

Diseño de un modelo matemático para el despacho de vehículos de emergencias médicas en Colombia.

Alvaro José Rengifo Campo

Pontificia Universidad Javeriana, Cali, Colombia
Calle 18 N° 118-250 Vía a Pance, Cali
ajrengifo@javerianacali.edu.co

María Gulnara Baldoquin de la Peña

Pontificia Universidad Javeriana, Cali, Colombia
Calle 18 N° 118-250 Vía a Pance, Cali
mgulnara@javerianacali.edu.co

John Wilmer Escobar

Pontificia Universidad Javeriana, Cali, Colombia
Calle 18 N° 118-250 Vía a Pance, Cali
jwescobar@javerianacali.edu.co

RESUMEN

El trabajo es parte de un proyecto en Logística Hospitalaria para la solución de problemas de localización, despacho y relocalización de vehículos de una importante empresa prestadora de servicios de emergencias médicas en Cali, Colombia. No se ha encontrado en la literatura trabajos en esta área que se ajusten al modelo de operación de dicha compañía. Se presenta un modelo de Programación Lineal Entera Mixta para el problema de despacho de vehículos teniendo en cuenta las restricciones impuestas por la compañía. El modelo responde a un MDVRPTW (Multi Depot Vehicle Routing Problem with Time Windows) que se resuelve cada vez que existe algún cambio en el estado del sistema, como recepción de una nueva llamada, avería de un vehículo, etc. Se generan soluciones del modelo para un intervalo de tiempo representativo de despacho, las cuales mejoran soluciones heurísticas obtenidas de diversas políticas de despacho actuales de dicha compañía.

Palabras clave: modelación, logística, ruteo.

Área principal: Logística y Transporte, OR en la salud, Optimización Combinatoria

ABSTRACT

This work is part of a project in Hospital Logistics for the solution of localization problems, dispatch and relocation of vehicles for an important company of services of medical emergencies in Cali, Colombia. In the literature has not been found any work that is adjusted to the operation pattern of this company. A model of Mixed Integer Lineal Programming is presented for the problem of dispatch of vehicles taking into account the restrictions imposed by the company. The model responds to a MDVRPTW (Multi Depot Vehicle Routing Problem with Time Windows) that it is solved every time some change exists in the state of the system, as reception of a new call, damages of a vehicle, etc. Solutions of the model were generated for a representative interval of time of dispatch, which improve obtained heuristic solutions of three actual politicians of this company.

Keywords: modelling, logistic, routing.

Main Area: L & T - Logistics and Transport, IND – OR in Health, OC – Combinatorial Optimization

1. Introducción

En el trabajo se propone un modelo matemático para la asignación y/o despacho de vehículos de emergencias médicas de una reconocida compañía que presta esos servicios en Colombia.

En la mayoría de modelos de asignación de vehículos a llamadas de emergencia, según el estado del arte revisado, se tiene en cuenta la utilización del vehículo disponible más cerca al área de atención. Con el uso del concepto de *preparedness* (nivel de preparación del sistema para atender las llamadas de emergencia en un tiempo razonable) artículos más recientes (Andersson et al. (2004, 2006, 2007), Lee (2011)) no necesariamente siguen el mismo principio teniendo en cuenta mantener el *preparedness* en un área determinada. Por ejemplo, en (Andersson (2007)) para llamadas de prioridad uno asignan la ambulancia más cercana, en llamadas de prioridad dos no siempre se sigue ese principio para mantener el *preparedness* en un condado dado. Las llamadas de prioridad tres, no urgentes, pueden ser recibidas y planificadas con anterioridad.

En estos trabajos revisados se tienen en cuenta solo flotas homogéneas de vehículos.

A partir del estudio del modelo de operación de dicha compañía se presenta un modelo de Programación Lineal Entera Mixta para el problema de despacho de vehículos que responde a un MDVRPTW (Multi Depot Vehicle Routing Problem with Time Windows).

El MDVRPTW es una extensión del clásico problema VRP (Polacek et al. (2004)). Además de las restricciones impuestas a un VRP, en un problema MDVRPTW hay múltiples depósitos y restricciones de ventanas de tiempo que especifican el tiempo disponible para el servicio a cada cliente. Existen múltiples aplicaciones de este tipo de problemas (Cordeau et al. (2002), Lau et al. (2003)). Debido a que estos problemas son del tipo NP de difícil resolución, se han aplicado diversos algoritmos aproximados, especialmente metaheurísticas en su resolución como por ejemplo Recocido Simulado (Czarnas et al. (2004)), Búsqueda Tabú (Amberg (2000)), Algoritmos Genéticos (Alba et al. (2004)), metaheurísticas híbridas usando por ejemplo Algoritmos Genéticos y Tabú Search (Noori y Ghannadpour (2012)).

El modelo propuesto se resuelve, usando el ILOG CPLEX, cada vez que existe algún cambio en el estado del sistema, como recepción de una nueva llamada, avería de un vehículo, etc. Con los datos de un intervalo de tiempo representativo de llamadas recibidas en el primer semestre del año 2011 en la ciudad de Cali se generaron 33 instancias, que corresponden al estado del sistema con cada nueva llamada entrante en dicho intervalo. Las soluciones del modelo para dichas instancias, se comparan con soluciones heurísticas obtenidas simulando las políticas actuales de despacho de dicha compañía.

