12 > 90\%)$	$P(X13 > 90\%)$	$P(X14 > 90\%)$
	5to Minuto	$P(X21 > 90\%)$	$P(X22 > 90\%)$	$P(X23 > 90\%)$	$P(X24 > 90\%)$
	8avo Minuto	$P(X31 > 90\%)$	$P(X32 > 90\%)$	$P(X33 > 90\%)$	$P(X34 > 90\%)$
	10mo Minuto	$P(X41 > 90\%)$	$P(X42 > 90\%)$	$P(X43 > 90\%)$	$P(X44 > 90\%)$

Finalmente entonces, la prueba de la hipótesis de investigación puede ser resuelta verificando los porcentajes de probabilidad que se obtengan para cada uno de los elementos en la matriz. Así por ejemplo, si $P(X34 > 90\%) = 93\%$, nosotros entonces podríamos inferir con un 93% de certeza que el Algoritmo Evolutivo Avanzado (4to instrumento) puede mejorar el 90% de optimalidad relativa en un tiempo menor a los 8 minutos de ejecución computacional (3era medición).

Tomando en cuenta el ejemplo anterior así como la estructura de la Tabla 8.2, se puede comprobar que en ningún momento se está haciendo referencia a la dimensionalidad ($100 \leq w \leq 120$) de las instancias que serán aplicadas en el experimento. Lo anterior es debido a que este atributo no será considerado como una variable a ser medida en el experimento. No obstante, es oportuno mencionar que para las 40 instancias a ser aplicadas en el experimento, el grado de dimensionalidad de cada instancia estará determinado a partir de una variable aleatoria “ w ” uniformemente distribuida entre 100 y 120. Tomando en cuenta lo anterior es predecible que para cada valor de “ w ” entre 100 y 120, existan dos instancias generadas que logren de este modo cubrir las 40 instancias (tamaño de la muestra “ n ”) para el experimento.

En el siguiente capítulo 9 nos abocaremos a exponer los resultados del experimento aquí descrito.

9. Resultados.

9.1 Introducción.

Empezaremos por definir las condiciones computacionales que fueron aplicadas durante el experimento:

1. Sistema operativo "Windows XP ©"
2. Procesador computacional INTEL © Pentium 4 a 2.4 Ghz (Gigahertz).
3. Memoria de acceso inmediato de 128 Mhz (Megahertz).

Los 4 instrumentos de medición fueron implementados computacionalmente. Las instancias aplicadas en el experimento fueron desarrolladas a partir de redes arriba de 100 nodos ($100 \leq w \leq 120$). A este respecto es importante recordar que en el trabajo de Ascheuer, Fischetti, y Grotschel en el 2001, ellos exponen en base a sus experiencias computacionales que las instancias del TSP-TW en el límite de 50 hasta 70 nodos, pueden ser resueltas hasta la solución óptima por el algoritmo de Ramificación y Corte BC. De esta manera, nuestro algoritmo para el Grupo Control basado también en el BC, sobrepasó esta frontera antes especificada.

Por igual, recordemos que en los resultados obtenidos en el trabajo de Ascheuer, Fischetti, y Grotschel en el 2001 se concluye que, en general, las instancias para el problema de investigación resultan particularmente difíciles de resolver para aquellas en las que la cantidad de nodos activos que contengan restricciones de ventana de horario estén por arriba del 50%. A ese respecto, nuestro experimento aplicó el tratamiento a unidades experimentales todas por arriba del 70% con nodos activos. Nuevamente, el reto computacional para nuestro algoritmo de Ramificación y Corte BC consistió en ir más allá de este límite antes expuesto.

Finalmente y no menos importante es el tema de la amplitud de las ventanas de horario para los nodos activos. Recordemos que matemáticamente hablando, entre más cerradas sean las ventanas de horario más pequeño es el espacio solución a ser explorado y por tanto más fácil es el problema en resolver. En el experimento desarrollado, todas las unidades experimentales que fueron utilizadas, además de tener

al menos un 70% de nodos con ventanas de horario activas, también éstas tienen al menos un 75% de amplitud en sus ventanas. Esto quiere decir que las instancias estuvieron caracterizadas por ventanas de horario con un alto porcentaje de traslape lo cual hace más difícil el problema para ser resuelto en términos de su solución óptima.

De todo lo anterior expuesto podemos confirmar que la condición de complejidad matemática estuvo plenamente cubierta y por tanto la prueba aplicada a nuestro Algoritmo Genético avanzado (4to instrumento de medición) fue significativa. Los resultados obtenidos en el experimento fueron satisfactorios.

Previo al análisis estadístico de la hipótesis mediante la aplicación de la prueba “T” de Student, estableceremos la siguiente Tabla 9.1 para dar a conocer los resultados generales del experimento:

Tabla 9.1 Matriz de frecuencias para los 16 grupos experimentales.

Resultados: Matriz de Frecuencias $F(x > Opt\%)$

		Instrumentos de Medición a ser comparados					
		Algoritmo BB (Grupo Control)	Algoritmo Genético Básico (Evolver)	Algoritmo Evolutivo Básico (Frontline)	Algoritmo Evolutivo Avanzado (Frontline)		
					F(x > 90%)	F(x > 92.5%)	F(x > 95%)
Mediciones a ser aplicadas después de cierto tiempo computacional	3er Minuto	NA	15	13	40	33	18
	5to Minuto	2	19	16	40	38	22
	8avo Minuto	6	23	25	40	40	33
	10mo Minuto	12	26	32	40	40	38

9.2 Prueba de la hipótesis de investigación.

Finalmente se procede a aplicar la prueba de la hipótesis. Los métodos estadísticos disponibles para hacer análisis básicamente son de dos tipos:

1. Los métodos descriptivos.
2. Los métodos inferenciales

A su vez los métodos estadísticos inferenciales pueden ser aplicados para dos propósitos muy bien diferenciados:

1. Pruebas de Hipótesis.
2. Estimación de intervalos de confianza

En nuestro caso lo que aplica es la prueba de hipótesis y para ello existen diversos métodos estadísticos que pueden ser de utilidad. Básicamente existen aquellos que tienen que ver con (1) Estadística Paramétrica y (2) Estadística NO Paramétrica. La estadística paramétrica es más potente debido a los supuestos de los cuales se parte al momento de iniciar el análisis de la muestra. Dichos supuestos son tales como el tamaño de la muestra o también alternativamente la propiedad de normalidad intrínseca de la población. En nuestro caso y tomando en cuenta ambos aspectos recién mencionados, haremos uso de la Prueba estadística “T” de student.

La prueba estadística “T” de student, a diferencia de la prueba “Z”, es más rigurosa debido a que no presupone el conocimiento acerca del valor de la desviación estándar de la población. En sustitución utiliza la desviación estándar de la muestra. No obstante, lo anterior tiene un costo estadístico. Lo anterior se manifiesta a través de los “Grados de Libertad” (df en inglés) que el estadístico va a permitir en los datos de la muestra. La prueba “T” que va a ser aplicada para la prueba de nuestra hipótesis será a través de una prueba de diferencias de medias. Los supuestos de los cuales se parte para dicho análisis son los siguientes:

1. Las poblaciones de donde parten las instancias que serán aplicadas para los 4 instrumentos algorítmicos están normalmente distribuidas.
2. Las muestras u observaciones que serán medidas están apareadas.
3. La desviación estándar de las poblaciones es desconocida.
4. Las muestras que serán tomadas para cada tratamiento es de tamaño 40.

Esencialmente en toda prueba de hipótesis es necesario comparar el valor del estadístico calculado versus el valor crítico del estadístico. En nuestro caso, la formula para calcular el valor del estadístico "T" para una prueba de diferencia de medias es como sigue a continuación:

$$t = \frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{n}}}, \text{ donde: } \bar{D} = \frac{\sum_{i=1}^n D_i}{n}, \text{ y } S_D = \sqrt{\frac{\sum_{i=1}^n (D_i - \bar{D})^2}{n-1}}$$

A continuación en la Tabla 9.2 (a), se muestran los valores calculados del estadístico "T".

Tabla 9.2 (a) Matriz de valores calculados para el estadístico "T".

		Instrumentos de Medición a ser comparados					
		Algoritmo BB (Grupo Control)	Algoritmo Genético Básico (Evolver)	Algoritmo Evolutivo Básico (Frontline)	Algoritmo Evolutivo Avanzado (Frontline)		
					P(x > 90%)	P(x > 92.5%)	P(x > 95%)
Mediciones a ser aplicadas después de cierto tiempo computacional	3er Minuto	NA	-0.4039	-0.5581	2.4261	1.0688	-0.0911
	5to Minuto	-2.4260	-0.1159	-0.3068	3.3130	1.5392	0.1111
	8avo Minuto	-1.2800	0.1616	0.3173	4.8510	4.1050	0.8300
	10mo Minuto	-0.7001	0.4001	0.9030	6.2980	5.5770	1.3277

Los valores calculados de "T" son los que se comparan contra los valores críticos de "T" a efectos de determinar la aceptación o el rechazo de la hipótesis nula. Por regla general tenemos que:

*Si $T_{Calc} \geq T_{Crítico}$ entonces Se acepta H_0
en caso contrario
Se acepta H_a*

Habitualmente el valor de $T_{Crítico}$ se puede obtener a través del uso de tablas estadísticas o alternativamente mediante software computacional. Invariablemente, se requiere considerar dos elementos. El primer elemento es el nivel del error permitido α (alfa) y el segundo elemento son los grado de libertad de la prueba (df). En cuanto al valor de α (alfa), este típicamente se considera sobre

la base de un 5% a un 1% de error. En otras palabras, si α (*alfa*) es el nivel del error permitido, entonces $1 - \alpha$ equivale al nivel de confiabilidad buscado en la prueba, o sea de un 95% a un 99% de confiabilidad. Por otro lado, la definición de los grados de libertad (*df*) se define de acuerdo al tipo de prueba que se va a implementar (cantidad de muestras involucradas), así como también, de acuerdo al tamaño de la muestra(s). Recordemos que en nuestro caso particular, la prueba a implementar es:

$$H_0: \mu_{\text{Grupo experimental}} \leq 90\% (\mu_{\text{Grupo Control}})$$

$$H_a: \mu_{\text{Grupo experimental}} > 90\% (\mu_{\text{Grupo Control}})$$

Como se puede apreciar en la formulación anterior, se trata de una prueba estadística de una cola y con una sola muestra a comparar con respecto al grupo control. Por tal motivo podemos entonces definir que los grados de libertad en nuestro caso se define como: $df = n - 1$, donde $n = \text{tamaño de la muestra}$. El uso de los grados de libertad (*df*), es lo que hace diferenciar la prueba “T” de la prueba “Z”. La distribución estadística en una prueba “Z” está representada por una distribución normal única. En cambio en la prueba “T”, la distribución estadística no es única ya que se trata en realidad de una familia de distribuciones cuasi-normales, en donde, cada una de estas distribuciones está diferenciada por los grados de libertad (*df*) que definen la prueba a realizar.

