

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICAS

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE
OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE
TRABAJO.

PROYECTO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL
TÍTULO DE MÁSTER EN CONTROL DE OPERACIONES Y
GESTIÓN LOGÍSTICA

POR:

Ing. Livino Manuel Armijos Toro

DIRIGIDO POR:

Dr. Fernando Francisco Sandoya Sánchez

2 0 1 5

Dedicatoria

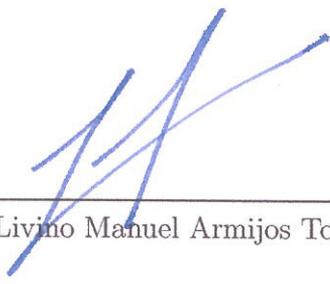
A mis padres y a Fabiola, por estar siempre a mi lado.

Agradecimientos

Primero quiero agradecer a Fernando por su guía, a mis padres por su comprensión, estímulo, apoyo y amor. A Fabiola por saber entenderme.

Declaración Expresa

La responsabilidad por los hechos y doctrinas expuestas en esta tesis de Graduación, me corresponde exclusivamente; el patrimonio intelectual del mismo, corresponde exclusivamente a la Facultad de Ciencias Naturales y Matemáticas, Departamento de Matemáticas de la Escuela Superior Politécnica del Litoral.



Ing. Livino Manuel Armijos Toro

TRIBUNAL DE GRADUACIÓN



Mgs. Carlos Martín Barreiro
PRESIDENTE DEL TRIBUNAL



Dr. Fernando Sandoya Sancho
DIRECTOR DE LA TESIS



Mgs. Guillermo Baquerizo Palma
VOCAL DEL TRIBUNAL

Índice general

Lista de figuras	VII
Lista de tablas	VIII
Introducción	3
1. Medidas de distancia y diversidad	6
1.1. Diversidad entre elementos	6
1.2. Medidas de diversidad	8
1.3. El problema de la diversidad máxima	9
1.4. Algoritmos de solución	9
1.5. GRASP	11
1.5.1. Algoritmo voraz	12
1.5.2. Mecanismo de generación de soluciones voraces aleatorias . . .	12
1.5.3. Técnica de búsqueda local	12
1.6. Análisis de conglomerados	12
1.6.1. Definición de la técnica: conglomerado jerárquico y aglomerativo	13
1.6.2. Condiciones de aplicación del análisis de conglomerados . . .	14
1.6.3. Procedimiento	15
2. Técnicas metaheurísticas y algoritmos basados en GRASP	19
2.1. Descripción de la metaheurística GRASP	19
2.2. Funcionamiento del algoritmo	21
2.3. Desarrollo experimental	24
2.3.1. Conformación de los equipos de trabajo	30
3. Uso del análisis de conglomerados en la conformación de equipos de trabajo	37
3.1. Ambato	38
3.2. Cuenca	39
3.3. Guayaquil	40
3.4. Machala	41
3.5. Quito	43
4. Conclusiones	46
5. Recomendaciones	49

A. Anexo de bases de datos y resultados	51
A.1. Ambato	51
A.2. Cuenca	54
A.3. Guayaquil	57
A.4. Machala	60
A.5. Quito	63

Índice de figuras

2.1.	Características promedio de los candidatos de la regional Ambato. . .	26
2.2.	Características promedio de los candidatos de la regional Cuenca. . .	27
2.3.	Características promedio de los candidatos de la regional Guayaquil. .	28
2.4.	Características promedio de los candidatos de la regional Machala. . .	29
2.5.	Características promedio de los candidatos de la regional Quito. . . .	30
2.6.	Capacidad del equipo uno de la regional Ambato.	31
2.7.	Capacidad del equipo dos de la regional Cuenca.	32
2.8.	Capacidad del equipo tres de la regional Guayaquil.	33
2.9.	Capacidad del equipo cuatro de la regional Machala.	34
2.10.	Capacidad del equipo quinto de la regional Quito.	35
3.1.	Resultados de las comparaciones de la regional Ambato.	38
3.2.	Resultados de las comparaciones de la regional Cuenca.	39
3.3.	Resultados de las comparaciones de la regional Guayaquil.	41
3.4.	Resultados de las comparaciones de la regional Machala.	42
3.5.	Resultados de las comparaciones de la regional Quito.	43

Índice de cuadros

1.1. Clasificación de las principales medidas de distancia, según el tipo de variable.	15
1.2. Clasificación de los métodos de agrupamiento.	17
2.1. Características a ser evaluadas.	23
2.2. Función objetivo para dos elementos.	23
2.3. Función objetivo para tres elementos.	24
2.4. Características de las variables de la regional Ambato.	25
2.5. Características de las variables de la regional Ambato.	25
2.6. Características de las variables de la regional Cuenca.	26
2.7. Características de las variables de la regional Cuenca.	26
2.8. Características de las variables de la regional Guayaquil.	27
2.9. Características de las variables de la regional Guayaquil.	27
2.10. Características de las variables de la regional Machala.	28
2.11. Características de las variables de la regional Machala.	28
2.12. Características de las variables de la regional Quito.	29
2.13. Características de las variables de la regional Quito.	29
2.14. Primer grupo regional Ambato.	31
2.15. Segundo equipo regional Cuenca.	32
2.16. Tercer equipo regional Guayaquil.	33
2.17. Cuarto equipo regional Machala.	34
2.18. Quinto equipo regional Quito.	35
3.1. Diversidad Ambato.	38
3.2. Intervalos para la diversidad obtenida.	39
3.3. Diversidad Cuenca.	39
3.4. Intervalos para la diversidad obtenida.	40
3.5. Diversidad Guayaquil.	40
3.6. Intervalos para la diversidad obtenida.	41
3.7. Diversidad Machala.	42
3.8. Intervalos para la diversidad obtenida.	42
3.9. Diversidad Quito.	43
3.10. Intervalos para la diversidad obtenida.	44
A.1. Características postulantes regional Ambato	51
A.1. Características postulantes regional Ambato	52
A.1. Características postulantes regional Ambato	53
A.2. Grupo de cinco equipos de trabajo formado utilizando el GRASP. . .	53
A.3. Equipos formados con análisis de conglomerados.	54
A.4. Características postulantes regional Ambato	54

A.4. Características postulantes regional Ambato	55
A.4. Características postulantes regional Cuenca	56
A.5. Grupo de cinco equipos de trabajo formado utilizando el GRASP. . .	56
A.6. Equipos formados con análisis de conglomerados.	57
A.7. Características postulantes regional Guayaquil	57
A.7. Características postulantes regional Guayaquil	58
A.7. Características postulantes regional Guayaquil	59
A.8. Grupo de cinco equipos de trabajo formado utilizando el GRASP. . .	60
A.9. Equipos formados con análisis de conglomerados.	60
A.10.Características postulantes regional Machala	61
A.10.Características postulantes regional Machala	62
A.11.Grupo de cinco equipos de trabajo formado utilizando el GRASP. . .	62
A.12.Equipos formados con análisis de conglomerados.	63
A.13.Características postulantes regional Machala	63
A.13.Características postulantes regional Machala	64
A.13.Características postulantes regional Quito	65
A.14.Grupo de cinco equipos de trabajo formado utilizando el GRASP. . .	66
A.15.Equipos formados con análisis de conglomerados.	66

Objetivo General

La presente tesis busca utilizar el problema de la diversidad máxima, en la variante que más se adecúe para resolver el problema de la conformación de equipos eficientes de trabajo, aplicado en este caso específicamente a la conformación de equipos de trabajo.

Objetivos Específicos

El principal objetivo de esta tesis es abordar el problema de la conformación de equipos de trabajo para el levantamiento de campo de la Encuesta Nacional de Empleo y Desempleo Urbano, ENEMDU, realizada trimestralmente por el Instituto Nacional de Estadísticas y Censos, INEC. Para el estudio de este problema se utilizará la metaheurística denominada GRASP, nombrada así por su acrónimo en inglés *Greedy Randomized Adaptive Search Procedure* (GRASP), que en castellano se puede traducir como procedimientos de búsqueda voraz, aleatorizados y adaptativos. Dentro del estudio de la metaheurística se esquematizará un algoritmo de solución para este tipo de problemas de optimización y se analizará su carácter genérico .

Con el objeto de constatar la fortaleza del método planteado en el contexto del problema en estudio se compararán los resultados obtenidos con la metaheurística GRASP, con un algoritmo alterno basado en el análisis de conglomerados.

Introducción

El proceso de selección de un subconjunto de objetos, personas, actividades, recursos, etc., desde una colección de posibilidades, basados en uno o más criterios es una actividad comúnmente realizada. Muchas veces, el objetivo de la selección es elegir a los mejores elementos, mejores en un sentido a ser definido, y en otras ocasiones, es de interés que los elementos seleccionados representen de forma efectiva la diversidad existente en la colección original de posibilidades.

En los últimos años, el interés en la explotación de los beneficios de la diversidad se ha incrementado ya que en la práctica, hay evidencia empírica que indica que la diversidad en un grupo incrementa la capacidad del grupo para resolver problemas. La construcción de grupos de trabajo es uno de estos ejemplos, pero también es uno de los principales problemas a la hora de iniciar un proyecto, ya que la conformación de estos equipos determina el éxito o fracaso de un emprendimiento, por consiguiente, la metodología utilizada para su conformación es determinante para su estructuración.

En la literatura de investigación de operaciones se han realizado trabajos en los que se demuestra que bajo ciertas hipótesis la diversidad triunfa sobre la habilidad [8] y aunque tradicionalmente se ha estudiado a la diversidad desde el punto de vista de la equidad o la representación, se puede también obtener beneficios prácticos de ella. La importancia de esta propuesta de estudio recae en la construcción de grupos de trabajo para distintas áreas de manera diversa, con la hipótesis de que a mayor diversidad de sus elementos se generarán también diferentes formas de solucionar un mismo problema; esto conlleva a que la eficiencia del grupo de trabajo en la resolución de los mismos mejore. El modelo de optimización de este tipo más estudiado es probablemente el Problema de Diversidad Máxima, notado MDP por sus siglas en inglés (Maximun Diversity Problem). Este tipo de problemas es también conocido como Problema de Max-Sum de diversidad y consiste en maximizar la medida de la diversidad que está definida como la suma de las distancias interelemento [4].

El presente trabajo busca utilizar el problema de la diversidad máxima, en la variante que más se adecúe para resolver el problema de la conformación de equipos eficientes de trabajo, aplicado en este caso específicamente a los equipos de trabajo que realizan operativos de campo en el Instituto Nacional de Estadística y Censos (INEC).

Este documento se organiza de la siguiente manera, en el capítulo uno se estudian los conceptos relacionados con distancia y medida, y como dichos conceptos son de utilidad en la generación de una noción de diversidad, se plantea el problema de la diversidad máxima y se analizan los diferentes algoritmos de solución. En el capítulo dos se plantea la problemática existente en la conformación de grupos de trabajo

para el caso específico del desarrollo de operativos de campo de la Encuesta Nacional de Empleo y Desempleo Urbano (ENEMDU). Se estudia la metaheurística GRASP, se analiza su funcionamiento, se plantea un algoritmo de solución, se ejemplifica su funcionamiento y se desarrolla la respectiva experimentación numérica. En el capítulo tres se propone un algoritmo alternativo de solución para el problema de la máxima diversidad basado en el análisis de conglomerados y se comparan los resultados obtenidos. Finalmente, en los capítulos cuatro y cinco se exponen las principales conclusiones y recomendaciones obtenidas en el desarrollo de esta tesis.

Capítulo 1

Medidas de distancia y diversidad

En este capítulo se introducen algunas definiciones y resultados sobre el análisis de la similitud y diversidad entre objetos (personas o cosas), el problema de máxima diversidad, y algunas nociones estadísticas que serán de utilidad en el desarrollo de este estudio. Además se discuten brevemente las técnicas implementadas en este trabajo.