El método de solución propuesto al problema de despacho de vehículos de emergencia médica no se ha encontrado en la literatura, tampoco consideraciones de políticas de despacho similares.

2. Estructura del problema real actual

La compañía presta, desde el 2002, servicios de consulta, consulta telefónica, urgencia, emergencia y traslado. Estimó una demanda de más de 25000 clientes-usuarios en el 2010 en la ciudad de Cali, con diferencias significativas en servicios prestados con vehículos médicos en cuanto a tres tipos de llamadas, denominadas en orden de prioridad: Emergencia (urgente, síntomas con amenaza a la vida), Urgencia (urgente, pero no existe amenaza a la vida) y Consultas (no urgentes). Dispone de una flota heterogénea de tres tipos de vehículos clasificados en vehículo de consulta (VC), Transporte de Ambulancia Básica (TAB) y Transporte de Ambulancia Medicalizada (TAM). Los vehículos de tipo TAM pueden atender cualquier tipo de llamada, los de tipo TAB no pueden atender emergencias y los VC solo pueden atender consultas.

La empresa tiene definido diferentes tiempos de respuesta (despacho + llegada) así como de atención para cada tipo de llamada. Para la localización de sus vehículos tiene la ciudad dividida en 5 sectores (sur, norte, este, oeste, centro) con un número de barrios en cada uno de ellos. El número de puntos potenciales y puntos reales donde ubica sus vehículos supera el número disponible de los mismos. Cuenta con un estudio de la demanda promedio durante las 24 horas de operación de un día, identificándose horas del día en que la demanda se presenta como alta y estacionaria ó baja, no importa cuál sea el día. De acuerdo con ello define dos intervalos de tiempo para cada uno de los cuales moviliza una flota diferente de los diferentes tipos de vehículos que dispone para operar.

La decisión actual de asignar llamadas a vehículos, de tipo operativa, así como reasignar vehículos en zonas estratégicas, depende de la experticia y habilidad del operador, donde una no adecuada decisión puede influir negativamente en decisiones de llamadas futuras. Por otro lado, el proceso de asignación es complejo, determinado por un conjunto de factores:

1. Cada tipo de llamada debe tener un tiempo máximo limitado de respuesta.
2. Poseer un parque limitado de vehículos, que haga que ante cada llamada no necesariamente puede existir un vehículo disponible para llegar al lugar de destino dentro del tiempo máximo reglamentado por tipo de llamada.
3. Siendo mucho mayor el número de llamadas de consulta, en relación con las de emergencia e urgencia, los tiempos máximos esperados de respuesta permiten la acumulación de llamadas en cola sin que necesariamente se incumpla con el tiempo máximo esperado de respuesta, en consecuencia cuando el operador recibe una llamada existen otras tantas pendientes de atención.

Pudiera pensarse en separar el sistema de asignación de vehículos a llamadas de emergencia médica Tipo 3 (Consultas) de las otras dos; sin embargo, ello le restaría posibilidades al sistema, como la que un vehículo tipo TAM ó TAB atienda una llamada de consulta, dado el caso que ante la cantidad de llamadas de consulta a atender en horas de demanda alta, todos los vehículos de Consulta VC se encuentren ocupados y sin posibilidad de llegar al lugar de atención del paciente dentro de los límites establecidos por el tiempo máximo de respuesta.

Una forma de incrementar el nivel de servicio es expandir la flota de vehículos; sin embargo esta decisión requiere niveles de inversión altos basados en decisiones estratégicas. La empresa requiere primero optimizar el uso de la flota existente alcanzando niveles de servicio superiores, para decidir la necesidad del incremento de su flota de vehículos.

3. Solución actual del problema

Dada la complejidad de las decisiones de asignación y/o despacho y la dificultad para evaluar el impacto de cada una de ellas en las siguientes, el operador toma las decisiones sobre la base de una de 3 políticas definidas por la compañía.

- Política 1: Asignar la llamada y/o despachar el vehículo que se encuentre en la localización a menor tiempo de desplazamiento de la localización de atención de la llamada entrante, independientemente de su estado (Ocupado o disponible).
- Política 2: Asignar la llamada y/o despachar el vehículo con menor tiempo de respuesta, es decir el vehículo que independientemente de su estado (Ocupado o disponible) puede llegar a la localización de atención de la llamada entrante en el menor tiempo posible.
- Política 3: Aplica el mismo principio de la Política 2 dando cierta preferencia a los vehículos con menor capacidad, entendiéndose por ello que pueden atender menos tipos de servicio.

La Política 1 se enfoca en la eficiencia buscando minimizar la suma de los tiempos de desplazamiento que invierten los vehículos en la atención de las llamadas, la Política 2 busca la eficacia local al tratar de minimizar los tiempos de respuesta a cada llamada, la Política 3 se

enfoca en la eficacia global incorporando un elemento que considera las llamadas futuras al dar preferencia a los vehículos que pueden atender menos tipos de servicio.