A medida que los grados de libertad de la muestra disminuyen, entonces las áreas bajo la curva de distribución que están contenidas en las colas aumentan. En términos generales, podemos decir que los valores críticos de la prueba “T” van aumentando en similitud a los valores críticos de la prueba “Z” en la medida en la cual el tamaño de la muestra crece por arriba de los “30” elementos. Para ejemplificar la similitud anterior, haremos uso de nuestra prueba de hipótesis una sola “cola”, en la cual las muestras experimentales son de tamaño “40”. Esto quiere decir que el grado de libertad (*df*) equivale a “39” en nuestro caso. A partir de lo anterior tendríamos los siguientes valores críticos:

	<u>$\alpha = 5\%$</u>	<u>$\alpha = 1\%$</u>
Valores Críticos de Z:	$df = \infty$ 1.645	2.33
Valores Críticos de T:	$df = 39$ 1.685	2.426

Con lo anterior podemos concluir que: a medida que disminuyen los elementos de la muestra, menor son los grados de libertad de la prueba y por ende, crecen los valores críticos de la prueba “T” que son requeridos para lograr el rechazo de la hipótesis nula.

A partir de los resultados expuestos en la Tabla 9.2 (a), en donde se expresaron los valores calculados del estadístico “T”, es posible entonces para cada elemento de la matriz, ir aceptando o rechazando la hipótesis de optimalidad según corresponda en cada caso. Tan solo hace falta comparar dichos valores calculados que aparecen en la tabla 9.2 (a) versus el valor crítico de “T” que nosotros aspiremos a probar. Si buscáramos un 95% de confiabilidad entonces la referencia mínima sería “1.685”. Si buscáramos un 99% de confiabilidad entonces la referencia mínima sería “2.426”.

En lugar de verificar tan solo la hipótesis a través de la comparación del valor calculado de “T” versus su valor crítico, alternativamente es posible para cada valor calculado de “T” establecer su correspondiente valor de significancia o confiabilidad. Esto quiere decir que aunque en algunos casos no sea posible alcanzar un nivel α (alfa) que sea suficiente para probar una hipótesis, si al menos es posible establecer su nivel de significancia (*P Valor*). A continuación, en la Tabla 9.2 (b) se muestran los porcentajes de probabilidad “P Valor” resultantes de la prueba “T” de student aplicados en el diseño del experimento para cada uno de los 16 grupos de la matriz. Estos valores de “P” corresponden uno a uno a los valores calculados del estadístico “T” que se reportaron en la Tabla 9.2 (a).

Tabla 9.2 (b) Matriz experimental de probabilidades “P Valor” para el Estadístico “T”, $P(x > Opt\%)$.

		Instrumentos de Medición a ser comparados					
		Algoritmo BB (Grupo Control)	Algoritmo Genético Básico (Evolver)	Algoritmo Evolutivo Básico (Frontline)	Algoritmo Evolutivo Avanzado (Frontline)		
					P(x > 90%)	P(x > 92.5%)	P(x > 95%)
Mediciones a ser aplicadas después de cierto tiempo computacional	3er Minuto	NA	34%	29%	99%	85%	46%
	5to Minuto	<1%	45%	38%	100%	93%	54%
	8avo Minuto	10%	56%	62%	100%	100%	79%
	10mo Minuto	24%	65%	81%	100%	100%	90%

A partir de los resultados numéricos expuestos en la tabla 9.2 (b), procedemos a continuación a desarrollar una breve explicación de los mismos:

1. El Algoritmo de Ramificación y Corte BC, como era de esperar, invariablemente obtiene sino la solución óptima, si al menos una mejor solución que nuestro Algoritmo Genético. Recordemos que para el caso del Algoritmo de Ramificación y Corte BC, aunque será utilizado como grupo control, no es requerido en el objetivo del proyecto de investigación comprobar su optimalidad.
2. Los tiempos computacionales registrados para el Algoritmo de Ramificación y Corte BC, oscilan entre los 20 minutos y las 5 horas dependiendo de cada unidad experimental que haya sido tratada.
3. Es interesante puntualizar que nuestra implementación algorítmica resulta estar mucho más influenciada por la cantidad de nodos con ventanas de horario activas que por la amplitud de las ventanas de horario. Esta propiedad resulta entonces útil por la característica abordada en el proyecto de investigación al hablar del término “Ventanas de tiempo negociables”.
4. En cuanto al experimento con las implementaciones basadas en Algoritmos Genéticos, debemos iniciar precisando que todos los parámetros (Tamaño de la población, Método de Selección, Operador de Cruce, porcentaje de Mutación y porcentaje de Elitismo) de los Algoritmos Genéticos correspondientes a los instrumentos 2, 3 y el 4, fueron fijados al mismo valor con la finalidad de hacer comparables sus desempeños.
5. Los valores que fueron fijados para los parámetros de los Algoritmos Genéticos provienen de la empírica del investigador así como también de los trabajos de investigación ya revisados en la bibliografía. El resumen de los valores aplicados a dichos parámetros puede ser revisado en la figura 7.1, la cual se expone en el apartado 7.7 correspondiente a la propuesta metodológica.
6. La implementación de nuestro Algoritmo Genético avanzado (instrumento # 4), logra consistentemente obtener soluciones satisfactorias (*arriba del 90% de optimalidad*) y en tiempos de ejecución computacional razonables ($\leq 3 t \leq 5$ minutos). Es importante agregar que estadísticamente hablando, las mediciones anteriores tienden a mejorar al hablar de sus promedios por razones derivadas del teorema del límite central.
7. Para el caso de los Algoritmos Genéticos Básicos (instrumento # 2 y # 3), los tiempos de convergencia registrados (tiempo computacional después del cual ya no hay mejora en la solución) son más o menos parecidos a los observados en el instrumento # 4. No obstante los porcentajes de optimalidad medidos resultan ser significativamente inferiores ya que estos nunca sobrepasan del 90%. Más bien tienden a oscilar entre el 80% y el 88% de optimalidad.

8. Es importante destacar en términos generales que aunque el instrumento # 2 (Algoritmo Genético de reproducción de estado estable) “evoluciona” mucho más rápido que el instrumento # 3 (Algoritmo Genético de reproducción generacional), el instrumento # 3 ofrece al final mejores porcentajes de optimalidad que el # 2.
9. De lo anterior podemos precisar que, aunque en los primeros 3 minutos de ejecución computacional el instrumento # 2 ofrece mejor desempeño que el instrumento # 3, en el experimento se observa que en el primer caso el algoritmo se “atasca” en un óptimo matemático local resultado de una degeneración prematura en la población de organismos.
10. La situación anterior es aprovechada por el instrumento # 3, el cual aunque empieza en las primeras generaciones muy lento, continúa mejorando la población de organismos y aprovechando el tiempo perdido por el instrumento # 2 durante la fase terminal de degeneración en el óptimo local.
11. Tomando en cuenta que deliberadamente todos los parámetros ya antes expuestos de los 3 Algoritmos Genéticos fueron fijados a un mismo valor, entonces la diferencia observada en el desempeño de estos 3 instrumentos puede ser inducida de la siguiente manera:
 - a. El instrumento # 3 obtiene mejores soluciones que el instrumento # 2 debido a la propiedad generacional en su método de reproducción.
 - b. El instrumento # 4 (propuesto) obtiene mejores soluciones que los otros dos debido a que aprovecha del instrumento # 3 la propiedad generacional de reproducción ya antes expuesta y además debido a la explotación en la estructura del problema que resulta ser el fundamento de la investigación doctoral en la que se aplica un algoritmo genético para resolver un problema muy particular.
12. Se aplicaron de manera complementaria algunas instancias “sencillas” para observar el desempeño del Algoritmo de Ramificación y Corte BC. Las instancias sencillas a las que nos estamos refiriendo podemos describirlas de manera general como aquellas por debajo de los 70 nodos y con menos del 60% de nodos activos en sus ventanas de horario. El resultado verificado en estos experimentos es un desempeño extraordinariamente rápido y eficiente del algoritmo. Para este tipo de instancias, se ocuparon tiempos computacionales por debajo de los 3 minutos.

9.3 Análisis costo-beneficio para la aplicación del Algoritmo Genético propuesto.

Las variables empleadas en el diseño experimental antes expuesto, tiempo computacional y % de optimalidad, han sido ampliamente promovidas para su uso en proyectos de investigación previos¹⁴⁸. Dicho diseño experimental, tuvo el objetivo de comprobar el desempeño que nuestra propuesta algorítmica ofrece con respecto al grupo control determinado por un método de solución exacta.

Las justificaciones del proyecto de investigación, expuestas en el capítulo 7, apuntan claramente a identificar la necesidad de herramientas de optimización para apoyar a la función de la logística de ruteo. No obstante lo anterior, es claramente verificable que el objetivo del proyecto de investigación nunca fue establecer una métrica para analizar el costo-beneficio que nuestra propuesta ofrece como ventaja económica al ser implementada en una problemática real de negocio.

El tratamiento del análisis costo-beneficio de la propuesta sobre la base de una justificación económica, puede ser resuelta a través de comparar el beneficio obtenido, versus el costo económico que tuviera que ser erogado para obtener dicho beneficio. Así entonces, en un sentido práctico diríamos que la propuesta es económicamente viable, si el beneficio económico obtenido por la propuesta algorítmica es “razonablemente” mayor que el costo requerido para su implementación.

El investigador desarrollará a continuación un análisis costo-beneficio, sobre la base de una serie de supuestos cuantitativos que han sido definidos, deliberadamente en una forma conservadora, con la finalidad de lograr minimizar el error que pueda existir en las conclusiones que de aquí se generen. Dichos supuestos provienen de la empírica del investigador, los cuales serán expuestos numéricamente a lo largo de la exposición de las Tablas 9.3, 9.4 y 9.5 a continuación.

Comenzamos con la Tabla 9.3 en la cual se exponen los elementos que el investigador está proponiendo para la implementación de la propuesta de solución. En dicha tabla, cada elemento involucrado se acompaña con una estimación típica del costo en cada caso correspondiente. Los costos están estipulados pensando en la implementación de la propuesta en una empresa con operaciones en el área metropolitana de ciudad de Monterrey (AMM).