1.1. Diversidad entre elementos

La distancia cuantifica la proximidad o lejanía de dos elementos de un conjunto cualquiera, entendiéndose que la noción de proximidad o lejanía está sujeta al problema en análisis, es por ello que por ejemplo en el estudio de las características de los individuos postulantes a formar un equipo de trabajo, la lejanía entre dos individuos se entiende como la complementariedad o diversidad de sus características estudiadas. Por lo expuesto surgió la necesidad de dar esta definición general sobre lo que podemos considerar como distancia o métrica, convergiendo en el siguiente criterio.

Definición 1.1. Sea X un conjunto cualquiera, una función $d : X \times X \rightarrow [0, +\infty[$,

se dice métrica o distancia sobre X si para todo $x, y, z \in X$ se tiene:

- $d(x, y) = 0$ si y solamente si $x = y$.
- $d(x, y) = d(y, x)$.
- $d(x, y) \leq d(x, z) + d(z, y)$.

En la literatura especializada existen varias definiciones de distancias definidas para una gran variedad de espacios, como es el caso del valor absoluto en \mathbb{R} , la distancia Euclideana en \mathbb{R}^n o la distancia del máximo en $\mathcal{C}[a, b]$. Dado que el propósito de esta sección es analizar la diversidad entre dos objetos, a continuación se muestran las distancias comunmente utilizadas para este fin [15].

Definición 1.2. Dados dos objetos x_i, x_j con las características $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$, $x_j = (x_{j1}, x_{j2}, \dots, x_{jp})^T$, se define la métrica del Coseno como:

$$d_{ij} = 1 - \frac{\sum_{k=1}^p x_{ik}x_{jk}}{\sqrt{\sum_{k=1}^p x_{ik}^2} \sqrt{\sum_{k=1}^p x_{jk}^2}} \quad (1.1)$$

En la bibliografía citada en [15], el autor explica el problema de como la diversidad presente en un grupo puede incrementar la eficiencia al resolver un problema, de hecho en su trabajo se presenta una relación que no satisface la definición formal de métrica, pues en ella se pretende determinar a través del signo, la similitud o diferencia entre las características de dos objetos, es por ello que hemos modificado dicha “métrica”, para que cuantifique solo la diferencia entre las características de dos objetos y de esta manera satisfaga la definición de métrica, de la siguiente manera

Definición 1.3. Dados dos objetos i, j con características descritas por los vectores $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$, $x_j = (x_{j1}, x_{j2}, \dots, x_{jp})^T$, se define la siguiente relación de disimilaridad entre la pareja i, j

$$d_{ij} = \frac{\sum_{l=1}^p \delta(x_{il}, x_{jl})}{p}, \quad (1.2)$$

donde

$$\delta(x_{il}, x_{jl}) = \begin{cases} 0 & \text{si } x_{il} = x_{jl} \\ |x_{il} - x_{jl}| & \text{de lo contrario.} \end{cases}$$

Se debe recalcar que una distancia o métrica es una relación u operación binaria, es decir que solo intervienen dos elementos de un mismo conjunto en dicha operación. En la siguiente sección se analizará la noción de medida, a través de la cual se puede cuantificar la diversidad de un equipo de trabajo.

1.2. Medidas de diversidad

Si bien, medir la diversidad entre dos objetos a través de una función de distancia es de gran utilidad, el objetivo de este estudio se fundamenta en cuantificar la diversidad entre un grupo de objetos, para esto se utilizará la noción de medida, por lo cual se plantea la siguiente definición.

Definición 1.4. Sea X un conjunto no vacío y \mathbb{X} una familia de subconjuntos del conjunto X , se dice σ -álgebra si:

- \emptyset, X pertenecen a \mathbb{X}
- Si A pertenece a \mathbb{X} entonces A^c pertenece a \mathbb{X}
- Si (A_n) es una sucesión de elementos de \mathbb{X} entonces $\cup_{n=1}^{\infty} A_n \in \mathbb{X}$

Una medida μ es una función a valores reales extendidos, es decir que puede tomar el valor de infinito, definida sobre una σ -álgebra \mathbb{X} de subconjuntos de X de tal manera que:

- $\mu(\emptyset) = 0$
- $\mu(E) \geq 0$ para todo $E \in \mathbb{X}$
- Si (E_n) es una sucesión de elementos de \mathbb{X} entonces $\mu(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu(E_n)$

Dados un conjunto $M = \{p_1, p_2, \dots, p_n\}$ de n objetos, con $n \in \mathbb{Z}^+$, donde p_{ik} es el atributo k -ésimo asociado al objeto p_i , con $1 \leq i \leq n$, $1 \leq k \leq l$ y $l \in \mathbb{Z}^+$ y una métrica d_{ij} definida entre cada pareja de elementos de M , basados en lo anterior se pueden definir las siguientes medidas de diversidad

Definición 1.5 (Medida de la dispersión de la suma). Esta medida se define como la suma de las distancias o disimilitudes entre cada uno de los elementos del subconjunto M , es decir:

$$div_s(M) = \sum_{p_i, p_j \in M} d_{ij}. \quad (1.3)$$

Definición 1.6 (Medida de la dispersión de la mínima distancia). Esta medida se define como la mínima distancia o disimilitud entre cada uno de los elementos del subconjunto M , es decir:

$$div_{md}(M) = \min_{p_i, p_j \in M} d_{ij}. \quad (1.4)$$

Definición 1.7 (Medida de la dispersión promedio). Esta medida se define como la distancia o disimilitud media de los elementos del subconjunto M , es decir:

$$div_{dp}(M) = \frac{\sum_{p_i, p_j \in M} d_{ij}}{|M|}, \quad (1.5)$$

donde $|M|$ denota la cardinalidad de M .

Una vez estudiadas algunas alternativas para estimar la diversidad de un subconjunto cualquiera M de un conjunto de objetos P , el siguiente paso consiste en establecer un problema de optimización cuya función objetivo busque determinar el subconjunto de P con la máxima diversidad posible. A este problema se lo denomina en la literatura como el Problema de Diversidad Máxima o simplemente MDP, como se definió anteriormente.

1.3. El problema de la diversidad máxima

Sea $X = \{x_1, x_2, \dots, x_n\}$, con $n \in \mathbb{Z}^+$, un conjunto de objetos cualquiera, y sea p_{ik} el atributo k -ésimo asociado al objeto x_i , con $1 \leq i \leq n$, $1 \leq k \leq l$ y $l \in \mathbb{Z}^+$. El problema de máxima diversidad (MDP) consiste en identificar un subconjunto M del conjunto P , con m elementos, de tal manera que los m elementos de M presenten la máxima diversidad posible. El problema MDP, si se considera la medida de dispersión de la suma, puede ser formulado de la siguiente manera:

$$\begin{aligned} \text{máx } & \sum_{i=1}^{n-1} \sum_{j=i+1}^n d_{ij} x_i x_j \\ \text{s.a. } & \sum_{i=1}^n x_i = m \\ & x_i \in \{0, 1\}; 1 \leq i \leq n, \end{aligned} \quad (1.6)$$

donde x_i se define por

$$x_i = \begin{cases} 1 & \text{si el elemento } p_i \text{ es seleccionado.} \\ 0 & \text{de lo contrario.} \end{cases} \quad (1.7)$$

El problema de diversidad máxima tiene aplicaciones como la administración de recursos humanos, medidas de biodiversidad entre otros [6].

1.4. Algoritmos de solución

En el mundo real existen multitud de problemas cotidianos que, desde un punto de vista ingenieril, precisan de una solución que cumpla un conjunto de requisitos de la manera más apropiada maximizando o minimizando determinado valor, es por ello que en este trabajo se define el concepto de optimización como el proceso de

búsqueda de la mejor solución a un problema de optimización definido de la siguiente manera:

Dado un funcional $J : H \rightarrow \mathbb{R}$ donde H es un espacio cualquiera, un problema de optimización consiste en encontrar un subconjunto R de H de manera que el funcional J alcance su valor máximo (o mínimo, dependiendo del interés del problema) para todo elemento de R . Al conjunto R se le denomina conjunto solución y a sus elementos, soluciones óptimas.

El funcional J puede estar acompañado de un conjunto de restricciones las cuales deben ser satisfechas por cada elemento de conjunto solución.

Como se menciona en [3], este tipo de problemas se pueden dividir en dos categorías: aquellos en los que la solución está codificada mediante valores reales y aquellos cuyas soluciones están codificadas por valores enteros. Dentro del segundo grupo mencionado se encuentran los problemas de optimización combinatoria. De acuerdo a la bibliografía citada en [3] este tipo de problemas consiste en encontrar una solución de un conjunto finito de posibilidades.

“Los problemas de optimización combinatoria presentan la peculiaridad que siempre existe un algoritmo exacto que permite obtener la solución óptima” [3], debido a que el tiempo de convergencia y los recursos computacionales son demasiado altos para la solución exacta del problema, se recurre al uso de algoritmos aproximados o heurísticos los cuales permiten obtener una solución de calidad en un tiempo razonable [14].

Los métodos heurísticos se pueden clasificar de la siguiente manera:

- 1 Métodos constructivos: Procedimientos que son capaces de construir una solución a un problema dado. La forma de construir la solución depende fuertemente de la estrategia seguida.
- 2 Métodos de búsqueda local: Parten de una solución factible dada y a partir de ella intentan mejorarla.

El principal problema que presentan los algoritmos heurísticos es su incapacidad para escapar de los óptimos locales. En general, ninguno de los métodos constructivos descritos en la sección anterior tendrían porque construir la solución óptima. Para solventar este problema se introducen otros algoritmos de búsqueda más inteligentes que eviten en la medida de lo posible quedar atrapados en óptimos locales. Estos algoritmos de búsqueda más inteligentes, denominados metaheurísticos, son procedimientos de alto nivel que guían a algoritmos heurísticos conocidos evitando que éstos caigan en óptimos locales.

Según J.P. Kelly et al.: “Las metaheurísticas son una clase de métodos aproximados que están diseñados para resolver problemas difíciles de optimización combinatoria, en los que los heurísticos clásicos no son efectivos. Las metaheurísticas proporcionan un marco general para crear nuevos algoritmos híbridos combinando diferentes conceptos derivados de la inteligencia artificial, la evolución biológica y los procedimientos estadísticos”.

Los métodos de búsqueda heurísticos (algoritmos metaheurísticos) como los listados abajo han sido usados para resolver problemas de este tipo.

- Búsqueda local
- Simulated annealing

- GRASP
- Swarm intelligence
- Tabu search
- Algoritmos genéticos
- Optimización basada en colonias de hormigas
- Reactive search

En esta tesis se utilizará la heurística de GRASP propuesta por Ghosh [4] para resolver el MDP.

1.5. GRASP

La metodología GRASP es una técnica iterativa, cuyo objetivo principal es resolver problemas difíciles en el campo de la optimización combinatoria, donde, en cada iteración del algoritmo se incluye una fase de construcción de una solución inicial, una de mejoramiento y una fase de actualización de la solución generada en fase inicial, es decir esta técnica construye inicialmente una solución que es mejorada inmediatamente en cada uno de las iteraciones del algoritmos. En cada iteración, los elementos $c \in C$ que no pertenecen a la solución son evaluados a través de una función $g : C \rightarrow \mathcal{R}_+$ que estima la variación al incluir este elemento en la solución parcial; estos elementos son ordenados por su valor de estimación en una lista llamada lista restringida de candidatos notada RCL por sus siglas en inglés (Restricted Candidate List) y uno de éstos es escogido aleatoriamente e incluido en la solución.