4. Metodología propuesta

El problema que tiene un carácter dinámico se plantea resolverlo de forma determinístico, modelándolo como un caso particular de MDVRPTW, modificando los parámetros y datos del problema cada vez que ocurra algún cambio en el estado del sistema, tales como recepción de una nueva llamada, avería de un vehículo, etc. El MDVRPTW planteado se formula como un modelo de Programación Lineal Entera Mixta.

El caso particular de tipo MDVRPTW al que responde el problema propuesto considera que la cantidad de depósitos coincide con la cantidad de vehículos disponibles en el sistema y cada depósito dispone exactamente de un solo vehículo. Cada vehículo realizará máximo un *tour* por las localizaciones de atención de las llamadas aún por atender regresando de manera ficticia al lugar de origen, pues el impacto en la función a optimizar será solo hasta la última localización visitada, desde donde comenzará su nuevo *tour* en la próxima aplicación del modelo por cambios en el sistema. El realizar un *tour* “cerrado” pretende mantener la condición de que un vehículo salga desde toda localización de atención, lo cual no se cumpliría al llegar a la última localización visitada.

Cada vez que el modelo se resuelve se considera un conjunto de llamadas no atendidas aún, asociadas con su nivel de prioridad clasificado por el sistema en los tipos de llamada Emergencia, Urgencia y Consulta y el tiempo transcurrido desde que fueron realizadas. Se conoce la localización de los diferentes tipos de vehículo, que son las posiciones en que culminaron o deberían encontrarse según la ejecución anterior del modelo, las ventanas de tiempo (sin límite inferior) para atender una llamada en una localización dada, y horarios de disponibilidad de los vehículos. La solución de cada MDVRPTW es un *tour* de cada vehículo disponible a un conjunto de las localizaciones que responden a las llamadas de emergencia de forma tal de no violar las restricciones del sistema. Se busca asignar todas las llamadas, utilizando el vehículo más adecuado para cada una de ellas, invirtiendo el menor tiempo de desplazamiento posible, y reduciendo el tiempo de inicio de cada *tour*.

5. Modelo matemático propuesto para la asignación de vehículos

El modelo propuesto, que a continuación se presenta, se formula con elementos de cada una de las políticas actuales de la compañía, buscando integrar las fortalezas de cada una de ellas.

Conjuntos:

$V = \{1, 2, \dots, m\}$ Conjunto de m números naturales que representan los vehículos de la flota.

D : Conjunto de m números naturales que representan las localizaciones de los vehículos en el momento de resolverse el modelo de despacho, es decir, D_k es la localización del vehículo k en el momento de resolverse el modelo de despacho.

$C = \{1, 2, \dots, c\}$ Conjunto de c números naturales que representan las localizaciones de los lugares de atención de cada una de las llamadas por atender en el momento de resolverse el modelo de despacho.

$N = D \cup C$, conjunto de todas las localizaciones, tanto de los lugares de atención de las llamadas por atender como del lugar donde se encuentran los vehículos (llamadas “seudobases”).

Parámetros:

m : número total de vehículos disponibles.

c : cantidad de llamadas pendientes de atención en el momento de resolverse el modelo de despacho.

T_{ij} : Número natural que representa el tiempo esperado de desplazamiento (en segundos) de cualquiera de los vehículos desde la localización i a la j .

L_j : Representa el tipo de servicio que requiere la llamada a atenderse en la localización $j \in C$.

Q_k : Representa el tipo de vehículo $k, k \in V$

E_i : Número entero que representa el tiempo esperado de permanencia (en segundos) de cualquiera de los vehículos en la localización $i, i \in N$.

A_j : Número entero que representa el instante de tiempo máximo (en segundos) en el que cualquiera de los vehículos deberá arribar a la localización $j, j \in C$.

B_k y W_k : Números enteros que representan los instantes de tiempo (en segundos) desde y hasta que el vehículo k está en turno para atender llamadas, $k \in V$.

M : Número entero (en segundos) que representa una cota superior adecuada, para condicionar los valores que toma la variable s_{ik} .

O : Número natural, en segundos, que representa el instante de tiempo en el cual se resuelve el modelo de despacho.

Los valores que tomarán L_j y Q_k , cualesquiera sean las localizaciones j y vehículos k , se muestran en la Tabla 1 y responden a la política de la empresa. Para el tipo de vehículo, mientras el valor es mayor, puede atender más tipos de servicio. En el caso del tipo de servicio, mientras el número es mayor, el tipo de servicio es más prioritario. Siempre debe cumplirse que Q_k debe ser mayor o igual a L_j , para una llamada en la localización j .

Tipo de vehículo	Valor de Q	Tipo de servicio	Valor de L
TAM	3	Emergencia	3
TAB	2	Urgencia	2
VC	1	Consulta	1

Tabla 1: Valores posibles por tipo de vehículo y servicio

Índices usuales:

i : Localizaciones tanto de las seudobases como de los lugares de atención de las llamadas, es decir los elementos que pertenecen a N .

j : Localizaciones de los lugares de atención de las llamadas, es decir los elementos que pertenecen a C .

k : Vehículos del sistema, es decir los elementos que pertenecen a V .