¹⁴⁸ Cheng, R. (1997), *Genetic Algorithms and Engineering Design*, John Wiley Interscience, Boston Ma, EUA, pag 43-57.

Tabla 9.3 Costos relacionados a la propuesta de solución basado en la aplicación de un Algoritmo Genético para solucionar el problema de ruteo SPDP-sTW en una empresa del AMM.

	COMPONENTES	INVERSION	DEPRECIACION / GASTO ANUAL	Depre- ciación
Costo de Oportunidad: Profesionalista competente en la aplicación del OR/MS	Profesionalización de la función de ruteo Logístico	1 Coordinador	\$240,000	
	Prestaciones		\$120,000	
	Habilitadores: Seminarios y Talleres		\$10,000	
SUB-TOTAL		67%	\$370,000	
Costo Tecnológico: Hardware, Software, Investigación y Desarrollo	Computadora Pentium IV (>=2.0 GHz)	\$25,000	\$6,250	4 Años
	Mapas Digitales (calles de la Cd. Monterrey)	\$20,000	\$10,000	2 Años
	Licenciamiento x uso de Algoritmo Genético	\$200,000	\$66,667	3 Años
	Soporte Técnico y Capacitación		\$100,000	
SUB-TOTAL		33%	\$182,917	
COSTO TOTAL DEL PROYECTO			\$552,917	

Como se puede apreciar en la tabla anterior, el costo relacionado al desarrollo de un profesionista que logre aplicar la propuesta algorítmica para la obtención de rutas de distribución mejoradas, ocupa el 67% del total del costo del proyecto. El desarrollo de dichas competencias en el profesionista no es trivial, ya que como se verá más adelante en el apartado 9.4, la configuración apropiada de los parámetros en el algoritmo genético será crucial para mejorar la relación costo-beneficio del proyecto. Hablar de que las dos terceras partes (67%) del costo del proyecto, recaen en el ámbito de la aplicación del algoritmo genético por parte de un profesionista con características específicas, justifican en gran medida la importancia de haber incluido el factor educativo y cultural del elemento humano en los antecedentes del proyecto de tesis.

Una vez cubierto el componente del costo en la justificación económica, entonces lo que procede a continuación, es dar tratamiento al componente del beneficio esperado. El cálculo del beneficio esperado puede ser atendido en dos aspectos principales los cuales son:

1. Eficiencia en el uso de las capacidades de transporte.
2. Efectividad en los horarios de servicio comprometidos con el cliente.

a. Eficiencia en el uso de las capacidades de transporte:

La función objetivo del problema SPDP-sTW, está determinado por la minimización del costo de distribución en el ruteo propuesto. De esta manera es fácil entender que, al rutear óptimamente un vehículo de transporte en términos del costo de distribución, entonces puede existir la factibilidad en tiempo y en capacidad de carga, de que una mayor cantidad de clientes puedan ser cubiertos con el mismo recurso de transportación. A su vez, esto quiere decir que al haber más clientes siendo atendidos por cada vehículo, entonces los vehículos se vuelven más productivos. Si los vehículos son más productivos, entonces el administrador de un negocio puede optar por hacer una de dos alternativas:

- a. Aprovechar el exceso de capacidad de transporte para incrementar el nivel de cobertura en la distribución y ventas del negocio.
- b. Optimizar (disminuir) la flota de equipo de transporte.

Ahora bien, los beneficios provenientes de la primera alternativa son difíciles de cuantificar, ya que en la realidad, no existiría evidencia suficiente acerca de cuantos clientes, y por consiguiente qué volumen de venta adicional podría llegar a materializarse sobre la base de solo un potencial excedente en la capacidad de transportación. Dicho de otro modo, sería más conservador e ineludible, disponer de un recurso de transporte solo hasta que los requerimientos del mercado así lo exigieran. Sobre el razonamiento anterior aplicaremos entonces la segunda alternativa, en la cual, los beneficios resultan mucho más objetivos y tangibles. A continuación en la Tabla 9.4, se muestran los costos anuales que típicamente podrían ser evitables por cada ruta que fuese disminuida, como consecuencia de la optimización en la capacidad de transporte vía un ruteo más eficiente.

Tabla 9.4 Ahorro potencial por ruta a ser disminuida de la flota vía eficiencia en el ruteo.

COMPONENTES DEL COSTO DE UNA RUTA	INVERSION	Depre- ciación	Depreciación / Gasto Anual	% Aplica- ción	Costo Anual
Vehículo de Transporte (Camión + Carrocería)	\$500,000	10 Años	\$50,000	100%	\$50,000
Mantenimiento			\$20,000	50%	\$10,000
Combustible			\$30,000	20%	\$6,000
Personal Operador	1 Chofer + 1		\$240,000	100%	\$240,000
Prestaciones	Ayudante		\$120,000	100%	\$120,000
TOTAL: GASTO DE OPERACIÓN ANUAL POR RUTA DE TRANSPORTE					\$426,000

b. Efectividad en los horarios de servicio comprometidos con el cliente:

El tratamiento de las ventanas de horario a ser consideradas para el servicio de entrega y recolección a los clientes, no forma parte de la formulación del costo económico en la función objetivo del problema SPDP-sTW. No obstante, esta consideración aparece formulada como una restricción dentro del modelo matemático del problema (revisar apartado 6.2). Lo anterior significa que, cualquier solución de ruteo que proponga el algoritmo genético, cumplirá rigurosamente el aspecto de los horarios de servicio para la totalidad de los “*n*” clientes que estén siendo tratados en la instancia del problema.

De acuerdo a lo anterior y dejando de lado cualquier circunstancia estocástica que pudiera presentarse en la realidad, para las soluciones generadas por el algoritmo genético podemos afirmar que, la probabilidad de llegar fuera del horario estipulado es del 0%. De acuerdo a la caracterización de las instancias que han sido atendidas en el presente proyecto de investigación (revisar apartado 9.1), tenemos que la cantidad promedio de clientes por ruta es de 100 nodos, con al menos un 70% de nodos activos con ventanas de horario. Esto quiere decir que tenemos 70 clientes a los que debemos arribar dentro de un horario de atención específico. Entonces la pregunta que nos tendríamos que hacer sería:

¿Cual es la probabilidad de que un despachador de rutas en una empresa, sin herramientas de optimización, logre secuenciar un vehículo de transporte atinando a los horarios de llegada de 70 clientes que requieren considerar ventanas de horario?

Dar respuesta a lo anterior resulta arriesgado. Para ejemplificar esto, podemos partir de un escenario conservador en el cual, gracias a la experiencia del despachador, éste secuencie las llegadas a los clientes de tal manera que solo en 1 cliente de cada 100, no logre empatar el horario de llegada requerido. Así entonces, si la confiabilidad en el procesamiento de cada ventana de horario correspondiente a cada cliente fuese del 99%, entonces tendríamos que la probabilidad de que una ruta estuviese libre de errores sería igual a: $99\%^{70} = 49\%$. De lo anterior se puede concluir que, estimar un beneficio por disponer de una herramienta de optimización equivale a estimar el costo de oportunidad por dejar de vender producto por no llegar a tiempo a la cita. Ambos elementos siguen siendo subjetivos a la vista de los cuestionamientos. Para remedir lo anterior, lo que proponemos hacer es recurrir a los registros históricos de la compañía en búsqueda de las ventas perdidas que pudiéramos definir que existen por motivo del producto que las rutas devuelven.

Vamos a continuación a generar los siguientes supuestos para el análisis:

1. Suponemos que solo el 50% del producto devuelto sea imputable a las llegadas fuera de horario.
 2. Si de éste, solo el 50% pudiera ser resuelto a través de la implementación de un ruteo más eficaz, entonces tendríamos un potencial del 25% de reducción en el producto que se devuelve.
 3. Si el porcentaje de producto devuelto que una empresa actualmente tiene fuese del 10%, entonces a través de un ruteo más inteligente podríamos aumentar las ventas en un 2.5%.
 4. Si la contribución neta a utilidades (EBITDA) por la venta del producto fuese, digamos del 15%, entonces tendríamos un incremento del 0.4% en las utilidades netas de la compañía.
 5. Si partimos de una estructura de costos en la cual, el costo de operación de una ruta no debiera exceder del 20% de lo que vende, entonces tendríamos que una ruta, como la que se expuso en la tabla 9.4, debería estar vendiendo aprox. \$2 Millones al año (ó \$6,500 por día).
 6. Si una empresa opera con 50 rutas, entonces ésta facturaría \$100 Millones al año.
 7. Con lo anterior, tendríamos un beneficio marginal sobre utilidades de \$375,000 por año.
- Con la finalidad de resumir los supuestos comentados, en la Tabla 9.5 se exponen los factores que hacen sensibilizar el beneficio potencial a obtenerse mediante el uso del algoritmo propuesto.

Tabla 9.5 Análisis Costo-Beneficio derivado de la implementación del algoritmo propuesto.

Matriz Análisis Costo-Beneficio (Miles \$ / Año)			% de Incremento en Ventas (Mejoramiento en el Servicio)					
			1%	2%	3%	4%		
ECONOMIA DE ESCALA DE LA EMPRESA	50 Rutas	% de Rutas a Rutas a Optimizar	1%	-\$190	-\$40	\$110	\$260	
			2%	\$23	\$173	\$323	\$473	
			5%	\$662	\$812	\$962	\$1,112	
			10%	\$1,727	\$1,877	\$2,027	\$2,177	
	100 Rutas	% de Rutas a Rutas a Optimizar	1%	\$173	\$473	\$773	\$1,073	
			2%	\$599	\$899	\$1,199	\$1,499	
			5%	\$1,877	\$2,177	\$2,477	\$2,777	
			10%	\$4,007	\$4,307	\$4,607	\$4,907	
	200 Rutas	% de Rutas a Rutas a Optimizar	1%	\$899	\$1,499	\$2,099	\$2,699	
			2%	\$1,751	\$2,351	\$2,951	\$3,551	
			5%	\$4,307	\$4,907	\$5,507	\$6,107	
			10%	\$8,567	\$9,167	\$9,767	\$10,367	
	500 Rutas	% de Rutas a Rutas a Optimizar	1%	\$3,077	\$4,577	\$6,077	\$7,577	
			2%	\$5,207	\$6,707	\$8,207	\$9,707	
			5%	\$11,597	\$13,097	\$14,597	\$16,097	
			10%	\$22,247	\$23,747	\$25,247	\$26,747	
	Costo implementación del algoritmo (\$ x Año) =						\$552,917	
	Costo de Operación por Ruta (\$ x Año) =						\$426,000	21%
	Venta por Ruta (\$ x Año) =						\$2,000,000	
	Contribución Marginal de Ventas a Utilidad =						15%	

Como se puede apreciar en los escenarios de sensibilidad expuestos en la tabla anterior, existen básicamente 3 factores que están siendo combinados para efectos de estimar el costo-beneficio:

1. Economía de escala de la empresa (# de Rutas en operación)
2. Factor de Rutas a optimizar (optimización de la capacidad de transporte)
3. Factor de Ventas a incrementar (optimización en el servicio al cliente)

En cuanto al 1er factor, no queda duda de que entre mayor sea la escala de operación de la empresa mayor es el beneficio potencial. En cuanto al tratamiento de los otros dos factores, hasta este momento, todo está analizado numéricamente en el ámbito de los supuestos ya antes mencionados, lo cual no deja de ser inquietante, ya que:

“No es posible a priori, saber que tan lejos de la solución óptima puede estar cualquier ruta que actualmente esté operando en las empresas de distribución, sobre la base de una práctica operativa basada tan solo en la experiencia de las personas”.