Se puede establecer una analogía con el problema de la Programación Lineal, donde primero se construye una solución factible y después se aplica el algoritmo Simplex, sin embargo, en GRASP se le da bastante importancia a la calidad de la solución generada inicialmente. La estructura básica de un algoritmo GRASP es la siguiente:

Mientras no se satisfaga el criterio de parada

- 1 Construir una solución voraz aleatoria.
- 2 Aplicar una técnica de búsqueda local a la solución voraz aleatoria obtenida en el paso anterior (para mejorarla).
- 3 Actualizar la mejor solución encontrada.

Se puede extender este algoritmo si se le añade un operador de mutación (mecanismo de generación de vecinos) al procedimiento GRASP básico:

Mientras no se satisfaga el criterio de parada

- 1 Solución = Solución voraz aleatoria.
- 2 Repetir L veces

- Solución = Búsqueda local (Solución).
- Actualizar la mejor solución (si corresponde).
- Solución = Mutación(Solución).

1.5.1. Algoritmo voraz

Como algoritmo voraz se puede utilizar una versión simplificada del algoritmo de Batchelor y Wilkins. El algoritmo escoge en cada paso al "mejor candidato", para formar parte de la solución y comprueba si al incluir este elemento el conjunto es una solución factible. De ser así, se mantiene esta estructura o de lo contrario se descarta al elemento seleccionado del conjunto de elegibles y se escoge al siguiente "mejor candidato".

1.5.2. Mecanismo de generación de soluciones voraces aleatorias

En la construcción de soluciones voraces aleatorias se utiliza una lista de candidatos RCL en la que se incluyen los mejores aspirantes a formar parte de la solución. De esa lista se escoge un candidato aleatoriamente. Un posible algoritmo voraz aleatorio consiste en utilizar como RCL la lista de los patrones más alejados a los centroides ya escogidos. El tamaño de esta lista puede ser, por ejemplo, igual al 5% de las muestras disponibles. El algoritmo voraz aleatorio es análogo al algoritmo voraz clásico, teniendo en cuenta que, en cada momento, se escoge aleatoriamente uno de los patrones más alejados al conjunto de centroides ya establecido.

1.5.3. Técnica de búsqueda local

La estrategia de búsqueda local se utiliza para mejorar la solución obtenida mediante el mecanismo de generación de soluciones voraces aleatorias. Son de utilidad cuando el camino al objetivo es irrelevante ya que partiendo del estado actual se mueven hacia los vecinos en busca de mejorar la función objetivo.

1.6. Análisis de conglomerados

El análisis de conglomerados es una técnica de clasificación que sirve para poder detectar y describir subgrupos de sujetos o variables homogéneas en función de los valores observados dentro de un conjunto aparentemente heterogéneo. Se fundamenta en el estudio de las distancias entre ellos, permitiendo en el análisis, cuantificar el grado de similitud, en el caso de las proximidades, y el grado de diferencia, en el caso de las distancias. Como resultado aparecen agrupaciones o conglomerados homogéneos.

A diferencia de otras pruebas de clasificación, como por ejemplo el análisis discriminante, las agrupaciones o conglomerados que se establecen se configuran a posteriori. En este caso el investigador no tiene conocimiento de la existencia de los subgrupos o conglomerados, ni del número resultante, ni mucho menos de las características que los definen. Es una técnica, por lo tanto, eminentemente exploratoria y descriptiva sin variables dependientes.

En el proceso de aplicación de esta técnica el investigador tiene que tomar una serie de decisiones: en primer lugar, debe decidir cuantos conglomerados se formarán, en segundo lugar, ha de establecer las medidas de similitud para controlar las medidas de proximidad entre las unidades y, finalmente, debe fijar que procedimiento o método de agrupación se utilizará para configurar los conglomerados.

En cuanto al procedimiento, el análisis de conglomerados permite hacer una agrupación jerarquizada o no. El procedimiento jerárquico es más adecuado para muestras pequeñas en las que los grupos se configuran por agrupaciones sucesivas de individuo a individuo, o de individuo a grupo configurando una estructura arborescente con niveles, que desemboca en una jerarquización de conglomerados. El procedimiento no jerárquico (K-medias), parte de la determinación de un número de grupos y asigna los casos a grupos diferenciados sin que unos dependan de otros (conglomerados no jerárquicos).

El método de la técnica puede ser aglomerativo o divisivo en función del punto de partida. Si se parte de tantos grupos como objetos se tengan y se va obteniendo las agrupaciones pertinentes, es el método aglomerativo. Sin embargo, si se parte de un único grupo y éste se va subdividiendo, se tiene el método divisivo. El método más utilizado es el aglomerativo.

A continuación se centra la explicación en el análisis de conglomerado jerárquico y aglomerativo.

1.6.1. Definición de la técnica: conglomerado jerárquico y aglomerativo

Esta técnica parte de la premisa de que todo fenómeno debe ser ordenado para ser entendible. Sokal y Sneath (1963) son los autores que más han influido en el desarrollo de esta técnica todavía vigente. Se trata de un conjunto de técnicas (fundamentalmente algoritmos) y métodos estadísticos multivariantes de clasificación automática de datos. Concretamente, a partir de una tabla se trata de situar todos los casos en grupos homogéneos (conglomerados o clusters) no conocidos previamente pero que los propios datos sugieren (los individuos parecidos serán asignados a un mismo conglomerado, mientras que los diferentes estarán en conglomerados distintos). Es una técnica que se puede utilizar para agrupar individuos o variables. El procedimiento e interpretación de los resultados siguen el mismo patrón. La única diferencia es el contexto donde se ubica la clasificación.

Es habitual utilizar otras técnicas que complementan o confirman la formación de grupos homogéneos de conglomerados, a saber, el análisis discriminante (técnica ad hoc) permite verificar la existencia de relaciones causales entre la pertenencia a un conglomerado determinado y los valores de las variables (Sneath y Sokal, 1973). Los resultados de un conglomerado pueden ampliarse también con regresión logística,

añadiendo siempre nuevas variables independientes, constituyendo los conglomerados agrupando individuos y no solo variables.

Los principios básicos del análisis de conglomerados recogidos por Pérez (2009) son:

- Es un método estadístico multivariante de clasificación automática de datos.
- Tiene la finalidad de revelar concentraciones en los datos para un agrupamiento eficiente en conglomerados según su homogeneidad.
- El agrupamiento puede ser para casos o variables (cualitativas o cuantitativas).
- Es esencial un uso adecuado del concepto de distancia, ya que los grupos se realizan según la proximidad o lejanía de unos con otros.
- Es fundamental que dentro de un conglomerado los elementos sean homogéneos, y lo más diferentes a los contenidos del resto.
- Es una técnica de clasificación post hoc: el número de conglomerados se determina en función de los datos, y puede no ser definido de antemano.

1.6.2. Condiciones de aplicación del análisis de conglomerados

El análisis de conglomerados suele iniciarse estimando las similitudes entre los individuos mediante la correlación de las distintas variables (cualitativas o cuantitativas). A continuación se procede a comparar los grupos según las similitudes, y finalmente se decide cuántos grupos se construyen. El objetivo será formar el mínimo número de grupos posible, lo más homogéneos dentro de cada grupo, y lo más heterogéneos entre grupos.

De entrada, para aplicar esta técnica no se debe cumplir ningún supuesto paramétrico, sin embargo, sí se deben considerar una serie de axiomas previos:

- Si las variables están en escalas muy diferentes será necesario estandarizar las variables (o trabajar con las desviaciones respecto de la media). También puede hacerse un análisis factorial previo y trabajar con puntuaciones factoriales.
- Observar valores perdidos y atípicos, ya que los valores atípicos deforman las distancias y producen conglomerados unitarios.
- Análisis previo de multicolinealidad, ya que las variables correlacionadas son nocivas para el análisis de conglomerados.
- El número de observaciones en cada conglomerado debe ser relevante, ya que pueden haber valores atípicos que difuminen las agrupaciones.
- El resultado debe tener sentido conceptual.
- Se pueden realizar otros análisis de forma complementaria: discriminante, regresión logística, etc.

1.6.3. Procedimiento

Como se mencionó anteriormente, el objetivo del análisis de conglomerados consiste en conseguir una o más particiones de un conjunto de individuos a partir de determinadas características de los constituyentes. Se podrá decir que los individuos son similares si pertenecen a la misma clase, grupo o conglomerado de tal forma que todos los individuos que están en el mismo conglomerado se parecen entre sí, y son diferentes a los individuos de otro conglomerado. Los miembros de un conglomerado tienen características genéricas comunes, que difícilmente pueden resumirse en una única variable. A partir de la identificación de grupos o conglomerados se podrá trabajar de forma segmentada sin necesidad de seguir trabajando con toda la muestra.

El procedimiento básico es el siguiente:

a. Selección de los individuos objeto de estudio. Se debe prestar especial atención a los casos atípicos que pueden distorsionar la agrupación de grupos homogéneos.

b. Selección de las variables o datos que describen y caracterizan la muestra. Se deben incluir transformaciones a partir de las variables que los definen (tipificación de variables, etc.). La selección de las variables relevantes a los objetivos del estudio es decisiva para que los resultados tengan sentido. La elección de las variables debe ser coherente con la investigación que se lleva a cabo. La inclusión de variables no relevantes aumenta la posibilidad de tener casos atípicos. De entrada, pues, optaremos por un mismo tipo de variables; en caso de no ser así se deberá proceder a su estandarización.

c. Elección de la medida de proximidad entre los individuos. El conocimiento de las distancias ayudará a interpretar las agrupaciones resultantes y a determinar cuál es el punto de corte más adecuado. Las medidas de similitud/distancia definen la proximidad y no covariación, y vienen determinada por la escala de medida de las variables (ordinal o de intervalo-razón). El resultado de la técnica depende de la medida de asociación-similitud-distancia utilizada; así distintas medidas de proximidad pueden ofrecer resultados distintos.

Variables cuantitativas	Variables cualitativas	Datos dicotómicos
Distancia Euclídea	Chi cuadrado	Jaccard
Distancia Euclídea al cuadrado	Phi cuadrado	Russel y Rao
Coseno de vectores		Sokal y Sneath
Correlación de Pearson		Rogers y Tanimoto
Distancia métrica de Chebynev		
Bloque, Manhattan o City-block		
Distancia de Minkowski		

Cuadro 1.1: Clasificación de las principales medidas de distancia, según el tipo de variable.

Fuente: [1]. Elaboración propia.

d. Elección del criterio para agrupar los individuos en conglomerados y la ejecución del algoritmo. No existe un criterio único para seleccionar el mejor algoritmo y, por tanto, la decisión es subjetiva en función del objetivo pretendido. Se puede obtener una agrupación progresiva (jerárquico). Dentro de esta tipología, podemos identificar diferentes métodos (Ver Cuadro (1.2)). De entre todos el más utilizado es el método promedio entre grupos, el resto de métodos requieren trabajar con la distancia Euclídea como criterio.

e. Identificar las agrupaciones resultantes. Mediante una representación gráfica de los conglomerados obtenidos para visualizar los resultados, mediante un dendograma o árbol lógico. Este gráfico resume el proceso de agrupación: en el eje de abscisas se sitúan los sujetos y en el eje de ordenadas aparecen las distancias utilizadas para agrupar conglomerados. Los sujetos o variables similares se conectan mediante enlaces a partir del método aglomerativo. La posición del enlace determina el nivel de similitud entre los objetos.

f. Interpretación de los resultados obtenidos. La decisión e interpretación final sobre el número adecuado de agrupaciones es totalmente subjetiva. En la decisión final tiene que primar un equilibrio entre un número reducido de conglomerados, con lo que obtendremos agrupaciones heterogéneas, y un número excesivo de agrupaciones con la dificultad de interpretación que lleva asociada. Generalmente se asocia la distancia óptima como punto de corte cuando en el nivel de agrupación se producen saltos bruscos.