Variables:

Se definen dos tipos de variables, binarias y enteras:

x_{ijk} : binaria, toma el valor 1 si se realiza un desplazamiento del vehículo k desde la localización i hasta la localización j , 0 de lo contrario.

s_{ik} : entera, representa el instante de tiempo (en segundos) en que el vehículo k se encontrará en la localización i

Modelo matemático

$$\min z = \sum_{i \in N} \sum_{j \neq i \in C} \sum_{k \in V} T_{ij} * x_{ijk} * Q_k + \frac{1}{O} \sum_{j \in C} \sum_{k \in V} s_{jk} \quad (0)$$

Restricciones de demanda:

$$\sum_{i \neq j \in N} \sum_{k \in V} x_{ijk} = 1 \quad \forall j \in C \quad (1)$$

Restricciones de cumplimiento de los tours:

$$\sum_{i \neq h \in N} x_{ihk} - \sum_{j \neq h \in N} x_{hjk} = 0 \quad \forall h \in C \quad \forall k \in V \quad (2)$$

$$\sum_{j \in C} x_{d(k)jk} \leq 1 \quad \forall k \in V \quad (3)$$

$$\sum_{j \in C} x_{d(k)jk} = \sum_{j \in C} x_{jd(k)k} \quad \forall k \in V \quad (4)$$

Restricciones de política de uso de los vehículos:

$$\sum_{i \neq j \in N} L_j * x_{ijk} \leq Q_k \quad \forall j \in C \quad \forall k \in V \quad (5)$$

Restricciones de satisfacción de ventanas de tiempo:

$$s_{d(k)k} = O \quad \forall k \in V \quad (6)$$

$$s_{ik} + E_i + T_{ij} - M*(1-x_{ijk}) \leq s_{jk} \quad \forall i \neq j \in N, \forall j \in C, \forall k \in V \quad (7)$$

$$s_{jk} \leq A_j \quad \forall j \in C, \forall k \in V \quad (8)$$

Restricciones de disponibilidad de vehículos

$$B_k \leq s_{jk} \leq W_k \quad \forall j \in C, \forall k \in V \quad (9)$$

La función objetivo consta de dos términos, que responden a dos objetivos a lograr. El primer término, el más relevante, se construye a partir de la Política 1 cuyo objetivo fundamental a lograr es atender todas las llamadas invirtiendo el menor tiempo posible, pero a la vez utilizando el vehículo más apropiado. Así, por ejemplo, asumiendo iguales tiempos de desplazamiento, se penaliza más enviar a una localización que requiere un servicio de consulta una Ambulancia de Transporte Asistencial Medicalizado (TAM) que una Ambulancia de Transporte Asistencial Básico (TAB), y ésta a su vez se penaliza más que enviar un Vehículo de Consulta (VC). El segundo término tiene en cuenta el mismo principio de la Política 2, buscando reducir el tiempo de inicio de cada tour, considerando que entre menor sea el instante de tiempo en el cual son atendidas las llamadas, menor será el costo para el sistema. Una relación adecuada entre cada término s_{jk} y O , le da un peso mucho menor al segundo término en la función a optimizar que al primero. Otra expresión que pudiera pensarse es por ejemplo $s_{jk}-O$, pero equipararía ambos objetivos.

Observar que aunque cada tour realizado por los diferentes vehículos regresa de manera “ficticia” al lugar de donde partió, ello no se tiene en cuenta en la función objetivo.

Las restricciones (1) obligan a que todas las llamadas pendientes de atención sean atendidas una única vez, por cualquiera de los vehículos ubicados en cualquier localización, sea su seudobase (donde se encuentra al inicio) ó una que corresponde a la de una llamada que ya atendió.

Las restricciones (2) aseguran que cada vehículo que visite un lugar de atención de una llamada, inicie un nuevo desplazamiento desde ese mismo lugar.

Las restricciones (3) y (4), para mantener dicho balance, obligan a que cada vehículo que salió de su seudobase hacia alguna localización, regrese a ella.

Las restricciones (5) aseguran que una localización no sea visitada por un tipo de vehículo que no tiene las condiciones para ello, de acuerdo a los valores L_j y Q_k . En la Tabla 2 se muestran las posibilidades existentes.

Tipo de Vehículo	Tipo de Servicio		
	Emergencia (3)	Urgencia (2)	Consulta (1)
TAM (3)	Si	Si	Si
TAB (2)	No	Si	Si
VC (1)	No	No	Si

Tabla 2: Prestación de servicio por tipo de llamada y vehículo

Las restricciones de tipo (6) aseguran que el instante de tiempo en el cual cada vehículo se encuentra en su seudobase sea igual al parámetro O , el cual representa el instante en que se ejecuta el modelo de despacho. Esto garantiza correspondencia entre los tiempos reales y del modelo.

Las restricciones de tipo (7) garantizan que el instante s_{jk} en el cual el vehículo k atiende una de las llamadas j pendientes sea mayor o igual, a la suma del instante s_{ik} , en que el vehículo k se encuentra en la localización i , el tiempo esperado de permanencia E_i en la localización i y el tiempo esperado de desplazamiento T_{ij} , desde la localización i hasta la localización j . Un vehículo no puede atender una llamada, sin que se hayan cumplido los tiempos esperados de atención y desplazamiento estipulados.