Estimar la brecha (o el grado de mejoramiento), que pueda existir entre una ruta desarrollada en forma manual versus el óptimo matemático, requeriría la aplicación de una de las siguientes dos estrategias:

1. Desarrollar un estudio comparativo para calcular el efecto (supuesto beneficio) que se obtendría al sustituir las rutas que actualmente estén siendo operadas en una empresa por las soluciones que sean obtenidas mediante la propuesta basada en el algoritmo genético.
2. Desarrollar un estudio analítico que permita estimar el comportamiento que pudiera ser obtenido derivado del uso del algoritmo, a través del estudio de la superficie de respuesta del algoritmo genético.

En lo referente al uso de la primera estrategia, si quisiéramos generar resultados concluyentes respecto a la brecha de optimalidad, tendríamos que desarrollar un muestreo para varias empresas. La razón de lo anterior, vuelve a ser sobre la base de que las rutas que actualmente se estén generando en forma empírica por una persona, seguramente serían diferentes en función a la experiencia de otra persona. Por la razón anterior y de acuerdo al alcance ya antes definido en el objetivo de la tesis, el investigador aplicará la segunda estrategia para lograr generar datos más concluyentes al respecto.

No obstante lo anterior, el investigador ha desarrollado un breve pero significativo análisis comparativo acerca de la brecha de optimalidad que puede existir entre un conjunto de rutas que anteriormente estuvieron en operación con respecto a lo que el algoritmo genético propuso como ruta mejorada.

Dicho análisis se aplicó para una muestra de 20 rutas pertenecientes a una misma compañía de distribución de producto con operaciones en el área metropolitana de la ciudad de Monterrey (AMM). Los resultados comparativos se exponen a continuación en la Tabla 9.6

Tabla 9.6 Análisis comparativo de los beneficios obtenidos en términos de las distancias a recorrer entre una ruta de distribución actual versus la ruta optimizada propuesta por el algoritmo genético.

Ruta muestreada	Kms de Trayecto del almacén a la zona	Kms de Ruteo en la Zona		Beneficio imputable al Ruteo		Kms Totales de Transportación		
		Actual	Optimizado	Beneficio en Kms	Beneficio %	Actual	Optimizado	Beneficio %
1	11.6	6.6	5.0	1.6	24%	18.2	16.6	9%
2	10.9	9.1	6.9	2.2	24%	20.0	17.8	11%
3	10.1	23.1	21.1	2.0	9%	33.2	31.2	6%
4	20.6	12.4	7.9	4.5	36%	33.0	28.5	14%
5	28.8	9.4	8.0	1.4	15%	38.2	36.8	4%
6	28.4	11.4	8.9	2.5	22%	39.8	37.3	6%
7	22.5	7.3	5.4	1.9	26%	29.8	27.9	6%
8	14.8	19.4	16.2	3.2	16%	34.2	31.0	9%
9	7.4	10.2	9.2	1.0	10%	17.6	16.6	6%
10	8.6	9.5	5.8	3.7	39%	18.1	14.4	20%
11	22.0	8.3	6.1	2.2	27%	30.3	28.1	7%
12	19.0	14.3	11.4	2.9	20%	33.3	30.4	9%
13	12.0	12.2	8.3	3.9	32%	24.2	20.3	16%
14	13.6	6.8	4.7	2.1	31%	20.4	18.3	10%
15	20.8	7.1	5.2	1.9	27%	27.9	26.0	7%
16	20.5	6.6	4.3	2.3	35%	27.1	24.8	8%
17	8.7	21.3	14.8	6.5	31%	30.0	23.5	22%
18	8.4	7.3	4.7	2.6	36%	15.7	13.1	17%
19	9.6	4.6	3.7	0.9	20%	14.2	13.3	6%
20	14.3	9.9	7.6	2.3	23%	24.2	21.9	10%
Prom	15.6	10.8	8.3	2.6	25%	26.5	23.9	10%

Como puede apreciarse en la tabla anterior, existen dos perspectivas de beneficio que pueden ser útiles para efectos de cuantificar la cantidad de recursos de transporte que pueden ser optimizados (reducidos) de la flota. El primero está sobre la cantidad de kilómetros que recorren los vehículos durante su ruteo entre cada uno de los clientes que están incluidos dentro de la zona (columnas 5 y 6). La columna 5 hace referencia al beneficio absoluto medido en kilómetros, mientras que la columna 6 lo re-expresa en términos porcentuales. La segunda perspectiva del beneficio potencial esta determinada

en la última columna de la tabla, la cual hace referencia a la cantidad de kilómetros totales que recorre la ruta, incluyendo los trayectos desde que sale del almacén hasta que regresa al mismo. En ese sentido obtenemos un 10% de beneficio global en comparación con el beneficio imputable únicamente al ruteo entre cliente y cliente, el cual alcanza hasta un 25% de beneficio global para la muestra de las 20 rutas.

El análisis comparativo anterior, sea con el beneficio al 25% o al 10%, no deja duda de que conservadoramente podemos alcanzar al menos un 10% de beneficio en el tema de reducción de costos económicos debido a la optimización de la flota de transporte. Dicha argumentación apoya plenamente el análisis costo-beneficio expuesto ya antes en la Tabla 9.5. En dicha tabla puede ser verificado que se desarrollaron 4 escenarios de optimización económica, desde un 1% hasta un 10% de brecha de optimalidad.

Con todo lo expuesto hasta este momento podemos resumir lo siguiente en referencia al análisis costo-beneficio de la propuesta:

1. De acuerdo al análisis expuesto en la tabla 9.5, la conveniencia económica para la aplicación del algoritmo genético está influenciado básicamente por tres factores:
 - a. Economía de escala de la empresa (# de Rutas en operación = EE)
 - b. Factor de Rutas a optimizar (optimización de la capacidad de transporte = FR)
 - c. Factor de Ventas a incrementar (optimización en el servicio al cliente = FV)
2. Aún en empresas debajo de 50 rutas ($EE = 50$), resulta recomendable la aplicación del algoritmo. Su contribución se verifica significativamente para valores de $FR \geq 5\%$. Alternativamente también puede verificarse dicha contribución para valores de $FV \geq 3\%$ y de una manera mucho más importante cuando ambos factores se conjugan.
3. El 1er factor referido a la economía de la escala de la empresa (EE), es el que resulta más significativo al hablar de la conveniencia económica de la aplicación del algoritmo genético.
4. Del factor (EE), tenemos que para el 2do cuartil, es decir $50 \leq EE \leq 100$, el máximo beneficio económico sería de \$4.9 Millones y el promedio de \$2.1 Millones por año.
5. Para el 3er cuartil ($100 \leq EE \leq 200$), tenemos que el máximo beneficio económico sería de \$10.4 Millones y el promedio de \$4.8 Millones por año.
6. Finalmente, para el 4to cuartil ($200 \leq EE \leq 500$), tenemos que el máximo beneficio económico sería de \$26.7 Millones y el promedio de \$12.8 Millones por año.

Con la tranquilidad que nos brinda el haber aplicado un muestreo para calcular el beneficio esperado, ahora la investigación se dispone a aplicar la segunda estrategia para lograr estimar datos complementarios que permitan ser concluyente al respecto de la justificación económica de la propuesta algorítmica. No debe quedar duda de que esta metodología es completamente válida. Lo anterior se fundamenta, sobre la base siguiente:

“Cualquier ruta, presuntamente no óptima, que actualmente esté siendo operada en una empresa sin el uso del algoritmo propuesto, está incluida dentro del espacio solución descrito por el poliedro del problema SPDP-sTW. A su vez, la topología de dicho poliedro, puede ser analizada a través del desarrollo de un modelo de regresión multi-variado. Para dicho modelo tenemos que:

- 1. La variable dependiente corresponde a la función objetivo del problema SPDP-sTW.*
- 2. Las variables independientes (parámetros estadísticos) corresponden a los parámetros de operación del algoritmo genético.*

El análisis costo-beneficio expuesto ya anteriormente en la Tabla 9.5 y verificado mediante un muestreo en la Tabla 9.6, debiera ser suficiente para evidenciar una justificación económica para la aplicación del algoritmo genético propuesto, al menos en términos prácticos. No obstante, el investigador está convencido de que la brecha de optimalidad que existe entre una ruta generada mediante métodos empíricos versus aquella obtenida mediante métodos matemáticos, es un valor diferencial que sigue siendo desconocido y que por tanto es interesante generar algún esfuerzo al menos para predecirlo. Para ello el investigador desarrollará un análisis experimental el cual será cubierto en los apartados 9.4 y 9.5. Dicho análisis experimental permitirá estudiar de una manera mucho más científica dicho potencial de mejora.

Así entonces, a continuación en el apartado 9.4, se exponen las bases sobre las cuales haremos la experimentación para el estudio analítico de los parámetros de operación del algoritmo genético. Los parámetros de operación del algoritmo genético a ser utilizados para el desarrollo del modelo de predicción serán expuestos más adelante en el apartado 9.5.

9.4 Estudio analítico para los parámetros de operación del Algoritmo Genético.

El objetivo primordial de cualquier estudio analítico consiste en determinar los parámetros y sus valores que mejor ofrecen un desempeño de operación buscado. Lo anterior resulta contrastante al revisar la falta de investigación enfocada en utilizar dichas métricas para el desarrollo de análisis comparativos entre diferentes tipos de meta-heurísticas para resolver problemas de naturaleza combinatoria¹⁴⁹. De lo anterior, tan solo se disponen de pocas referencias a proyectos tales como el desarrollado en 1995 para resolver el famoso problema combinatorio de Steiner. En dicho proyecto de investigación, Esbensen desarrolla un análisis comparativo entre el desempeño de un algoritmo genético propuesto versus dos heurísticas conocidas y ambas en relación a la solución óptima global obtenida mediante un procedimiento de solución exacta basado en algoritmos de ramificación y corte.