Tipo de métodos	Descripción del método
Distancia mínima o vecino más próximo	Los grupos se unen considerando la menor de las distancias existentes entre los miembros más cercanos de distintos grupos. Crea grupos más homogéneos. Ayuda a detectar outliers, pero no es útil para resumir datos. Los conglomerados son demasiado grandes. Es el método más sencillo.
Distancia máxima o vecino más lejano	Los grupos se unen considerando la menor de las distancias existentes entre los miembros más lejanos de distintos grupos. Los grupos resultantes son más heterogéneos. Es útil para detectar outliers y los conglomerados son pequeños y compactos.
Media o promedio entre grupos	La distancia entre los grupos se obtiene calculando la distancia promedio entre todos los pares de observaciones independientemente de que estén próximos o alejados. Agrupa los conglomerados con un tamaño óptimo y fusiona conglomerados con varianzas pequeñas. Es uno de los métodos más utilizados.
Vinculación intra grupos	Es una variante de la anterior aunque en este caso se combinan los grupos buscando que la distancia promedio dentro de cada conglomerado sea la menor posible.
Ward o Método de varianza mínima	La distancia entre dos conglomerados se calcula como la suma de cuadrados entre grupos en el ANOVA. Se persigue la minimización de la varianza intragrupal y maximiza la homogeneidad dentro de los grupos. Suele ser muy adecuado aunque los conglomerados que genera suelen ser pequeños y muy compactos. Es especialmente sensible a los outliers.
Centroide	La distancia entre dos conglomerados se calcula como la distancia entre sus centroides. Se trabaja con los valores originales. Las variables deben ser cuantitativas de intervalo. Este método es sensible si los tamaños de los grupos son muy diferentes
Agrupación de medianas	Variante del método anterior en la que no se considera el número de casos que forman cada uno de los agrupamientos, sino solo el número de conglomerados.

Cuadro 1.2: Clasificación de los métodos de agrupamiento.

Fuente: [1]. Elaboración propia.

Capítulo 2

Técnicas metaheurísticas y algoritmos basados en GRASP

La construcción de equipos de trabajo es uno de los principales problemas a la hora de iniciar un proyecto, ya que la conformación de estos equipos determina el éxito o fracaso de un emprendimiento. En esta tesis abordaremos el problema de la conformación de equipos eficientes de trabajo para el desarrollo de los operativos de campo de la Encuesta Nacional de Empleo y Desempleo Urbano, ENEMDU, desarrollada trimestralmente por el INEC en las 5 ciudades bases: Quito, Guayaquil, Cuenca, Ambato y Machala.

Para resolver este problema en esta tesis se utiliza la metaheurística GRASP, la cual consiste en algoritmos que proporcionan soluciones factibles (admisibles), que, aunque no alcancen necesariamente el óptimo, al menos se acercan a su valor en un tiempo de cálculo razonable. A pesar de que en un principio parezca lo contrario, en la mayoría de los casos nos encontramos con procedimientos muy simples. Adicionalmente utilizando variables genéricas las cuales pueden ser reemplazadas en busca de una mejor diversidad basados en los criterios de los especialistas de recursos humanos.

Para el presente análisis se considerará a un equipo de trabajo como un conjunto de personas que actúan de manera interdependiente, aportando habilidades complementarias para el logro de un propósito común. Esta definición está basada en la evidencia que muestra que bajo ciertas hipótesis la diversidad en un grupo incrementa la capacidad del mismo para resolver problemas [8].

2.1. Descripción de la metaheurística GRASP

La metodología GRASP fue desarrollada al final de la década de los ochenta como un procedimiento iterativo que consiste en una fase de construcción, una fase

de mejora y una fase de actualización. En la primera fase se construye iterativamente una solución inicial, añadiendo un elemento en cada paso, es decir en cada iteración del algoritmo. La elección del próximo elemento para ser añadido a la solución parcial viene determinado por una función voraz, denominada así al no considerar sus posibles soluciones en iteraciones siguientes y enfocarse solamente en su valor actual. Esta función mide el beneficio de añadir cada uno de los elementos a la solución construida. En cada iteración de esta fase, la elección del siguiente elemento que formará parte de la solución se determina elaborando una lista de candidatos (LC) con todos los elementos que pueden entrar a formar parte de la solución, descartando como es de suponer, los ya seleccionados en iteraciones anteriores. Los elementos de esta lista de candidatos se ordenan de forma decreciente respecto a los valores obtenidos en la función voraz que mide, como ya hemos comentado, el beneficio asociado a cada uno de los elementos. De esta lista de candidatos ordenada se toman los elementos que conformarán la llamada lista restringida de candidatos (RCL). Esta lista incluirá a aquellos elementos cuyos valores de la función voraz sean más beneficiosos desde el punto de vista del criterio de optimización definido. Una vez constituida esta lista, se seleccionará aleatoriamente un elemento de la misma, que automáticamente pasará a formar parte de la solución de partida. Se repite este procedimiento hasta conformar un grupo del tamaño deseado.

Es necesario resaltar el hecho de que esta medida no tiene en cuenta que ocurrirá en iteraciones sucesivas una vez que se hace una elección, sino únicamente en la iteración actual. Se dice además que el procedimiento es adaptativo porque en cada iteración se actualizan los beneficios obtenidos de añadir el elemento seleccionado a la solución parcial, es decir, la evaluación que se tenga de añadir un determinado elemento a la solución en la iteración j no coincidirá necesariamente con la que se tenga en la iteración $j + 1$.

Por otro lado, la heurística es aleatoria porque no selecciona el mejor candidato según la función voraz elegida, sino que, con el objeto de diversificar y no repetir soluciones en dos construcciones diferentes, se construye una lista con los mejores candidatos de entre los que se selecciona uno al azar.

Dado que la fase inicial no garantiza la optimalidad local respecto a la estructura del entorno en el que se está trabajando debido a las selecciones aleatorias, se aplica un procedimiento de búsqueda local (vecindades) como post-procedimiento para mejorar la solución obtenida.

Hemos de tener en cuenta que, en un problema de optimización combinatoria dado, cada solución del problema tiene un conjunto de soluciones asociadas al que se denomina entorno o vecindad de la solución. Así, dada una solución, podemos calcular cualquier solución perteneciente a su entorno mediante una operación denominada movimiento.

Un pseudocódigo de la fase de construcción se lo puede enunciar de la siguiente manera:

```
                % % % Fase de construcción % % %
Inicio
Solución = {}
    MIENTRAS La solución no esté completa
        Actualizar la función voraz
        Hallar LC
        Ordenar de forma decreciente LC
        Hallar RCL
        a = Selección Aleatoria
        Solución = Solución + {a}
    FIN MIENTRAS
Fin
```

Como se mencionó anteriormente las soluciones generadas en la fase de construcción del algoritmo GRASP no garantizan la optimalidad, de aquí el interés de realizar una búsqueda local posterior a la fase construcción.

En la fase de mejora se realiza un procedimiento de búsqueda local, la cual parte de la solución inicial obtenida en la fase de construcción y calcula una nueva solución considerando los elementos restantes de LC. Es decir, evalúa el cambio en la función objetivo al reemplazar un elemento de la solución construida por cualquier elemento de la lista de candidatos LC restante.

Un pseudocódigo genérico para la fase de mejora es:

```
                % % % Fase de mejora % % %
Inicio
    SolIni=Solución inicial
    MejorSolución=SolIni
        SolMej=Búsqueda local
        Si Valor (SolMej) < Valor (MejorSolución)
            entonces
                MejorSolución=SolMej
Fin
```

La fase de actualización consiste en realizar varias veces la fase de construcción y mejora y elegir de entre ellas a la mejor solución encontrada. Esto nos ofrece un muestreo del espacio de soluciones existente.

2.2. Funcionamiento del algoritmo

En esta sección, para mostrar el funcionamiento del algoritmo propuesto en el caso específico de la conformación de grupos de trabajo y considerando su carácter iterativo, se resolverá una iteración del problema.

Se debe considerar que los resultados obtenidos dependen en gran medida del valor que tomen una serie de parámetros con los que trabaja la metaheurística, para ello se iniciará estudiando los parámetros que tienen una mayor influencia sobre el funcionamiento de la metaheurística y en que valores concretos tenemos que fijar éstos de forma que se obtengan los mejores resultados.

Los parámetros que, por su importancia, han sido seleccionados para la realización del estudio son los siguientes:

Función índice: La función índice es el criterio de optimización elegido para la resolución de este problema. Como el objetivo de este trabajo es cuantificar las diferencias de habilidades entre un conjunto de postulantes a encuestadores utilizaremos la medida de dispersión de la suma como función índice.

Tamaño de la RCL: El tamaño de la RCL tiene un papel decisivo en la fase de construcción de la solución en la metaheurística GRASP, ya que de su valor depende el número de aspirantes a entrar en la solución en cada una de las iteraciones de esta fase. Dado que la selección del candidato es un proceso totalmente aleatorio, la elección de una RCL de gran tamaño, aunque aumenta la aleatoriedad, permite que aspirantes con valores de la función índice bajos puedan llegar a formar parte de la solución, mientras que una RCL de pequeño tamaño, aunque hace que sólo los mejores trabajos sean candidatos a entrar en la solución, esto repercute en que el método pierda aleatoriedad, por tanto, escoger el tamaño adecuado es fundamental para obtener resultados de calidad.

Número de iteraciones: El número de iteraciones, que en nuestro caso funcionará además como criterio de parada, se puede definir como el número de veces que la metaheurística realiza el proceso completo de construcción y mejora de la solución. Según esto, el algoritmo generará tantas soluciones como iteraciones realice y cuanto mayor sea dicho número mayor será la probabilidad de encontrar mejores soluciones. Ahora bien, hay que tener en cuenta que un número excesivo de iteraciones repercute negativamente en el tiempo empleado por el algoritmo en la obtención de la mejor solución.

Un número adecuado de iteraciones será pues aquél que nos permita obtener una solución de calidad en un tiempo razonable.

A continuación para ejemplificar el funcionamiento y bondades del algoritmo resolveremos el siguiente problema académico. Se supone que de una lista de 10 candidatos de la regional Ambato, con las características que se muestran en el Cuadro (2.1), se deber conformar un equipo de trabajo de 3 miembros con la mayor diversidad posible según la medida de la dispersión de la suma definida en el Capítulo 1.

Solo para este caso consideraremos el tamaño del conjunto RCL constante e igual a 4.

Inicialmente no se posee criterio de ingreso al conjunto RCL por parte de los elementos elegibles ya que por definición la distancia de un punto a sí mismo es cero, por lo cual el valor de la función objetivo es nula para todos los postulantes. Por esto, se considera como criterio de ingreso inicial la calificación obtenida en las pruebas técnicas. Se ordena de forma decreciente según su calificación obtenida, obteniendo la siguiente lista de candidatos

$$LC = \{4, 5, 6, 8, 10, 1, 2, 3, 7, 9\}. \quad (2.1)$$

Por lo tanto, se tiene que la lista de candidatos reducida para la elección del primer elemento está dado por

$$RCL = \{4, 5, 6, 8\}. \quad (2.2)$$

Del conjunto RCL, se toma un elemento al azar, supongamos el candidato número 6. Lo que nos da la siguiente lista de candidatos restante

$$LC = \{4, 5, 8, 10, 1, 2, 3, 7, 9\}. \quad (2.3)$$

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

ID	PRUEBAS TÉCNICAS	CALIFICACIÓN INSTRUCCIÓN	INICIATIVA	ORGANIZACIÓN DE INFORMACIÓN	COMPRENSIÓN ORAL	EXPRESIÓN ORAL
1	8	7	9	8	8	9
2	7	7	5	7	6	2
3	7	7	6	2	6	9
4	10	10	3	8	3	7
5	10	8	1	3	1	10
6	10	10	6	8	8	9
7	7	9	5	1	1	9
8	9	10	4	2	3	5
9	7	7	7	9	5	2
10	9	9	5	2	1	1

Cuadro 2.1: Características a ser evaluadas.