Las restricciones (8) imponen que el instante s_{jk} en el cual el vehículo k atiende una de las llamadas j pendiente no sea superior, al instante máximo A_j , en el cual debe atenderse la llamada.

Finalmente, las restricciones de tipo (9) aseguran que el instante s_{jk} , en el cual el vehículo k atiende una de las llamadas j pendientes, no sea inferior al instante B_k desde que el vehículo está en turno para atender llamadas, ni superior al instante W_k , hasta que el vehículo está en turno para atender llamadas.

El que cada vehículo que sale de su seudobase no vaya a otra seudobase se asegura con los valores dados a los T_{ij} .

5.1 Modelo matemático “suave”

Existe la posibilidad que el sistema no pueda cumplir con los tiempos máximos esperados de respuesta de todas las llamadas; es decir, que en algunos casos deberán incumplirse las restricciones de tipo (8), para encontrar una solución factible. Esto equivale a considerar ventanas de tiempo suaves. Para tal fin se crea una variable de holgura r_{jk} y se ajusta la restricción (8), como la diferencia entre el instante en el cual el vehículo k llega a la localización de atención j y el tiempo máximo esperado de respuesta para la llamada $j \in C$.

$$s_{jk} - r_{jk} \leq A_j \quad \forall j \in C, \forall k \in V \quad (8')$$

Mientras s_{jk} sea menor que A_j , r_{jk} no requiere tomar ningún valor; en caso de que deba incumplirse la restricción, r_{jk} toma el valor del tiempo de incumplimiento.

Teniendo en cuenta que no es deseable incumplir con el tiempo máximo esperado de respuesta, si se presenta, el impacto debe ser alto y acorde con el tipo de servicio. Para tal fin se incluye un tercer término en la función objetivo, que penaliza en términos representativos, el incumplimiento de los tiempos máximos esperados de respuesta, de acuerdo a los valores de la variable de holgura r_{jk} .

$$\min z = \sum_{i \in N} \sum_{j \neq i \in C} \sum_{k \in V} T_{ij} * x_{ijk} * Q_k + \frac{1}{O} \sum_{j \in C} \sum_{k \in V} s_{jk} + 60 \sum_{j \in C} \sum_{k \in V} L_j * r_{jk} \quad (0')$$

Considerando el método basado en la asignación de prioridades, la importancia relativa del tercer término, es significativamente mayor que la del primero. Incumplir en un minuto el tiempo máximo esperado de respuesta, equivale a invertir una hora más en tiempo de desplazamiento. Según lo expuesto acerca de la importancia relativa de los términos de la función objetivo, la prioridad de objetivos en ese orden es: tercer, primer y segundo término de la función.

6. Diseño de experimentos

El objetivo del diseño de experimentos es comparar el desempeño del mejor de los modelos propuestos con las políticas de despacho actuales. Para ello se desarrollaron tres algoritmos heurísticos que permiten simular las soluciones que se obtienen siguiendo cada una de esas políticas. Se identificó un intervalo de tiempo representativo con la mayor cantidad de llamadas recibidas (33) en un rango de tiempo de dos horas, en el primer semestre del año 2011 en la ciudad de Cali. Con estos datos se derivaron 33 instancias, que corresponden al estado del sistema con cada nueva llamada entrante. Para estas instancias se corrieron los 3 algoritmos heurísticos y se obtuvieron además las soluciones de los modelos propuestos con el CPLEX versión 10.2.

Los valores esperados de los tiempos de desplazamiento y de atención se estimaron a partir de 22.525 registros del sistema de información de la compañía, correspondientes a la totalidad de llamadas recepcionadas en el primer semestre del año 2011 en la ciudad de Cali. Para los casos en que no existían datos para estimar los tiempos de desplazamiento entre barrios, se definió un valor de 10.800 segundos, el cual excede el tiempo máximo de respuesta de los tres tipos de servicio. Los centroides de los barrios de la ciudad de Cali, representaron las posibles localizaciones. Las localizaciones iniciales de los vehículos se definieron aleatoriamente entre los barrios con mayor demanda de atenciones y no a partir de un modelo de relocalización.

Se consideraron 12 vehículos totales disponibles. Identificados los barrios con mayor demanda, se generó un número aleatorio entre 1 y 20 para cada uno de los 12 vehículos y se asignó la localización en el centroide del barrio a cada uno de ellos. Cuando un vehículo finaliza una atención, y no tiene alguna llamada pendiente por atender, permanece en el barrio donde finalizó la atención.