Los resultados de Esbensen apuntan a concluir que las soluciones obtenidas mediante la aplicación de su algoritmo genético, logran consistentemente alcanzar soluciones muy cercanas al óptimo global para instancias de pequeña dimensionalidad. En cuanto a las instancias de media o alta dimensionalidad, Esbensen concluye que el algoritmo genético logra obtener mejores soluciones que las heurísticas comparadas en el análisis comparativo. La razón de lo anterior se concluye, es debido a que el esfuerzo computacional requerido por un algoritmo genético se incrementa a una velocidad menor que respecto el crecimiento del espacio solución del problema¹⁵⁰. Dicha propiedad no ocurre con el procedimiento de solución exacta de ramificación y corte. Este último resulta extremadamente eficiente para problemas de baja dimensionalidad, pero su desempeño se deteriora rápidamente de acuerdo con el crecimiento del espacio solución.

No obstante la virtud anteriormente expuesta para el uso de los algoritmos genéticos en lo relacionado al tema de la dimensionalidad, esto no resulta evidencia suficiente para afirmar que este tipo de algoritmos puedan ser aplicados de manera inmediata a cualquier tipo de problema. Lo antes mencionado no es trivial sobre todo en problemas de investigación como el nuestro, en donde el espacio solución descrito por el poliedro, se encuentre delimitado y sometido a un conjunto importante de restricciones (revisar apartado 6.2). Tan solo hace falta pensar en la cantidad tan enorme de organismos, que debido a las restricciones del poliedro, seguramente resultarán infactibles a lo largo

¹⁴⁹ Reeves, C. (1995), *Genetics algorithms flow shop sequencing*, Computers & Operations Research, Vol 22, EUA, pag 5-13.

¹⁵⁰ Esbensen, H. (1995), *Computing near-optimal solutions to the Steiner problem using genetics algorithms*, Networks Interscience, Vol 26, No.4, EUA, pag 173-185.

del proceso de optimización evolutiva. La situación antes descrita ocasiona que el algoritmo ocupe mucho de su tiempo computacional en sustituir los organismos enfermos. Esto a su vez, provoca que el proceso de exploración del espacio solución sea más lento y costoso¹⁵¹.

En sintonía con lo expuesto en el párrafo anterior, podemos citar a Premkumar, quién afirma que en general, la aplicación exitosa de los algoritmos genéticos está mucho más influenciada por el contexto restrictivo del problema que por su dimensionalidad¹⁵². A pesar de la interesante cantidad de investigaciones referentes al estudio analítico de los parámetros de un algoritmo genético, podemos afirmar que no existe un criterio de fijación general, permaneciendo mucho de su aplicación sobre el contexto particular del problema que esté siendo tratado así como de la empírica del investigador¹⁵³.

Aunque el objetivo del proyecto de investigación nunca fue establecer el conjunto de parámetros óptimos para el funcionamiento del algoritmo genético, por las razones ya antes expuestas, creemos que resulta importante al menos desarrollar un esfuerzo para su análisis predictivo. Lo anterior motiva al investigador a predecir aquellos niveles en los parámetros de operación del algoritmo genético, que pudieran ofrecer mejores expectativas de desempeño sobre la base particular del contexto restrictivo y dimensional del problema que está siendo atendido. Para dar respuesta a lo anterior, el investigador desarrollará a continuación un diseño experimental adicional.

La teoría del diseño de experimentos nos indica que para obtener resultados que sean concluyentes, debemos minimizar la cantidad de parámetros a ser manipulados en el experimento así como la cantidad de niveles para cada uno de los parámetros. El proceso anterior tiene el objetivo de minimizar la cantidad de combinaciones a ser revisadas lo cual equivale a minimizar la cantidad de tratamientos a ser aplicados en el experimento. Desafortunadamente, existen ocasiones en los que identificar a priori dichas combinaciones entre parámetros y sus niveles resulta inapropiado. En tal caso, lo que se requiere es establecer algún método que sistemáticamente vaya acotando el procedimiento de búsqueda combinatoria con la finalidad de lograr explorar de un modo eficiente el espacio solución.

¹⁵¹ Levine, G. (1997), *Developing fitter genetic algorithms*, INFORMS Journal of Computing, Vol. 9, No. 5, EUA, pag 251-253.

¹⁵² Premkumar, G. (1999), *Telecommunications Network design: a genetic algorithm approach*, School of Information Sciences and Technology & Pennsylvania State College of Business Administration, EUA, pag 19.

¹⁵³ Streifel, R. (1999), *Dynamic Fussy control of genetic algorithm parameters*, IEEE Trans. Syst., EUA, pag 426-433.

En ese sentido de sistematización del diseño experimental para el análisis de los parámetros de un algoritmo genético, el principal esfuerzo jamás realizado en términos de investigación es lo hecho por Stewardson y Hicks. Ellos proponen el uso de un diseño de experimentos “esparcido” (sparse en inglés) basado en una estrategia secuencial de dos etapas. La primera etapa corresponde a una fase de “visualización” (screening en inglés) con la aplicación de una cantidad limitada de replicas, mediante la cual se identifican los niveles más alto y más bajo correspondientes a cada uno de los parámetros experimentados. Los resultados de esta primera fase son procesados a través de un modelo general de ANOVA con la finalidad de descartar los niveles de aquellos parámetros que estadísticamente resulten con menor significancia relativa. Con los resultados obtenidos mediante ANOVA, se procede a desarrollar una segunda fase de experimentación sobre la base de una menor cantidad de niveles entre los parámetros y una mayor cantidad de replicas que en la fase anterior. Lo anterior hace plausible que el poder estadístico de esta segunda fase secuencial sea mucho más concluyente con respecto a lo que se pueda obtener de la primera¹⁵⁴.

¹⁵⁴ Stewardson, D. (2002), *Overcoming Complexity: optimizing genetic algorithms for use in complex scheduling problems via designed experiments*, University of Newcastle, United Kingdom, pag 5-13.

9.5 Diseño experimental y resultados para la predicción de los parámetros del Algoritmo Genético.

Conviene recordar en este momento que el diseño experimental que fue planteado para dar respuesta a la hipótesis del proyecto de investigación, se desarrolló sobre la base de fijar convenientemente los parámetros principales del algoritmo genético de acuerdo a la empírica del investigador. Sobre la base de esta empírica, podemos proponer que los parámetros del algoritmo genético que sean experimentados para optimizar su calibración, son los siguientes:

1. Tamaño de la población de organismos (N)
2. Cantidad de generaciones (G)
3. Factor de cruzamiento (FC)
4. Factor de mutación (FM)

El diseño experimental a desarrollar deberá cubrir y balancear la cantidad de mediciones a ser aplicadas para cada uno de los parámetros (o también llamados predictores), así como de sus niveles entre si. Así entonces procederemos a ejecutar una serie de mediciones sobre la base combinatoria de estos 4 parámetros y convenientemente tomando en consideración 4 niveles para cada uno de estos. El diseño experimental será desarrollado a partir de 10 replicas para cada uno de los 16 tratamientos (4 niveles para cada parámetro) con la finalidad de incrementar el poder estadístico del experimento.

Las Tablas 9.7, 9.8 y 9.9 muestran el diseño experimental a ser desarrollado. La Tabla 9.7, muestra las combinaciones así como los significados cuantitativos, para cada uno de los 4 niveles con que operan los parámetros relacionados al tamaño de la población (N) y a la cantidad de generaciones (G).

Tabla 9.7 Diseño experimental para los parámetros N y G.

Tamaño de la Población (N)		Cantidad de Generaciones (G)				Valor Real
		5,000	10,000	20,000	40,000	Valor Real
Valor Real	Val.Estd.	1	2	3	4	Val.Estd.
50	1	A1= X11	A2= X12	A3= X13	A4= X14	
100	2	A5= X21	A6= X22	A7= X23	A8= X24	
150	3	A9= X31	A10= X32	A11= X33	A12= X34	
200	4	A13= X41	A14= X42	A15= X43	A16= X44	

De la Tabla 9.7, podemos entonces ejemplificar al elemento $A10= X32$, como aquel tratamiento correspondiente a un tamaño de población (N) fijado a 150 organismos y con una cantidad de generaciones (G) fijado a un límite de 10,000 generaciones. A continuación abajo en la Tabla 9.8, se muestra las combinaciones así como los significados cuantitativos, para cada uno de los 4 niveles con que operan los parámetros relacionados al factor de cruzamiento (FC) y al factor de mutación (FM).

Tabla 9.8 Diseño experimental para los parámetros FC y FM.

Factor de Cruzamiento (FC)		Factor de Mutación (FM)				Valor Real
		2.5%	5%	7.5%	10%	Valor Real
Valor Real	Val.Estd.	1	2	3	4	Val.Estd.
20%	1	B1= Y11	B2= Y12	B3= Y13	B4= Y14	
40%	2	B5= Y21	B6= Y22	B7= Y23	B8= Y24	
60%	3	B9= Y31	B10= Y32	B11= Y33	B12= Y34	
80%	4	B13= Y41	B14= Y42	B15= Y43	B16= Y44	

De la Tabla 9.8, podemos entonces ejemplificar al elemento $B14= Y42$, como aquel tratamiento correspondiente a un factor de cruzamiento (FC) fijado al 80% y con un factor de mutación (FM) fijado a un 5%. Finalmente en la Tabla 9.9, se muestra las combinaciones para cada uno de los 4 niveles con que operan los 4 parámetros involucrados en el experimento (N, G, FC y FM). Esto equivale entonces a un total de 256 combinaciones posibles, es decir 4 parámetros combinados a 4 niveles ($4^4 = 256$).

Tabla 9.9 Diseño experimental para los parámetros N, G, FC y FM.

INTERACCION N-G		INTERACCION FC-FM					Valor Real
		Y11	Y12	Y43	Y44	Valor Real
Valor Real	Val.Estd.	B1	B2	B15	B16	Val.Estd.
X11	A1	T1= X11,Y11	T2= X11,Y12	T15= X11,Y43	T16= X11,Y44	
X12	A2	T17= X12,Y11	T18= X12,Y12	T31= X12,Y43	T32= X12,Y44	
.....	
X43	A15	T225= X43,Y11	T226= X43,Y12	T239= X43,Y43	T240= X43,Y44	
X44	A16	T241= X44,Y11	T242= X44,Y12	T255= X44,Y43	T256= X44,Y44	

Sobre la base de un total de 256 tratamientos y una cantidad de 10 replicas, se tiene entonces un total de 2560 ejecuciones del algoritmo genético. Dicha información estadística es transferida al software computacional SPSS ver 11.0 para su análisis correspondiente. A continuación en la Tabla 9.10, se muestran los resultados provenientes del análisis de varianza:

Tabla 9.10 Análisis de Varianza Univariado: Nivel de significancia para los predictores.