Fuente: Generación aleatoria. Elaboración propia.

A continuación, se calcula cual sería el valor de la función objetivo al incorporar cualquiera de los elementos del conjunto LC de (2.3) a la solución construida. Esta función objetivo viene dada por la primera ecuación del sistema (1.6)

Elemento	1	2	3	4	5	7	8	9	10
Función Objetivo	0.04	0.15	0	0.09	0.22	0.06	0.06	0.15	0.1

Cuadro 2.2: Función objetivo para dos elementos.

Fuente: Algoritmo GRASP. Elaboración propia.

Ordenando la lista de candidatos LC, según su aporte a la función objetivo se tiene

$$LC = \{5, 2, 9, 10, 4, 7, 8, 1, 3\}, \quad (2.4)$$

y por lo tanto la lista de candidatos reducida es

$$RCL = \{5, 2, 9, 10\}. \quad (2.5)$$

Se elige un elemento al azar, supongamos que es el candidato número 9.

Se actualiza la lista de candidatos

$$LC = \{5, 2, 10, 4, 7, 8, 1, 3\}, \quad (2.6)$$

y se calcula su aporte a la función objetivo de ser considerados en la solución construida.

Se ordena la lista de candidatos según su aporte de forma descendente

$$LC = \{5, 7, 10, 8, 3, 4, 1, 2\}, \quad (2.7)$$

se obtiene la lista de candidatos reducida

$$LC = \{5, 7, 10, 8\}, \quad (2.8)$$

Elemento	1	2	3	4	5	7	8	10
Función Objetivo	0.21	0.18	0.4	0.27	0.65	0.6	0.40	0.51

Cuadro 2.3: Función objetivo para tres elementos.

Fuente: Algoritmo GRASP. Elaboración propia.

y se toma al azar el último miembro del grupo, obteniendo la siguiente solución construida

$$SC = \{6, 9, 8\}, \quad (2.9)$$

cuya función objetivo es $FO = 0,40$.

Actualizamos la lista de candidatos restantes

$$LC = \{1, 2, 3, 4, 5, 7, 10\}, \quad (2.10)$$

A continuación se realiza una búsqueda en las vecindades con el objetivo de mejorar el valor de la función objetivo, con este fin se evalúa el resultado de reemplazar cualquier elemento de SC por algún otro de LC . De este análisis se concluye que la función objetivo mejora, tomando un valor de $FO = 0,74$, al reemplazar el primer elemento por el candidato número 5. Es decir que la solución construida se torna

$$SC = \{5, 9, 8\}, \quad (2.11)$$

Actualizamos la lista de candidatos

$$LC = \{1, 2, 3, 4, 6, 7, 10\}, \quad (2.12)$$

Realizando una vez más este análisis por vecindades se obtiene que la función objetivo toma un valor de $FO = 0,92$ al reemplazar el tercer elemento por el candidato número 10, consiguiendo finalmente la siguiente solución construida

$$SC = \{5, 9, 10\}. \quad (2.13)$$

A continuación, la metaheurística comenzaría de nuevo todo el proceso, realizándolo tantas veces como iteraciones se hayan fijado de antemano. En cada una de estas iteraciones tomaría la mejor solución encontrada, la compararía con la mejor solución guardada hasta ese momento y ofrecería al final la mejor de todas.

2.3. Desarrollo experimental

En esta sección se conformarán los equipos de trabajo, en las regionales de Quito, Guayaquil, Cuenca, Ambato y Machala del INEC, que se encargarán del levantamiento de información necesaria para el desarrollo de la encuesta ENEMDU, considerando las siguientes variables:

- Calificación en las pruebas técnicas
- Capacidad para resolver problemas lógicos

- Iniciativa
- Capacidad para seguir órdenes
- Organización de información
- Orientación al liderazgo
- Comprensión oral
- Comprensión escrita
- Expresión oral
- Expresión escrita

Estas variables están codificadas numéricamente, donde un valor pequeño representa una baja capacidad y un valor alto representa una mejor capacidad.

A continuación se presenta una breve descripción de las variables para cada una de las regionales.

Ambato: Para el presente análisis se cuenta con información sobre 50 candidatos de la regional Ambato cuyas características se detallan en los siguientes Cuadros.

	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información
Promedio	7.88	5.46	5.40	5.70	5.56
Max	10	8	8	8	8
Min	6	3	3	3	3
Moda	8	4	4	8	3
Desviación Estándar	1.33	1.85	1.70	1.84	1.76

Cuadro 2.4: Características de las variables de la regional Ambato.

Fuente: Levantamiento del autor. Elaboración propia.

	Liderazgo	Comprensión oral	Comprensión escrita	Expresión oral	Expresión escrita
Promedio	5.48	5.58	5.60	5.58	5.54
Max	8	8	8	8	8
Min	3	3	3	3	3
Moda	7	8	4	7	6
Desviación Estándar	1.61	1.91	1.81	1.76	1.80

Cuadro 2.5: Características de las variables de la regional Ambato.

Fuente: Levantamiento del autor. Elaboración propia.

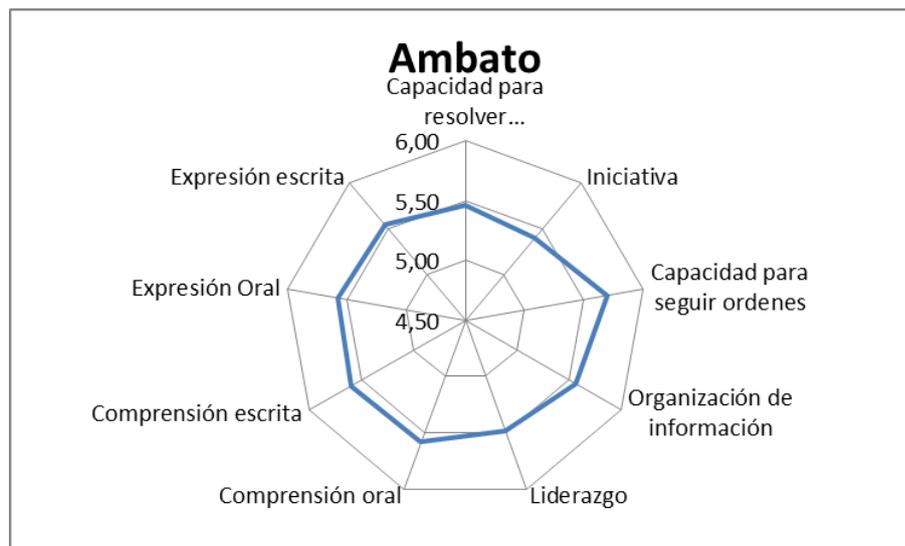


Figura 2.1: Características promedio de los candidatos de la regional Ambato.

Fuente: Levantamiento del autor. Elaboración propia.

Cuenca: Para el presente análisis se cuenta con información sobre 54 candidatos de la regional Cuenca cuyas características se detallan en los siguientes Cuadros.

	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información
Promedio	7,61	5,87	5,91	5,17	5,48
Max	10	8	8	8	8
Min	6	3	3	3	3
Moda	8	8	7	3	5
Desviación Estándar	1,20	1,71	1,64	1,80	1,58

Cuadro 2.6: Características de las variables de la regional Cuenca.

Fuente: Levantamiento del autor. Elaboración propia.

	Liderazgo	Comprensión oral	Comprensión escrita	Expresión oral	Expresión escrita
Promedio	5.26	5.94	5.61	5.50	5.13
Max	8	8	8	8	8
Min	3	3	3	3	3
Moda	3	7	7	4	3
Desviación Estándar	1.83	1.72	1.38	1.75	1.69

Cuadro 2.7: Características de las variables de la regional Cuenca.

Fuente: Levantamiento del autor. Elaboración propia.

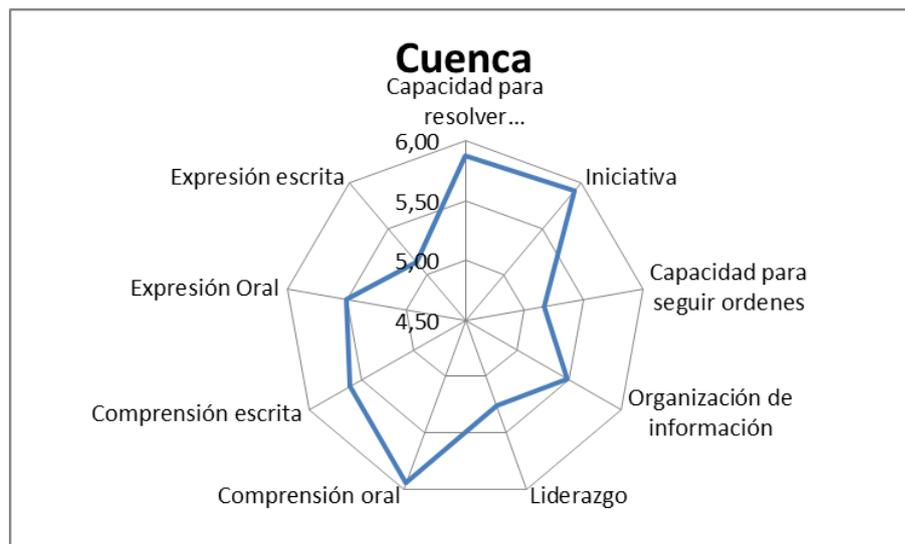


Figura 2.2: Características promedio de los candidatos de la regional Cuenca.

Fuente: Levantamiento del autor. Elaboración propia.

Guayaquil: Para el presente análisis se cuenta con información sobre 68 candidatos de la regional Guayaquil cuyas características se detallan en los siguientes Cuadros.

	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información
Promedio	7,61	6,54	6,07	5,18	5,35
Max	10	8	8	8	8
Min	6	4	4	2	2
Moda	8	7	8	8	8
Desviación Estándar	1,20	1,20	1,47	2,34	2,35

Cuadro 2.8: Características de las variables de la regional Guayaquil.

Fuente: Levantamiento del autor. Elaboración propia.

	Liderazgo	Comprensión oral	Comprensión escrita	Expresión oral	Expresión escrita
Promedio	5,00	5,97	6,15	5,84	6,22
Max	8	8	8	8	8
Min	2	3	3	3	3
Moda	7	8	8	8	7
Desviación Estándar	2,21	1,80	1,60	1,70	1,61

Cuadro 2.9: Características de las variables de la regional Guayaquil.

Fuente: Levantamiento del autor. Elaboración propia.

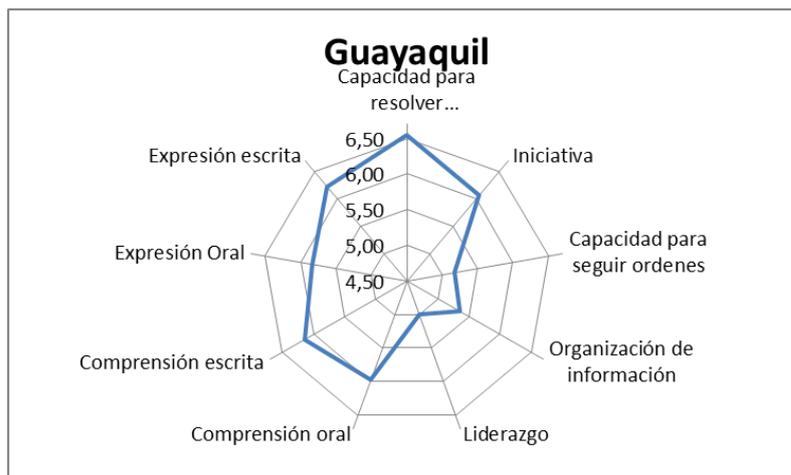


Figura 2.3: Características promedio de los candidatos de la regional Guayaquil.