Se compararon entre sí las soluciones obtenidas por ambas variantes del modelo propuesto. Los resultados de las tres políticas de la compañía y las del modelo “suave” (por ser el de mejor desempeño como se muestra en la Sección 6.1) se evaluaron por cinco criterios:

- Cumplimiento en tiempo de respuesta: Porcentaje del total de llamadas atendidas, en un tiempo menor o igual al tiempo máximo esperado de respuesta.
- Nivel de Incumplimiento del tiempo de respuesta: Teniendo en cuenta que el nivel de respuesta definido mide frecuencia más no impacto del incumplimiento, se planteó una medición del mismo como la suma de los productos del tiempo de incumplimiento por el valor asignado al tipo de servicio, de cada una de las llamadas.
- Tiempo de Ocupación de los vehículos: Tiempo total invertido por los vehículos para atender la totalidad de las llamadas (suma de los tiempos de desplazamiento y de los tiempos de atención de cada una de las llamadas).
- Nivel de ocupación de la flota: Suma del promedio de vehículos utilizados con la desviación estándar de los datos, en cada segundo del intervalo crítico de utilización de los vehículos dentro del intervalo total de análisis. Esto es para ver cómo la variabilidad de los datos impacta la medida de desempeño.
- Nivel de *Preparedness*: Resta del promedio del valor que corresponde al tipo de los vehículos no utilizados (Q_k) y la desviación estándar de los datos, en cada segundo del intervalo crítico de utilización de los vehículos dentro del intervalo total de análisis. Con esto se refleja la importancia no solo de los vehículos no utilizados, sino de su capacidad de atención de llamadas.

Los algoritmos heurísticos y los modelos fueron implementados en un Notebook HP G42-363LA con procesador Core i3-370M y 3GB de RAM.

6.1 Soluciones obtenidas

La Tabla 3 muestra los resultados de la aplicación de las dos variantes del modelo a las instancias donde ambos modelos obtuvieron soluciones factibles. De las 33 instancias, en 13 de ellas (de la 6 a la 18) solo se obtuvieron soluciones con el modelo suave. En todos los casos se cumple que $1 V 1 = 1 D 1 = 12$. Se diferencian solo los tiempos de ejecución del modelo y Modelo Suave (MS), dados en segundos, pues el valor de la función a optimizar es la misma en ambos casos.

Inst.	C	Tiempo	Tiempo MS	F.O.	Inst.	C	Tiempo	Tiempo MS	F.O.
1	1	0,062	0,078	985,7	24	14	3,214	27,519	7551,09
2	2	0,094	0,109	1611,2	25	15	14,461	2,200	7380,06
3	3	0,125	0,094	1947,9	26	16	65,458	56,223	7393,04
4	4	0,156	0,140	2184,2	27	15	19,672	26,286	7347,86
5	5	0,172	0,125	2840,1	28	12	8,335	17,800	6988,96
6	6	0,234	0,328	2510,9	29	13	26,099	44,102	7424,84
19	9	0,250	0,343	5664,6	30	14	25,303	64,740	7630,19
20	10	0,312	0,421	5597,7	31	14	2518,310	324,670	8527,97
21	11	0,421	0,452	5905,6	32	14	16,099	2,012	7922,25
22	12	0,499	0,530	7292,9	33	15	7,878	19,110	8325,22
23	13	1,108	13,557	7360,4					

Tabla 3: Comparación de resultados de instancias con soluciones factibles en ambos modelos

Las Tablas 4 a 8 resumen los resultados obtenidos según los criterios definidos anteriormente (5):

Criterio de Despacho	Llamadas Atendidas a Tiempo	Total de Llamadas	Cumplimiento en tiempo de respuesta
Política 1	29	33	87,88%
Política 2	31	33	93,94%
Política 3	31	33	93,94%
Modelo	32	33	96,97%

Tabla 4: Cumplimiento en tiempo de respuesta

Criterio de Despacho	Llamada Incumplida	Tipo de Servicio	Tiempo de Incumplimiento	Impacto por llamada	Nivel Incumplimiento
Política 1	19	2	1.450	2.900	6.578
	25	1	1.697	1.697	
	39	1	88	88	
	41	1	1.893	1.893	
Política 2	19	2	627	1.254	19.254
	34	2	9.000	18.000	
Política 3	19	2	627	1.254	21.152
	34	2	9.949	19.898	
Modelo	19	2	627	1.254	1.254

Tabla 5: Nivel de Incumplimiento del tiempo de respuesta

Criterio de Despacho	Tiempo Total de Desplazamiento	Tiempo Total de Atenciones	Tiempo Total Invertido
Política 1	23.828	49.463	73.291
Política 2	38.758	49.463	88.221
Política 3	40.368	49.463	89.831
Modelo	23.717	49.463	73.180

Tabla 6: Tiempo de ocupación de los vehículos

Criterio de Despacho	Promedio	Desviación Estándar	Nivel de Ocupación de Flota
Política 1	7,58	1,14	8,71
Política 2	9,81	1,42	11,23
Política 3	9,31	1,25	10,56
Modelo	7,20	1,01	8,21

Tabla 7: Nivel de desempeño de ocupación de la flota

Modelo de Despacho	Promedio	Desviación Estándar	Nivel de Preparedness
Política 1	8,20	1,79	6,42
Política 2	3,81	3,26	0,55
Política 3	4,78	1,86	2,92
Modelo	10,53	1,21	9,32

Tabla 8: Nivel de *preparedness* del sistema

6.2 Análisis de Resultados

En la comparación de ambas variantes del modelo (Tabla 3) el número de llamadas pendientes de atención en dichas instancias osciló entre 1 y 16. El modelo “suave” es más robusto por encontrar soluciones en todos los casos, y en tiempos de solución, cuando eran comparables las dos variantes del modelo, en general los tiempos fueron inferiores. La única instancia con valores atípicos en tiempo de solución con ambas variantes del modelo fue muy superior con el modelo original. El modelo “suave” obtuvo la solución en menos de 1 segundo en 11 instancias (33% de las mismas), en menos de 1 minuto en 28 instancias (85% de las mismas), entre 1 y 2 minutos en 4 instancias (12% de las mismas) y en una sola instancia (3%) lo obtuvo entre 5 y 6 minutos.