Dependent Variable: FOBJ

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2044506.597 ^a	4	511126.649	3268.080	.000
Intercept	1759480.821	1	1759480.821	11249.901	.000
TPOB	2040286.114	1	2040286.114	13045.335	.000
NGENE	976.483	1	976.483	6.244	.013
FCRUCE	11.357	1	11.357	.073	.788
FMUTA	619.156	1	619.156	3.959	.047
Error	399601.163	2555	156.400		
Total	10134069.1	2560			
Corrected Total	2444107.760	2559			

a. R Squared = .937 (Adjusted R Squared = .936)

El valor del estadístico “F” (penúltima columna) es calculado dividiendo el promedio cuadrado de cada parámetro (columna 4) entre el promedio cuadrado del error (619,15). El nivel de significancia (o valor de “p”) para cada parámetro, es verificado en la última columna. En general, un parámetro con un valor de “p” por abajo del 5% se le considera como significativo. De los coeficientes de la prueba “F” (penúltima columna) y de los valores de significancia o valores “p” (última columna), revisados en la tabla anterior se pueden obtener los siguientes resultados:

1. El parámetro “TPOB”, que se refiere al tamaño de la población (N), es el que obtiene el mayor nivel de significancia (1-p = 100%) de los 4 predictores que están siendo confrontados, lo que significa que el tamaño de la población (N) influye decididamente en el valor de la función objetivo obtenido por el algoritmo genético.
2. Los siguientes 2 parámetros de mayor significancia en ese mismo orden, son “NGENE” y “FMUTA”. El primero se refiere a la cantidad de generaciones (G) ejecutadas por el algoritmo genético el cual obtiene un nivel de significancia bastante alto (1-p = 98.7%). El segundo parámetro se refiere al factor de mutación (FM) que es empleado por el algoritmo genético para

mantener la diversidad en la población de organismos. En este caso obtenemos un nivel de significancia suficiente para las expectativas del investigador ($1-p= 95.3\%$)

3. En lo concerniente al último parámetro FCRUCE el cual corresponde al factor de cruzamiento (FC) utilizado por el algoritmo genético durante el proceso de reproducción de los organismos, tenemos que su valor de “p” excede del 5%, por lo que entonces podemos deducir que dicho parámetro no establece fuertemente una relación causa-efecto sobre la función objetivo obtenida en el algoritmo genético.
4. El valor de R^2 *ajustado* es de 93.6%. Al valor anterior se le conoce como coeficiente de determinación el cual se le utiliza como una métrica de bondad de ajuste respecto al modelo lineal. Dicho de otro modo, éste valor puede ser entendido como la proporción de la variabilidad de la variable dependiente (en nuestro caso es el valor de la función objetivo) que puede ser explicada a través del modelo de regresión que esté siendo utilizado. Un porcentaje de determinación pequeño, significaría que el modelo no se ajusta adecuadamente a los datos. Podemos finalmente decir que, el coeficiente de determinación obtenido en nuestro caso satisface plenamente los requerimientos del experimento.

Finalmente a continuación en la Tabla 9.11, se muestran los resultados para la estimación de los coeficientes de predicción para cada uno de los parámetros. Cabe mencionar que el cálculo de dichos coeficientes se realizó sobre la base de un modelo lineal de predicción el cual se encuentra apoyado sobre la base del coeficiente de determinación del 93.6% obtenido en el análisis anterior.

Tabla 9.11 Estimación estadística de los coeficientes e intervalos de confianza para los parámetros de operación del algoritmo genético.

Dependent Variable: FOBJ

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	119.064	1.123	106.066	.000	116.863	121.265
TPOB	-24.923	.218	-114.216	.000	-25.351	-24.495
NGENE	-.549	.220	-2.499	.013	-.979	-.118
FCRUCE	5.957E-02	.221	.269	.788	-.374	.493
FMUTA	-.435	.219	-1.990	.047	-.863	-6.290E-03

La 2da. columna de la tabla anterior, corresponde al valor del coeficiente del modelo lineal que está siendo estimado para cada uno de los parámetros que aparecen a su vez en la 1era. columna. La 4ta. columna corresponde al valor de la prueba “t” de student, el cual se calcula dividiendo el valor del coeficiente de cada parámetro entre su error estándar. En términos generales la prueba “t” de student, se utiliza para determinar el grado de importancia relativa que cada uno de los parámetros guarda en el modelo. Como regla general podemos establecer que aquellos parámetros cuyo valor absoluto de “t” sea mayor o igual a 2, resultan ser buenos predictores respecto a la variable dependiente.

El valor de “t” antes explicado a su vez, se asocia con el nivel de significancia (valor de “p”) el cual se reporta en la 5ta. columna. Dicho valor guarda una interpretación equivalente a la reportada en la prueba “F” de análisis de varianza de la Tabla 9.10 ya antes expuesta. Finalmente los valores que aparecen en la 6ta y 7ma. columna corresponden a los intervalos de confianza al 95% para los coeficientes del modelo lineal calculados para cada parámetro en la columna 2. De los coeficientes reportados en la Tabla 9.11 podemos interpretar los siguientes resultados:

1. El parámetro “TPOB”, o tamaño de la población (N), tiene un coeficiente de estimación negativo (-24.923), lo cual significa una proporcionalidad inversa. Lo anterior comprueba que a medida que la población de organismos en el algoritmo genético se incrementa, en esa medida el valor de la función objetivo (costo) se logra disminuir en términos numéricos, lo cual debe ser entendido como un efecto favorable en la calidad de la solución. También podemos comprobar que el valor absoluto del coeficiente es el más grande que se reporta en el modelo, lo cual quiere decir que es el que también logra mayormente influir en el resultado de la función objetivo del algoritmo genético. Lo anterior se comprueba por igual al revisar el coeficiente “t” el cual en términos absolutos, también resulta ser el mayor en todo el modelo.
2. El parámetro “NGENE”, que se refiere a la cantidad de generaciones (G), tiene un coeficiente de estimación mucho menor en términos absolutos (-0.549), lo cual significa que aunque, a medida que transcurren las generaciones en el algoritmo genético se mejora el resultado de la función objetivo, no obstante, el efecto no resulta ser tan impactante en el resultado. De cualquier manera al revisar su coeficiente “t” (-2.499), éste nos sugiere que el parámetro no debe ser excluido del modelo.

3. El parámetro “FMUTA”, que se refiere al factor de mutación (FM), tiene un coeficiente de estimación muy comparable (-0.435) al revisado en el caso de la cantidad de generaciones (G). Esto nuevamente hace deducir a mayor factor de mutación en el algoritmo genético también se obtendrán mejores soluciones en la función objetivo. El valor de su coeficiente “t” (-1.990), resulta apenas ser suficiente para que el parámetro pueda catalogarse como relevante en el modelo de predicción. Lo que si resulta interesante observar en este parámetro respecto a lo verificado en los otros dos previamente expuestos, es en lo referente a los intervalos de confianza que aparecen registrados. El límite inferior del coeficiente (lower bound), favorece la hipótesis de que cualquier incremento en el factor de mutación mejorará la solución obtenida por el algoritmo genético. No obstante el valor de (-6.290 E-03) en el límite superior del coeficiente (upper bound), nos hace dudar de lo anterior sobre la base del intervalo de confianza fijado en el experimento al 95%. Este coeficiente (-0.000629) es tan pequeño, que bajo estas circunstancias, la influencia que el factor de mutación pudiera ejercer en la solución obtenida por el algoritmo genético sería muy marginal.

4. Finalmente el parámetro “FCRUCES”, que se refiere al factor de cruzamiento (FC), tiene el coeficiente de estimación en valor absoluto más pequeño de todos (5.957 E-02). Lo que de este valor podemos interpretar, es que dicho parámetro afecta marginalmente el resultado obtenido en la solución del algoritmo genético. El valor de su coeficiente “t” (-0.269), comprueba lo anterior. Nuevamente resulta interesante observar en este parámetro lo referente a los intervalos de confianza que aparecen registrados. El límite inferior del coeficiente es negativo (-0.374), mientras que el valor de su límite superior es positivo (0.493). Esta ambivalencia lo que nos hace deducir con respecto a este parámetro, es que al nivel de confiabilidad del 95% que había sido pretendido, no es posible probar la hipótesis que indique la medida en la que este parámetro afecta a la solución del algoritmo genético. Lo anterior no significa que debemos aceptar la hipótesis de que el factor de cruzamiento no tiene influencia en la solución del algoritmo genético, sino que para probarlo o rechazarlo se debería utilizar algún otro diseño experimental lo cual escapa al alcance y objetivo trazado en nuestro proyecto de investigación.

Con lo expuesto hasta aquí, establecemos entonces el siguiente modelo lineal de predicción para el comportamiento de la función objetivo solucionada a través del algoritmo genético propuesto:

$$FObj.Est = 119.064 - 24.923 tpob - 0.549 ngene - 0.435 fmuta$$

Aunque el modelo lineal de predicción tiene un coeficiente de determinación R^2 *ajustado* igual al 93.6% (revisar tabla 9.10), el cual es bastante razonable, no obstante, el investigador a continuación propone un modelo de regresión No-Lineal con el propósito de finalmente estimar la brecha de optimalidad en el poliedro del problema SPDP-sTW, el cual a su vez esta descrito a través del diseño factorial que fue aplicado a los parámetros de operación del algoritmo genético propuesto.

El modelo de regresión NO-lineal que a continuación se propone está compuesto por 3 componentes polinomiales para cada uno de los parámetros de operación del algoritmo genético. Los 3 componentes del polinomio son: uno lineal, otro cuadrático y finalmente uno más del tipo hiperbólico. A continuación se expone el modelo propuesto:

$$\begin{aligned}
 FObj.Est = & A_0 + A_1 * tpob + A_2 * tpob^2 + \frac{A_3}{tpob} + \\
 & B_1 * ngene + B_2 * ngene^2 + \frac{B_3}{ngene} + \\
 & C_1 * fcruza + C_2 * fcruza^2 + \frac{C_3}{fcruza} + \\
 & D_1 * fmuta + D_2 * fmuta^2 + \frac{D_3}{fmuta}
 \end{aligned}$$

El modelo anterior fue resuelto mediante la minimización de la suma de diferencias cuadráticas a través de un algoritmo convencional de programación lineal. El coeficiente de determinación R^2 *ajustado* arrojó 98.2%, lo cual supera al obtenido a través del modelo de predicción lineal. A continuación en la Tabla 9.12, se presentan los coeficientes obtenidos para el modelo:

Tabla 9.12 Coeficientes obtenidos para el modelo polinomial.