Fuente: Levantamiento del autor. Elaboración propia.

Machala: Para el presente análisis se cuenta con información sobre 48 candidatos de la regional Machala cuyas características se detallan en los siguientes Cuadros.

	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información
Promedio	7.06	6.04	5.92	6.00	6.02
Max	10	8	8	8	8
Min	6	4	4	4	4
Moda	6	5	4	6	5
Desviación Estándar	1.16	1.38	1.44	1.38	1.44

Cuadro 2.10: Características de las variables de la regional Machala.

Fuente: Levantamiento del autor. Elaboración propia.

	Liderazgo	Comprensión oral	Comprensión escrita	Expresión oral	Expresión escrita
Promedio	5.71	5.79	5.75	6.00	6.15
Max	8	8	8	8	8
Min	4	4	4	4	4
Moda	4	4	4	6	6
Desviación Estándar	1.46	1.50	1.47	1.38	1.32

Cuadro 2.11: Características de las variables de la regional Machala.

Fuente: Levantamiento del autor. Elaboración propia.

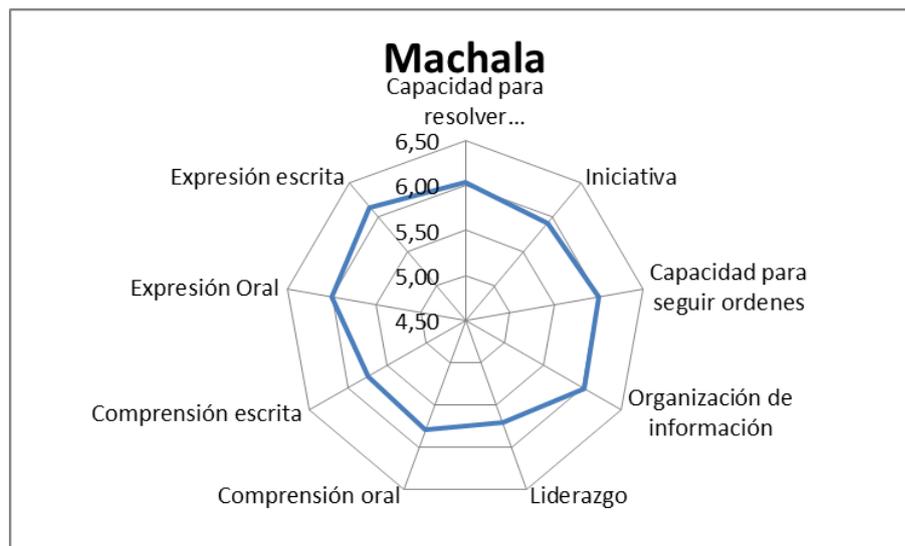


Figura 2.4: Características promedio de los candidatos de la regional Machala.

Fuente: Levantamiento del autor. Elaboración propia.

Quito: Para el presente análisis se cuenta con información sobre 62 candidatos de la regional Quito cuyas características se detallan en los siguientes Cuadros.

	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información
Promedio	7.90	6.05	5.58	5.48	5.95
Max	10	8	8	8	8
Min	6	4	4	4	4
Moda	7	7	4	4	7
Desviación Estándar 1.10	1.40	1.31	1.39	1.30	

Cuadro 2.12: Características de las variables de la regional Quito.

Fuente: Levantamiento del autor. Elaboración propia.

	Liderazgo	Comprensión oral	Comprensión escrita	Expresión oral	Expresión escrita
Promedio	5.61	6.15	5.90	6.11	5.65
Max	8	8	8	8	8
Min	4	4	4	4	4
Moda	4	7	6	7	4
Desviación Estándar	1.44	1.40	1.34	1.12	1.42

Cuadro 2.13: Características de las variables de la regional Quito.

Fuente: Levantamiento del autor. Elaboración propia.

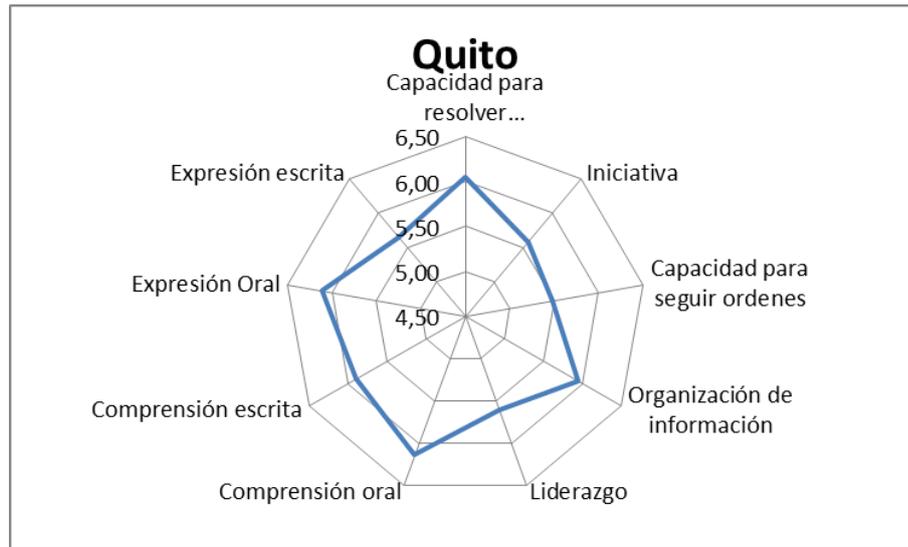


Figura 2.5: Características promedio de los candidatos de la regional Quito.
Fuente: Levantamiento del autor. Elaboración propia.

2.3.1. Conformación de los equipos de trabajo

Tras el breve análisis de las características de los candidatos de las diferentes regionales, a continuación se mostrarán los resultados obtenidos al conformar 5 equipos de trabajo de 5 miembros cada uno, para las distintas regionales del INEC, utilizando la metaheurística GRASP. Para este ejercicio se procederá de la siguiente manera: se conformará el primer equipo de trabajo y se eliminan los elementos ya seleccionados de la lista de candidatos para la conformación del segundo equipo. Se continúa con este procedimiento hasta conformar los 5 equipos necesarios. El algoritmo se corrió considerando los siguientes parámetros: Sean FO^* y FO_* el valor máximo y mínimo de la función voraz, el conjunto RCL está formado por todos los elementos x de conjunto LC de manera que su función objetivo FO cumpla la condición $FO(x) \geq FO^* + \alpha(FO^* - FO_*)$. Nótese que si $\alpha = 0$, entonces este esquema de selección es un algoritmo miope, mientras que si $\alpha = 1$, entonces es totalmente aleatorio. Para nuestra experimentación consideraremos un $\alpha = 1$, garantizando la mayor aleatoriedad posible y un número de 20 iteraciones de entre las cuales se elige a la mejor respuesta.

A continuación se muestran los equipos formados al utilizar la metaheurística GRASP sobre la totalidad de la base de candidatos de las diferentes regionales del INEC.

Conformación del primer equipo De un total de 49 candidatos presentados en la regional Ambato, se conformó el primer equipo de trabajo. En el Cuadro (2.14) se describen las características de los candidatos elegidos.

Al recordar que el objetivo de esta tesis es conformar grupos de trabajo que se complementen, consideraremos como la capacidad del equipo, la capacidad máxima individual de los miembros del grupo en cada una de las variables analizadas. En la siguiente Figura se muestra la capacidad del equipo uno de la regional Ambato.

ID	Cap. para resolver prob. lógicos	Iniciativa	Capacidad para seguir órdenes	Org. de inf.	Liderazgo	C. oral	C. escrita	E. oral	E. escrita
2	4	8	8	3	3	5	8	3	7
13	8	4	7	6	7	3	3	7	3
20	3	7	3	3	7	8	4	3	8
25	3	5	8	8	4	5	4	3	3
45	3	3	3	3	3	7	7	6	4
MAX	8	8	8	8	7	8	8	7	8

Cuadro 2.14: Primer grupo regional Ambato.

Fuente: Algoritmo GRASP. Elaboración propia.

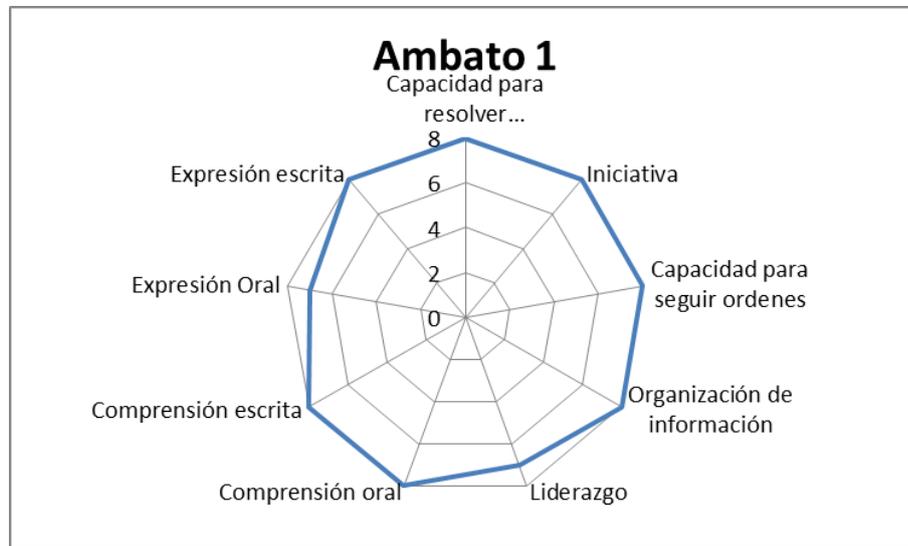


Figura 2.6: Capacidad del equipo uno de la regional Ambato.

Fuente: Algoritmo GRASP. Elaboración propia.

Conformación del segundo equipo Para la conformación del segundo equipo de trabajo consideraremos a la regional Cuenca. Para este caso de un total de 54 candidatos, se obtuvo que el primer equipo de trabajo está conformado por los candidatos $Cuenca_1 = \{9, 20, 24, 33, 51\}$. Como el objetivo en esta sección es describir el segundo equipo conformado, se tiene que de 49 candidatos restantes, tras eliminar del conjunto elegible a los 5 elementos que conforman el equipo $Cuenca_1$, se obtiene que el segundo equipo está conformado por los elementos $Cuenca_2 = \{13, 19, 26, 38, 53\}$. En el Cuadro (2.15) se describen las características de los candidatos elegidos en el segundo equipo.

En la siguiente Figura se muestra las capacidad del equipo dos de la regional Cuenca.

ID	Cap. para resolver prob. lógicos	Iniciativa	Capacidad para seguir órdenes	Org. de inf.	Liderazgo	C. oral	C. escrita	E. oral	E. escrita
13	3	3	8	6	5	6	7	4	3
19	6	7	8	7	7	4	3	8	3
26	8	6	3	3	4	5	4	3	7
38	8	5	3	4	3	8	3	6	7
53	3	8	3	5	8	7	7	4	6
MAX	8	8	8	7	8	8	7	8	7

Cuadro 2.15: Segundo equipo regional Cuenca.

Fuente: Algoritmo GRASP. Elaboración propia.