En lo adelante (igual que en las Tablas 4 a la 8) al referirse al modelo se asume la versión suave del mismo.

Según los resultados de los cinco criterios reflejados en las Tablas 4 a la 8 se puede concluir que el modelo propuesto arrojó en todos los casos los mejores resultados. De las políticas actuales de la compañía no hay ninguna que sea la “segunda mejor” en todos los criterios, la que más se acerca es la Política 1 salvo en un criterio, en cumplimiento en tiempo de respuesta (Tabla 4).

El modelo propuesto es más adecuado que las políticas actuales aplicadas por la compañía, por 3 razones fundamentales:

- Incluye las fortalezas de las 3 políticas.
- Permite reasignar llamadas pendientes con cada ejecución del modelo a partir de cada estado del sistema.
- Da soluciones exactas en un tiempo razonable.

Dos de los resultados gráficos de los criterios señalados anteriormente se reflejan en las Figuras 1 y 2. El eje de las abscisas en ambas figuras representa tiempos en segundos a partir del instante inicial. El eje de las ordenadas en la Figura 1 es el número de vehículos disponibles, en el caso de la Figura 2 es el nivel del *preparedness*.

En la Figura 1 se observa que las Políticas 2 y 3 tienen instantes de tiempo donde todos los vehículos se encuentran ocupados. La máxima cantidad de vehículos ocupados en un mismo instante de tiempo, por la Política 1 y el modelo propuesto, son 10 y 8, respectivamente.

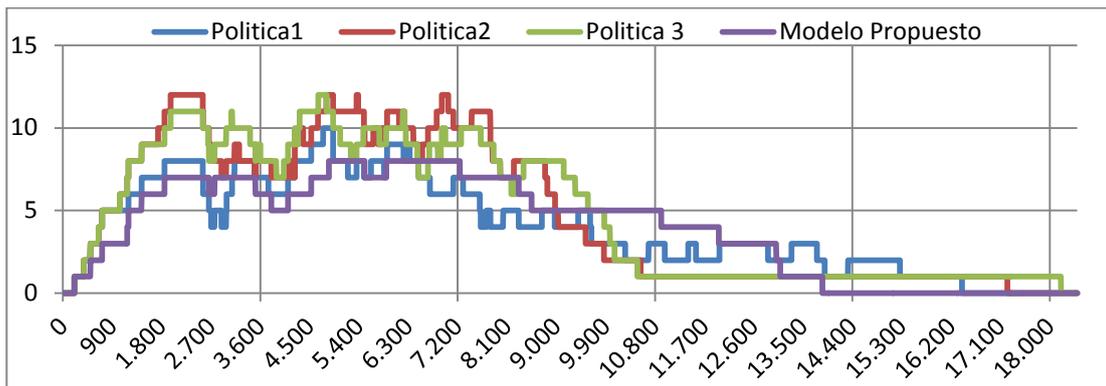


Figura 1: Ocupación de la flota de vehículos

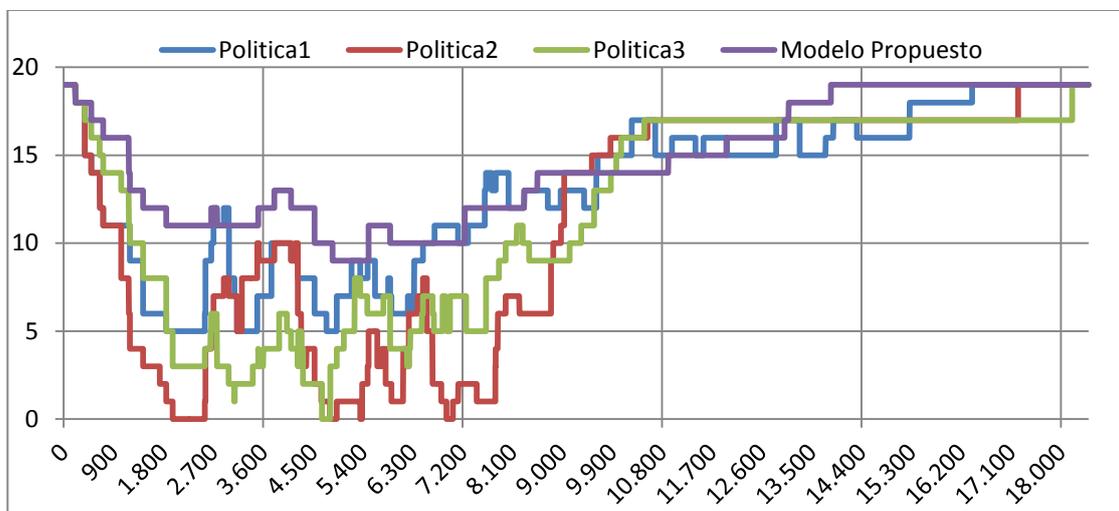


Figura 2: *Preparedness* del sistema

En la Figura 2 se observa que las Políticas 2 y 3 tienen instantes de tiempo donde el nivel de *preparedness* es 0. El mínimo nivel de *preparedness* en un instante de tiempo dado, para la Política 1 y el modelo propuesto, son 5 y 9, respectivamente.