	X	X ²	1/X	PARAM. OPT
	LINEAL	CUADRÁ- TICO	HIPER- BÓLICO	
TPOB (A _i)	2. 1865	-0. 2581	106. 1476	4. 0
NGENE (B _i)	-1. 6492	0. 1966	-1. 2288	4. 0
FCRUCE (C _i)	-1. 3911	0. 1740	-2. 0155	3. 5348
FMUTA (D _i)	-6. 6890	1. 0300	-4. 3781	3. 0130
A ₀	13. 2691			

Las columnas 2, 3 y 4 corresponden a la exposición de los coeficientes del modelo polinomial antes descrito. La columna 2 corresponde al término lineal. La siguiente columna al término cuadrático y la cuarta al término hiperbólico. Finalmente la última columna corresponde a los valores óptimos calculados para cada uno de los parámetros del algoritmo genético descritos a través del modelo polinomial antes mencionado. Existen básicamente dos métodos para calcular los valores óptimos para los parámetros del modelo:

1. Aplicando al modelo polinomial 4 derivadas parciales. Cada derivada parcial corresponde a cada uno de los parámetros del algoritmo genético. Cada derivada parcial a su vez, es igualada a cero (gradiente igual a cero) para determinar el punto de inflexión de la curva de costo de la función objetivo con respecto al parámetro en cuestión. De esta manera se obtienen 4 ecuaciones con 4 incógnitas (corresponden a los 4 parámetros del modelo polinomial), las cuales pueden ser resueltas por cualquier método de convergencia.
2. Aplicando cualquier algoritmo de optimización que pueda resolver problemas cuya superficie de respuesta NO esté delimitada por hiper-planos lineales. Naturalmente, este tipo de problemas no puede ser atendido a través de métodos de solución exacta sino solo a través de métodos de aproximación.

En el caso nuestro el investigador no dejó pasar la oportunidad para volver a aplicar, una vez más, el algoritmo genético propuesto para aproximar los valores óptimos correspondientes a los parámetros del modelo polinomial. Así entonces, si sustituyéramos los valores óptimos calculados en el modelo de regresión polinomial, obtendríamos entonces un valor mínimo en el costo de la función objetivo equivalente a “25.094”, el cual es el valor que utilizaremos como cota inferior para estimar la brecha de optimalidad en el poliedro del problema SPDP-sTW.

Hecho lo anterior, ya solo faltaría calcular un valor apropiado para el costo de la función objetivo sobre la cual, al compararla versus la cota inferior, dicha diferencia cuantitativa fuese entonces interpretada analíticamente como la brecha de optimalidad a poder potenciar como beneficio económico. Para poder establecer dicho costo referencial analizaremos a continuación en la Tabla 9.13, la Distribución estadística del costo en la función objetivo para el espacio solución.

Tabla 9.13 Distribución estadística del costo en la función objetivo para el espacio solución.

Distribución Estadística del Costo en la Función Objetivo para el Espacio Solución							
Suma Aritmética de los Parámetros	Rangos de Valores obtenidos en el Costo de la Función Objetivo						TOTALES
	25-30	30-35	35-40	50-55	60-105	105-110	
4	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.4%
5	0.0%	0.0%	0.0%	0.5%	0.7%	0.4%	1.6%
6	0.0%	0.0%	0.3%	1.3%	1.6%	0.7%	3.9%
7	0.2%	0.0%	1.2%	2.0%	4.1%	0.5%	7.9%
8	1.5%	0.5%	1.8%	4.1%	4.7%	0.0%	12.5%
9	2.4%	0.9%	2.7%	4.6%	4.9%	0.0%	15.4%
10	3.5%	1.8%	2.9%	4.7%	4.3%	0.0%	17.1%
11	4.3%	2.4%	1.9%	3.8%	2.5%	0.0%	15.0%
12	5.3%	2.9%	0.9%	2.4%	1.1%	0.0%	12.5%
13	3.8%	2.0%	0.5%	1.1%	0.3%	0.0%	7.7%
14	2.5%	0.9%	0.0%	0.3%	0.0%	0.0%	3.8%
15	1.7%	0.3%	0.0%	0.0%	0.0%	0.0%	2.0%
16	0.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%
TOTALES	25.5%	11.6%	12.0%	24.6%	24.3%	2.0%	100.0%

En la Figura 9.1 se puede comprobar la normalidad que existe en la distribución de frecuencias correspondientes al diseño factorial que fue empleado para el experimento.

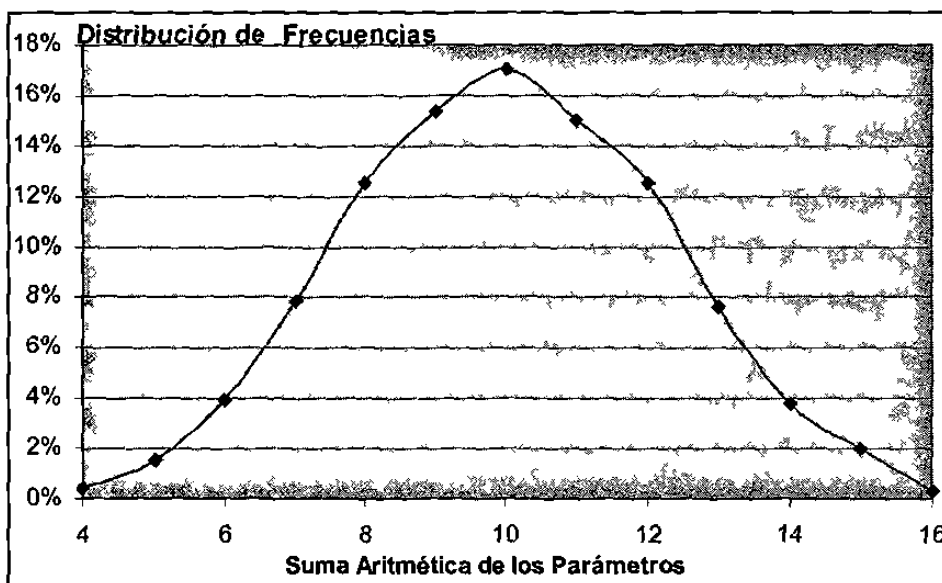


Figura 9.1 Distribución de frecuencias para la suma aritmética de parámetros del algoritmo genético.

Finalmente en la Figura 9.2, se presenta la Distribución Acumulada del espectro del Espacio Solución del problema SPDP-sTW descrito por los parámetros del algoritmo genético.

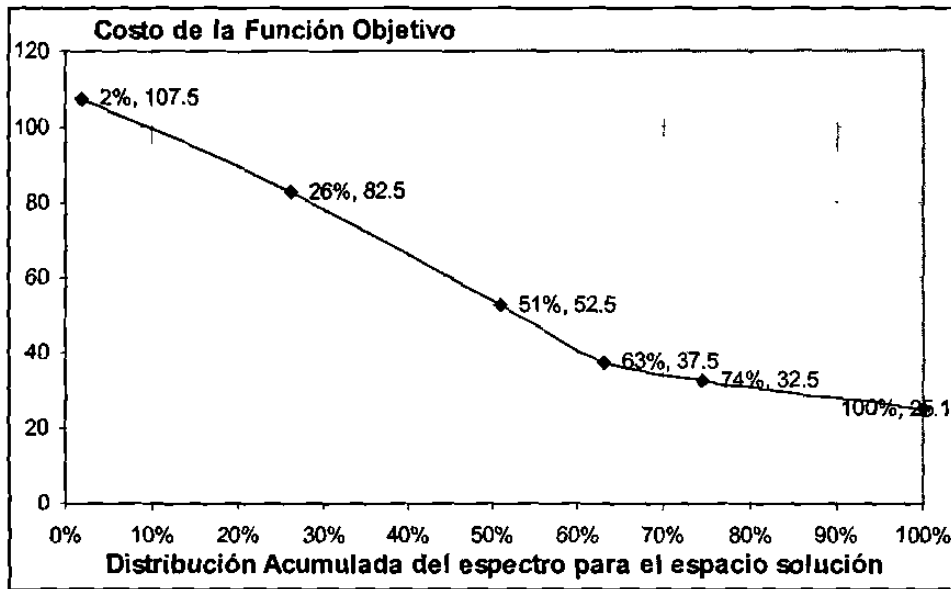


Figura 9.2 Distribución Acumulada del espectro del Espacio Solución del problema.

Para explicar la interpretación de la curva anterior podemos dar el siguiente ejemplo. Más del 40% del espacio solución del problema oscila apenas entre un valor del costo aproximado a 40 y el valor mínimo de 25.1, identificado ya anteriormente como el valor óptimo esperado (25.094). Dicho de otro modo el 40% de las soluciones del espacio total, apenas logran mejorar el porcentaje de optimalidad en un 13.9% al disminuir 14.9 puntos el valor del costo en la función objetivo con respecto a la cota superior (107.5), el cual es calculado a través de la curva descrita por el modelo polinomial.

De la Figura 9.2, podemos establecer que una adecuada referencia contra la cual comparar la cota inferior para determinar la brecha de optimalidad, pudiera ser a través del valor promedio del costo en la función objetivo descrito por la curva del modelo polinomial. Dicho costo promedio estaría ponderado por el tamaño del espacio solución en la función objetivo del problema. De la Figura 9.2, podemos verificar que dicho costo promedio equivale a 52.9 aproximadamente. Entonces la brecha de optimalidad puede ser finalmente calculada de la siguiente manera:

$$\frac{52.9 - 25.094}{107.5} = 25.87\%$$

Como puede ser revisado, éste valor estimado supera el 10% y el 25% que habíamos calculado mediante el muestro en el apartado 9.3.

De lo anterior, podemos entonces identificar solo dos parámetros restantes que son el tamaño de la población (TPOB) y el factor de mutación (FMUTA). Por tanto, el análisis a desarrollar consiste en estimar la relación de estos dos parámetros del algoritmo genético con respecto a la eficacia en el porcentaje de optimalidad alcanzado. Aunque en el apartado 9.5 se comprueba que ambos parámetros son significativos con respecto al valor de la función objetivo, no obstante, no es posible afirmar lo mismo con respecto a la eficacia en el porcentaje de optimalidad alcanzado. Para este análisis se utilizarán los mismos datos que se aplicaron en el apartado 9.5 con la excepción de que se añadirá un parámetro adicional que es el que se refiere al tiempo computacional (TIEMPO). Dicho tiempo computacional corresponde a aquel que fue registrado durante el levantamiento de los datos que fueron ocupados para el experimento expuesto en el apartado 9.5.