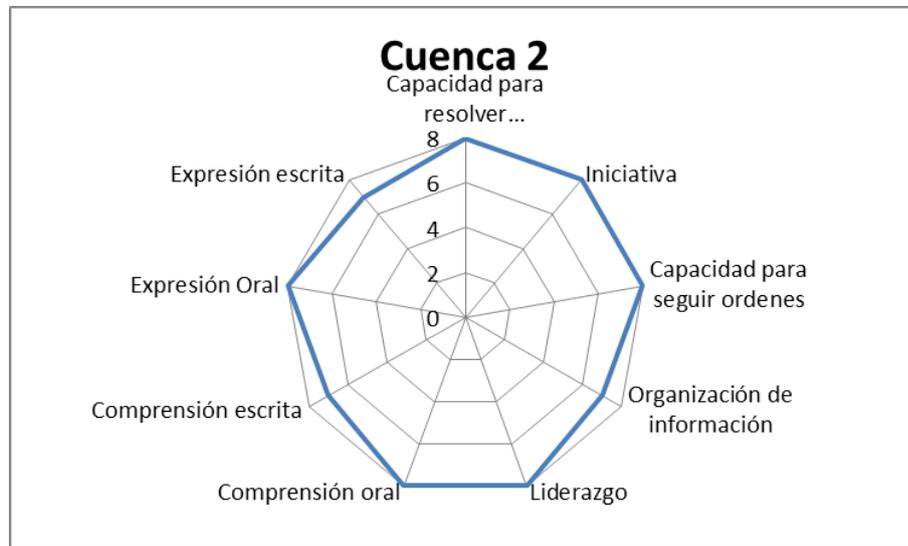


Figura 2.7: Capacidad del equipo dos de la regional Cuenca.

Fuente: Algoritmo GRASP. Elaboración propia.

Conformación del tercer equipo Para la conformación del tercer equipo de trabajo consideraremos a la regional Guayaquil. Para este caso se consideró un total de 68 candidatos. Los dos primeros equipos de trabajo conformados fueron $Guayaquil_1 = \{1, 2, 22, 34, 47\}$, $Guayaquil_2 = \{9, 36, 40, 50, 61\}$. Como el objetivo en esta sección es describir el tercer grupo conformado, se tiene que de 58 candidatos restantes, tras eliminar del conjunto elegible a los elementos ya seleccionados, se obtiene que el tercer equipo está conformado por los elementos $Guayaquil_3 = \{13, 30, 31, 32, 58\}$. En el Cuadro (2.16) se describen las características de los candidatos elegidos en el tercer equipo.

En la siguiente Figura se muestra las capacidades del equipo tres de la regional Guayaquil.

ID	Cap. para resolver prob. lógicos	Iniciativa	Capacidad para seguir órdenes	Org. de inf.	Liderazgo	C. oral	C. escrita	E. oral	E. escrita
13	8	4	8	2	2	6	6	4	6
30	5	8	7	8	4	6	8	4	3
31	5	4	8	7	7	3	7	7	5
32	8	6	2	4	3	4	4	7	8
58	6	7	2	2	7	7	5	3	5
MAX	8	8	8	8	7	7	8	7	8

Cuadro 2.16: Tercer equipo regional Guayaquil.

Fuente: Algoritmo GRASP. Elaboración propia.

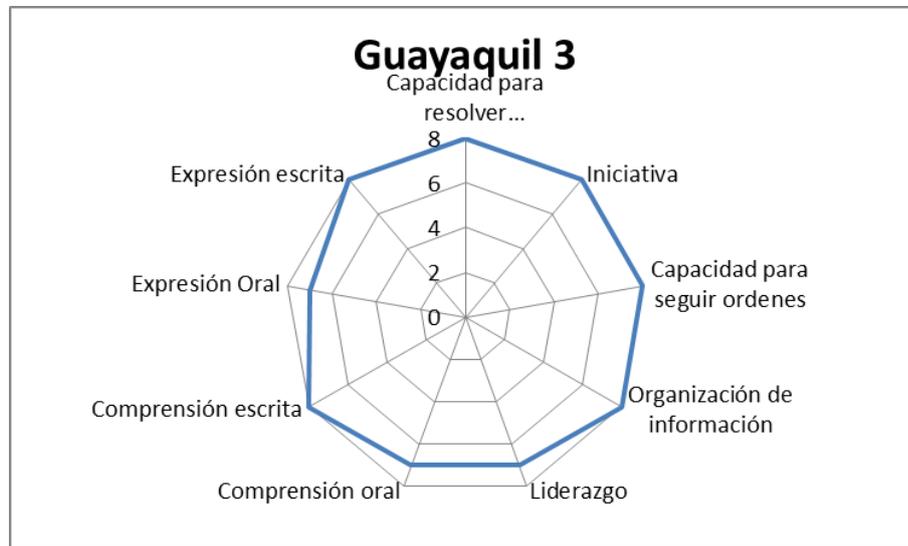


Figura 2.8: Capacidad del equipo tres de la regional Guayaquil.

Fuente: Algoritmo GRASP. Elaboración propia.

Conformación del cuarto equipo Para la conformación del cuarto equipo de trabajo consideraremos a la regional Machala. Para este caso se consideró un total de 48 candidatos. Los tres primeros equipos de trabajo conformados fueron $Machala_1 = \{1, 5, 25, 38, 41\}$, $Machala_2 = \{8, 10, 27, 40, 48\}$, $Machala_3 = \{11, 29, 32, 42, 43\}$. Como el objetivo en esta sección es describir el cuarto grupo conformado, se tiene que de 33 candidatos restantes, tras eliminar del conjunto elegible a los elementos ya seleccionados, se obtiene que el cuarto equipo está conformado por los elementos $Machala_4 = \{3, 15, 26, 35, 44\}$. En el Cuadro (2.17) se describen las características de los candidatos elegidos en el cuarto equipo.

En la siguiente Figura se muestra las capacidad del equipo cuatro de la regional Machala.

ID	Cap. para resolver prob. lógicos	Iniciativa	Capacidad para seguir órdenes	Org. de inf.	Liderazgo	C. oral	C. escrita	E. oral	E. escrita
3	7	8	7	5	8	8	4	7	5
15	8	7	4	5	5	6	8	5	8
26	4	5	8	8	4	7	6	7	6
35	7	8	6	4	8	5	6	8	7
44	5	4	5	8	4	4	5	4	6
MAX	8	8	8	8	8	8	8	8	8

Cuadro 2.17: Cuarto equipo regional Machala.

Fuente: Algoritmo GRASP. Elaboración propia.

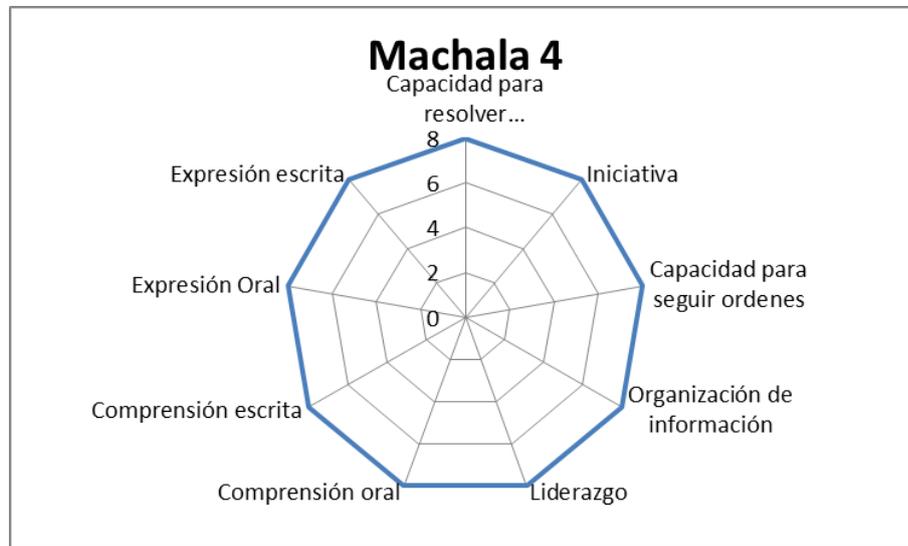


Figura 2.9: Capacidad del equipo cuatro de la regional Machala.

Fuente: Algoritmo GRASP. Elaboración propia.

Conformación del quinto equipo Para la conformación del quinto equipo de trabajo consideraremos a la regional Quito. Para este caso se consideró un total de 62 candidatos. Los cuatro primeros equipos de trabajo conformados fueron $Quito_1 = \{6, 9, 13, 18, 21\}$, $Quito_2 = \{16, 36, 48, 49, 54\}$, $Quito_3 = \{4, 26, 31, 46, 50\}$, $Quito_4 = \{10, 12, 25, 27, 32\}$. Como el objetivo en esta sección es describir el quinto grupo conformado, se tiene que de 42 candidatos restantes, tras eliminar del conjunto elegible a los elementos ya seleccionados, se obtiene que el cuarto equipo está conformado por los elementos $Quito_5 = \{2, 5, 28, 43, 55\}$. En el Cuadro (2.18) se describen las características de los candidatos elegidos en el cuarto equipo.

En la siguiente Figura se muestra las capacidad del equipo cinco de la regional Quito.

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

ID	Cap. para resolver prob. lógicos	Iniciativa	Capacidad para seguir órdenes	Org. de inf.	Liderazgo	C. oral	C. escrita	E. oral	E. escrita
2	7	4	4	6	8	8	7	6	8
5	5	4	4	7	4	7	7	8	6
28	4	8	4	6	5	6	6	4	4
43	7	5	8	6	7	4	4	7	6
55	8	6	7	5	4	8	5	5	4
MAX	8	8	8	7	8	8	7	8	8

Cuadro 2.18: Quinto equipo regional Quito.

Fuente: Algoritmo GRASP. Elaboración propia.

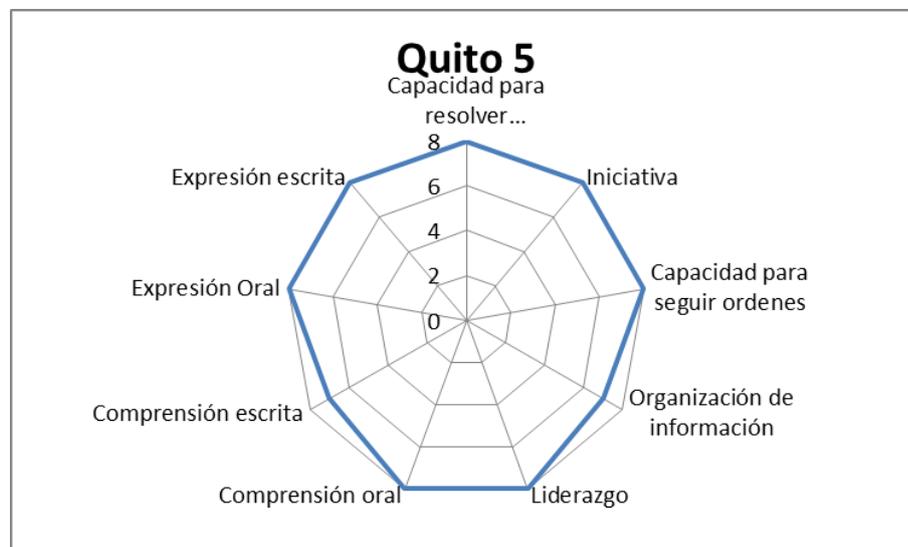


Figura 2.10: Capacidad del equipo quinto de la regional Quito.

Fuente: Algoritmo GRASP. Elaboración propia.

La totalidad de los equipos formados para cada una de las regionales y las bases utilizadas para su conformación son presentadas en el Anexo.

Capítulo 3

Uso del análisis de conglomerados en la conformación de equipos de trabajo

Como se estudió en el Capítulo 1, el propósito del análisis de conglomerados es el agrupar los sujetos o variables homogéneas en función de los valores observados dentro de un conjunto aparentemente heterogéneo de forma que los datos agrupados dentro de un mismo conglomerado sean homogéneos entre sí (mínima varianza) y que estos conglomerados sean lo más heterogéneos posibles entre ellos (máxima varianza). Con esta consideración es casi inmediato el suponer bajo el contexto del problema en análisis, que un método alternativo para la conformación de equipos de trabajo de n miembros, consiste en la conformación de n conglomerados, y tomando aleatoriamente un representante de cada uno para conformar el equipo deseado, garantizando de esta manera la mejor elección de un grupo de trabajo de manera aleatoria, ya que la teoría estadística garantiza la máxima varianza (diversidad) entre elementos de distintos conglomerados. Es por ello que en este capítulo se realizará una comparación entre el algoritmo GRASP y el análisis de conglomerados para la conformación de equipos de trabajos con la mayor diversidad posible.