7. Conclusiones

Se desarrollaron dos versiones de un modelo matemático de Programación Lineal Entera Mixta, para el problema de despacho de vehículos en servicios de emergencia médica de una compañía prestadora de ese servicio en Colombia, que se ajustan a las restricciones impuestas por dicha compañía. La versión “suave” se introduce ante la posibilidad que el sistema no pueda cumplir con los tiempos máximos esperados de respuesta de todas las llamadas. Los modelos corresponden a un caso particular de problema de ruteo de tipo MDVRPTW. La función objetivo definida incorpora las fortalezas de las 3 políticas de operación actuales usadas por la compañía.

De acuerdo al diseño de experimentos realizado donde se comparan las dos versiones del modelo propuesto y la mejor de ellas con las 3 políticas actuales de la empresa (evaluadas mediante la implementación de 3 algoritmos heurísticos que las simulan) se obtienen dos conclusiones fundamentales:

1. De las dos versiones del modelo propuesto, la aplicación del modelo “suave” es más realista para su implementación, debido al número de infactibilidades encontradas en un representativo

número de instancias (12 de 33, el 33%) con el modelo original, independientemente de que es más robusto en tiempos de solución y obtiene las mismas soluciones que el modelo original cuando éste da soluciones factibles.

2. El modelo “suave” propuesto resulta más adecuado que cada una de tres políticas de despacho utilizadas por la compañía actualmente, según cinco criterios de desempeño planteados.

8. Trabajo futuro

Los tiempos esperados de desplazamiento de los vehículos desde una a otra localización dependen de factores tales como: mes del año, día de la semana, día feriado o no, rango de tiempo del día en el cual se está ejecutando el modelo de despacho así como el tipo de vehículo. Se deben estimar estos valores usando un modelo de pronósticos.

Integrar el modelo de despacho con un modelo de relocalización. Una vez que sea ejecutado el modelo de despacho, y se toma la decisión de despachar algún vehículo, el modelo de relocalización validaría si la relocalización de los vehículos pudiera maximizar la capacidad de respuesta del sistema. Si se toma la decisión de relocalizar, las nuevas localizaciones deben ser tenidas en cuenta para una nueva ejecución del modelo de despacho.

Incluir el servicio de traslado de pacientes teniendo en cuenta los desplazamientos hasta el lugar de recogida y atención del paciente, el posterior desplazamiento hasta el destino del paciente y los tiempos de entrega del paciente.

Ampliar el diseño de experimento realizado, con diversos intervalos de tiempo y días de despacho.

9. Referencias bibliográficas

- Alba, E. and Dorronsoro, B. (2004), Solving the Vehicle Routing Problem by Using Cellular Genetic Algorithms, *EvoCOP, 2004, LNCS, Vol. 3004, 11-20, Portugal, Springer-Verlag.*
- Andersson, T., Petersson, S., Värbrand, P. (2004) Calculating the preparedness for an efficient ambulance health care, *Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems, Page 785-790, ISBN: 0-7803-8500-4.*
- Andersson, T., Petersson S., and Varbrand, P. (2006) “Dynamic Ambulance Relocation for a Higher Preparedness”, *Proceedings of DSI 2004 : the 35th Annual Meeting of the Decision Sciences Institute , November 20-23, Boston, USA, 2004. ISBN: 0-9667118-0-7*
- Andersson, T. and Varbrand, P. (2007). “Decision support tools for ambulance dispatch and relocation”, *Journal of the Operational Research Society Vol. 58(2), pp. 195–201.*
- Amberg, A., Domschke, W. and Voss, S. (2000), Multiple Center Capacitated Arc Routing Problems: A Tabu Search Algorithm using Capacitated Trees”, *European Journal of Operational Research, Vol.124, 360-376.*
- Cordeau, J.F., Desaulniers, G., Desrosiers, J., Solomon, M.M. and Soumis, F. (2002), *VRP with Time Windows, The Vehicle Routing Problem, SIAM Monographs on Discrete Mathematics and Applications, Vol. 9, 157-193.*
- Czarnas, P, Czech, Z.J., and Gocyla, P. (2004), *Parallel Simulated Annealing for Bi-criterion Optimization Problems, PPAM 2003, LNCS 3019, Springer-Verlag, 233-240.*
- Lau, H.C. , Sim, M. and Teo, K.M.(2003), *Vehicle Routing Problem with Time Windows and a Limited Number of Vehicles, European Journal of Operational Research, Vol.148, 559-569.*
- Lee, S. (2011), *The Role of Preparedness in Ambulance Dispatching, Journal of the Operational Research Society, Vol. 62, Issue 10, pp. 1888-1897.*
- Polacek, M., Hartl R.F. and Doerner, K. (2004), *A variable neighborhood search for the multi depot vehicle routing problem with time windows, Journal of Heuristics, Vol.10, No.6, 613-627.*