La nueva variable que será utilizada para medir la eficacia en el porcentaje de optimalidad alcanzado, la denominaremos simplemente como (EFICACIA). Como es de esperarse, la forma como mediremos la variable EICACIA será a través de la interacción del valor en la función objetivo (FOBJ) obtenido mediante el algoritmo genético, junto con el valor del tiempo computacional registrado (TIEMPO). Así entonces la formulación que usaremos para calcular la variable EFICACIA es la siguiente:

$$EFICACIA_i = \sqrt{\frac{FOBJ_{opt} * TIEMPO_{min}}{FOBJ_i * TIEMPO_i}} \quad \forall \quad i = 1 ..n \quad \text{donde } n = 1..2560$$

$$FOBJ_{opt} = 25.096$$

$$TIEMPO_{min} = \min(TIEMPO_i)$$

Como puede ser visto en la formulación anterior, la variable EFICACIA mide el desempeño del algoritmo genético simultáneamente en dos dimensiones, es decir, con respecto al valor óptimo esperado de la función objetivo y con respecto a un valor referencial de tiempo. Dicha referencia de tiempo, convenientemente se fija al identificar el tiempo mínimo que haya sido registrado en el experimento. Así entonces, la variable EFICACIA mide a la vez dos aspectos:

1. Que tan lejos queda el valor de la función objetivo respecto al valor óptimo esperado.
2. Con que rapidez se obtiene dicho valor en la función objetivo

Ambos factores de desempeño se multiplican entre sí, con la finalidad de calcular el efecto conjunto que es el que realmente nos interesa medir. La raíz cuadrada en la fórmula, simplemente se utiliza para obtener el desempeño promedio de ambos factores conjuntamente.

Hecho lo anterior, ejecutamos el análisis de significancia en el software estadístico SPSS Ver 11.0 con la finalidad de verificar los resultados. A continuación se muestra en la Tabla 9.14 el análisis de varianza.

Tabla 9.14 Nivel de significancia para los predictores de la variable Eficacia.

Tests of Between-Subjects Effects

Dependent Variable: EFICACIA

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	23.102 ^a	4	5.776	1594.384	.000
Intercept	3.609	1	3.609	996.381	.000
TPOB	6.597E-03	1	6.597E-03	1.821	.177
FMUTA	.511	1	.511	141.113	.000
FOBJ	3.871E-02	1	3.871E-02	10.687	.001
TIEMPO	3.638	1	3.638	1004.272	.000
Error	9.255	2555	3.622E-03		
Total	270.664	2560			
Corrected Total	32.357	2559			

a. R Squared = .814 (Adjusted R Squared = .814)

De la tabla anterior, podemos destacar los siguientes resultados:

1. El valor de la función objetivo (FOBJ) así como el tiempo computacional aplicado (TIEMPO), ambos elementos resultan ser significativos al calcular la eficacia en el desempeño del algoritmo genético (EFICACIA). La variable EFICACIA debe ser entendida como la propiedad activa en el algoritmo genético para lograr obtener soluciones de buena calidad en tiempos computacionales pequeños (calidad vs tiempo).
2. El parámetro "Tamaño de la población" (TPOB), no es significativo al 5% de confiabilidad ($1-p = 82.3\%$), como predictor de la eficacia en el desempeño del algoritmo genético. Lo anterior comprueba que aunque, el manejar una población grande de individuos contribuye a aumentar las posibilidades de obtener mejores soluciones debido a la diversidad de material genético presente en la población, no obstante, el costo incurrido para sostener una población más grande se repercute en el tiempo computacional para su procesamiento.

3. El “factor de mutación” (FMUTA), si resulta ser un componente significativo en lo referente al balance entre la obtención de soluciones de buena calidad versus el tiempo (ó costo) requerido para lograrlo. Lo anterior, no significa que al aumentar el valor del parámetro FMUTA, necesariamente se obtendrá un mejor desempeño en la variable EFICACIA. Lo que quiere decir, es que el valor de FMUTA, tiene un rol definitivamente significativo en el desempeño del algoritmo genético al hablar del binomio (calidad vs tiempo).
4. El coeficiente de determinación (R^2 ajustado) es del 81.4%. Esto quiere decir que la proporción de la variabilidad registrada en la variable EFICACIA puede ser explicada en un 81.4% a través de los predictores (TPOB y FMUTA). Aunque el coeficiente es razonablemente bueno, no obstante, lo anterior hace prever que deben existir otros elementos que están influenciando el resultado de la eficacia en el desempeño del algoritmo genético. El tratamiento para el análisis de dicha fuente de variabilidad está fuera del alcance de la presente tesis. No obstante, el investigador sugiere que dicha variabilidad está originada primordialmente por el manejo del resto de las fases algorítmicas (heurísticas) que están siendo incluidas en adición al propio algoritmo genético¹⁵⁵.

A continuación se muestra en la Tabla 9.15 el cálculo de los coeficientes de estimación así como de los intervalos de confianza.

Tabla 9.15 Coeficientes de estimación e intervalos de confianza.

Parameter Estimates

Dependent Variable: EFICACIA

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	.413	.013	31.566	.000	.388	.439
TPOB	6.006E-03	.004	1.349	.177	-2.721E-03	1.473E-02
FMUTA	1.415E-02	.001	11.879	.000	1.181E-02	1.648E-02
FOBJ	-3.41E-04	.000	-3.269	.001	-5.457E-04	-1.365E-04
TIEMPO	-9.34E-03	.000	-31.690	.000	-9.922E-03	-8.766E-03

¹⁵⁵ Supra: revisar el apartado 7.7

De la tabla anterior, podemos destacar los siguientes resultados:

1. El coeficiente de TPOB ($6.006E-03$), evidencia que dicho parámetro ofrece una influencia predictiva muy baja sobre el valor numérico de la variable EFICACIA.
2. El intervalo de confianza obtenido para el coeficiente TPOB ($-2.721E-03$ a $1.473E-02$) deja claro que la influencia de dicho parámetro en la variable EFICACIA, aparte de ser pequeña, también es ambigua. Lo anterior se explica por el cambio de signo que existe entre el intervalo inferior y el superior del coeficiente.
3. Aunque los parámetros FOBJ y TIEMPO tienen coeficientes muy pequeños, esto es previsible, ya que éstos fueron utilizados para calcular directamente el valor de la variable EFICACIA.
4. Los valores de ambos coeficientes (FOBJ y TIEMPO) son de signo negativo, por lo que se confirma la tendencia como ambos parámetros influyen el valor de la variable EFICACIA. Lo anterior quiere decir, que en ambos parámetros la afectación es inversamente proporcional, es decir, a menor valor en la función objetivo (o sea mejor calidad) y a menor tiempo registrado (más rápido), entonces la eficacia del algoritmo aumenta. Los signos reportados en los intervalos de confianza (superior e inferior) no dejan ninguna sospecha al respecto.
5. El coeficiente del parámetro relacionado al factor de mutación (FMUTA), es el que ofrece en términos relativos, la mayor influencia ($1.415E-02$) sobre la eficacia en el desempeño del algoritmo genético. Los signos reportados en el intervalo de confianza, también muestran que este parámetro mantiene una tendencia consistente en relación a la variable EFICACIA.
6. El coeficiente del parámetro (FMUTA) es positivo, lo cual significa que la tendencia es directamente proporcional con respecto a la variable EFICACIA. Esto quiere decir que sería previsible que al aumentar el factor de mutación en el algoritmo genético, también fuese esperable que aumentase la eficacia del algoritmo genético. El valor del coeficiente indica un gradiente del 1.42% en el mejoramiento de la eficacia con respecto al factor de mutación
7. A pesar del valor del gradiente anterior, no es posible concluir que por cada incremento del 1% en el factor de mutación fuese posible indefinidamente obtener un mejoramiento del 1.42% en la eficiencia del algoritmo genético. El razonamiento anterior es sobre la base de que el modelo de predicción que fue aplicado para el análisis de significancia es lineal y por tanto solo es posible establecer un criterio de tendencia y no de optimalidad como ocurrió con el modelo polinomial desarrollado en la tabla 9.12 ya antes expuesto.

Con el análisis anterior, damos por concluido el presente capítulo quedando satisfechos con los resultados expuestos.

9.7 Comentarios finales.

Con la exposición anterior, damos por terminado el capítulo referente a los resultados obtenidos en la fase de experimentación. Como pudimos comprobar, el diseño experimental expuesto, dio tratamiento a los siguientes aspectos:

1. En el apartado 9.2 se dio respuesta a la hipótesis de investigación mediante un análisis comparativo de las soluciones obtenidas mediante el algoritmo genético propuesto con respecto al grupo control. El grupo control fue constituido de las soluciones óptimas verificables obtenidas mediante el uso de un algoritmo de ramificación y corte.
2. En el apartado 9.3, se desarrolló un análisis costo beneficio. Este contribuyó a ofrecer una justificación económica acerca de la viabilidad del algoritmo propuesto para mejorar la capacidad de transporte y el servicio al cliente sobre la base de supuestos conservadores definidos empíricamente por el investigador.
3. En el apartado 9.4 se establecieron las bases para el desarrollo de un modelo de regresión lineal con la finalidad de determinar la significancia estadística así como la importancia relativa que cada uno de los parámetros del algoritmo genético tienen con respecto a la obtención de mejores soluciones en la función objetivo.
4. A partir del modelo de regresión lineal anterior, en el apartado 9.5 se propone un modelo de regresión polinomial. Mediante éste último se comprueban los supuestos de mejora en la métrica de optimalidad que fueron propuestos por el investigador en el apartado 9.3. El modelo de predicción permite estimar el tamaño de la brecha de optimalidad. Los resultados del modelo analítico sobre pasan los supuestos de mejora establecidos en el apartado 9.3.
5. En el apartado 9.6, se desarrolla un análisis para determinar la significancia de los parámetros del algoritmo genético con respecto a su eficacia en el tratamiento del binomio “calidad vs tiempo”. Los resultados muestran que aunque el tamaño de la población sí resulta ser significativo en cuanto al valor de la función objetivo, no obstante, el tiempo computacional queda comprometido.

Finalmente, es importante destacar que para poder establecer conjeturas definitivas acerca de la definición de los niveles óptimos de operación para los parámetros involucrados en el algoritmo genético propuesto, sería necesario desarrollar estudios y diseños experimentales particulares tales como los expuestos en el apartado 9.4 desarrollados por Stewardson y Hicks.