Para este análisis se utilizará una vez más las bases de los candidatos a encuestadores de las 5 regionales del INEC. Además se utilizarán los siguientes parámetros de comparación. Se construirán 5000 equipos de trabajo con cada uno de los métodos mencionados y se analizará la diversidad medida a través de la distancia del coseno definida en el Capítulo 1 de los diferentes equipos formados. El tamaño del conjunto RCL utilizado en la metodología GRASP formará con la máxima diversidad posible, es decir, con un $\alpha = 1$ y se realizarán 10 iteraciones de las cuales se elegirá a la mejor respuesta.

A continuación se muestran los resultados obtenidos para las diferentes regiona-

les.

3.1. Ambato

Para el proceso de selección de encuestadores se presentaron 50 candidatos, cuyas características promedio fueron mostradas en los Cuadros (2.4) y (2.5). Tras conformar 5000 grupos diferentes con el método de conglomerados propuesto y compararlos con los resultados obtenidos al correr el algoritmo GRASP el mismo número de veces se obtuvieron los siguientes resultados.

DIVERSIDAD		
	Conglomerados	GRASP
Mínimo	0,66	1,56
Máximo	1,47	1,57
Promedio	1,01	1,57
Moda	N/A	1,57
Rango	0,81	0,02
Desviación Estándar	0,12	0,00

Cuadro 3.1: Diversidad Ambato.

Fuente: Método de conglomerados y algoritmo GRASP. Elaboración propia.

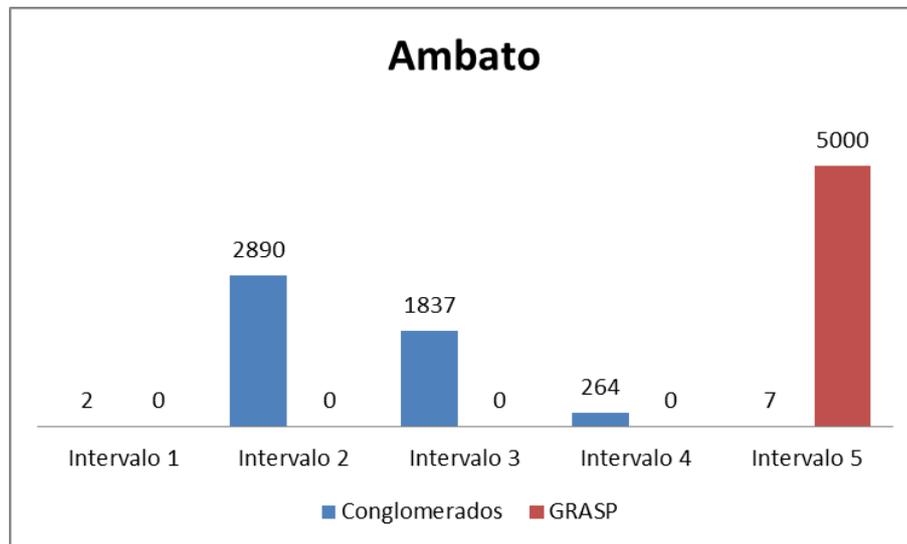


Figura 3.1: Resultados de las comparaciones de la regional Ambato.

Fuente: Método de conglomerados y algoritmo GRASP. Elaboración propia.

Donde los intervalos mostrados en la gráfica anterior vienen dados de la siguiente manera:

De los resultados obtenidos se tiene que el 99,99 % de los resultados obtenidos con el GRASP alcanzan una diversidad de 1,57, mientras que más del 50 % de grupos formados utilizando análisis de conglomerados se ubican en el Intervalo 2.

Intervalo 1	[0, 66; 0, 84]
Intervalo 2	(0, 84; 1, 02]
Intervalo 3	(1, 02; 1, 21]
Intervalo 4	(1, 21; 1, 39]
Intervalo 5	(1, 39; 1, 57]

Cuadro 3.2: Intervalos para la diversidad obtenida.

Fuente: Método de conglomerados y algoritmo GRASP. Elaboración propia.

3.2. Cuenca

Para el proceso de selección de encuestadores en la regional Cuenca se presentaron 54 candidatos, cuyas características promedio fueron mostradas en los Cuadros (2.6) y (2.7). Tras conformar 5000 grupos diferentes con el método de conglomerados propuesto y compararlos con los resultados obtenidos al correr el algoritmo GRASP el mismo número de veces se obtuvieron los siguientes resultados.

DIVERSIDAD		
	Conglomerados	GRASP
Mínimo	0,53	1,36
Máximo	1,25	1,40
Promedio	0,87	1,40
Moda	N/A	1,40
Rango	0,72	0,05
Desviación Estándar	0,12	0,00

Cuadro 3.3: Diversidad Cuenca.

Fuente: Método de conglomerados y algoritmo GRASP. Elaboración propia.

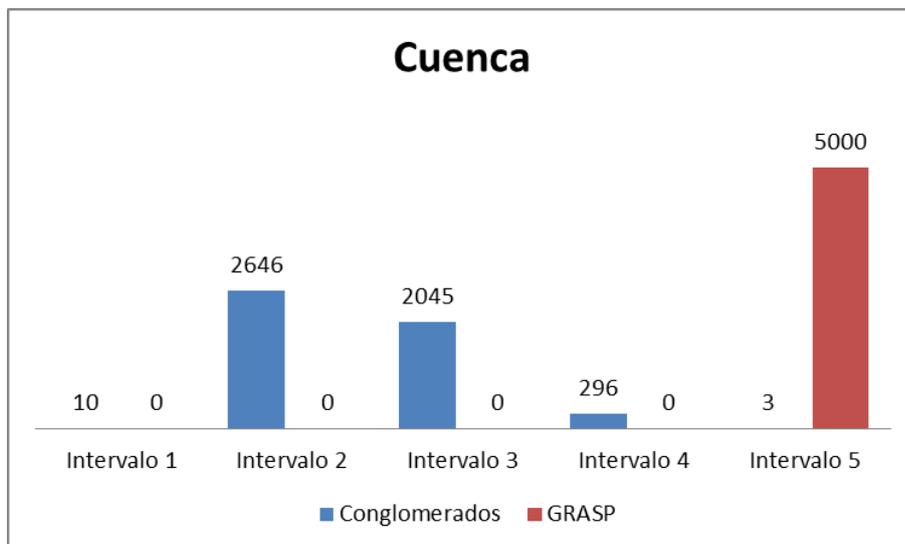


Figura 3.2: Resultados de las comparaciones de la regional Cuenca.

Fuente: Método de conglomerados y algoritmo GRASP. Elaboración propia.

Donde los intervalos mostrados en la gráfica anterior vienen dados de la siguiente manera:

Intervalo 1	[0, 53; 0, 71]
Intervalo 2	(0, 71; 0, 88]
Intervalo 3	(0, 88; 1, 06]
Intervalo 4	(1, 06; 1, 23]
Intervalo 5	(1, 23; 1, 40]

Cuadro 3.4: Intervalos para la diversidad obtenida.

Fuente: Método de conglomerados y algoritmo GRASP. Elaboración propia.

De los resultados obtenidos se tiene que el 99,99 % de los resultados obtenidos con el GRASP alcanzan una diversidad de 1,40, mientras que más del 50 % de grupos formados utilizando análisis de conglomerados se ubican en el Intervalo 2.

3.3. Guayaquil

Para el proceso de selección de encuestadores en la regional Guayaquil se presentaron 68 candidatos, cuyas características promedio fueron mostradas en los Cuadros (2.8) y (2.9). Tras conformar 5000 grupos diferentes con el método de conglomerados propuesto y compararlos con los resultados obtenidos al correr el algoritmo GRASP el mismo número de veces se obtuvieron los siguientes resultados.

DIVERSIDAD		
	Conglomerados	GRASP
Mínimo	0,52	1,58
Máximo	1,40	1,60
Promedio	0,96	1,60
Moda	N/A	1,60
Rango	0,88	0,02
Desviación Estándar	0,13	0,00

Cuadro 3.5: Diversidad Guayaquil.

Fuente: Método de conglomerados y algoritmo GRASP. Elaboración propia.

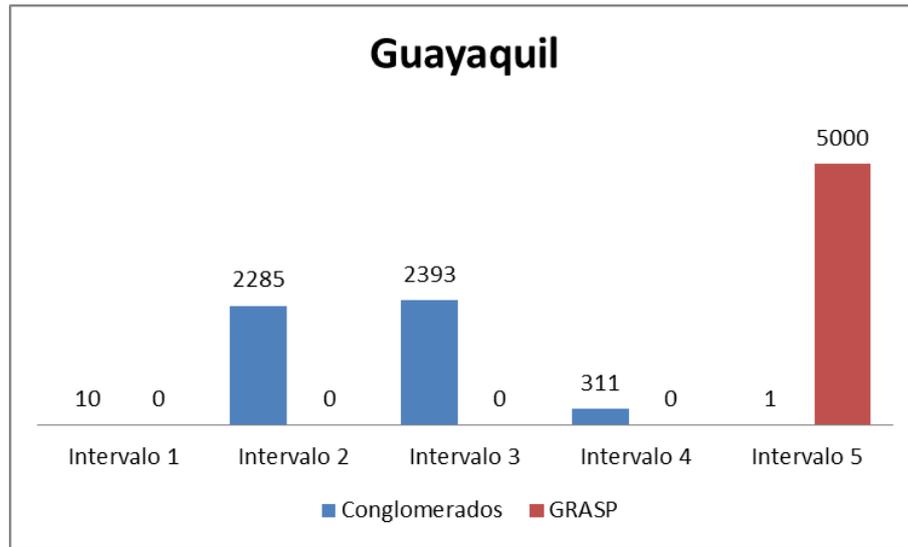


Figura 3.3: Resultados de las comparaciones de la regional Guayaquil.

Fuente: Método de conglomerados y algoritmo GRASP. Elaboración propia.

Donde los intervalos mostrados en la gráfica anterior vienen dados de la siguiente manera:

Intervalo 1	[0, 52; 0, 73]
Intervalo 2	(0, 73; 0, 95]
Intervalo 3	(0, 95; 1, 17]
Intervalo 4	(1, 17; 1, 38]
Intervalo 5	(1, 38; 1, 60]

Cuadro 3.6: Intervalos para la diversidad obtenida.

Fuente: Método de conglomerados y algoritmo GRASP. Elaboración propia.

De los resultados obtenidos se tiene que el 96,72 % de los resultados obtenidos con el GRASP alcanzan una diversidad de 1, 60, mientras que más del 93,56 % de grupos formados utilizando análisis de conglomerados se ubican en los Intervalo 2 y 3.

3.4. Machala

Para el proceso de selección de encuestadores en la regional Machala se presentaron 48 candidatos, cuyas características promedio fueron mostradas en los Cuadros (2.10) y (2.11). Tras conformar 5000 grupos diferentes con el método de conglomerados propuesto y compararlos con los resultados obtenidos al correr el algoritmo GRASP el mismo número de veces se obtuvieron los siguientes resultados.

DIVERSIDAD		
	Conglomerados	GRASP
Mínimo	0,27	0,88
Máximo	0,78	0,90
Promedio	0,50	0,90
Moda	N/A	0,90
Rango	0,51	0,02
Desviación Estándar	0,08	0,00

Cuadro 3.7: Diversidad Machala.

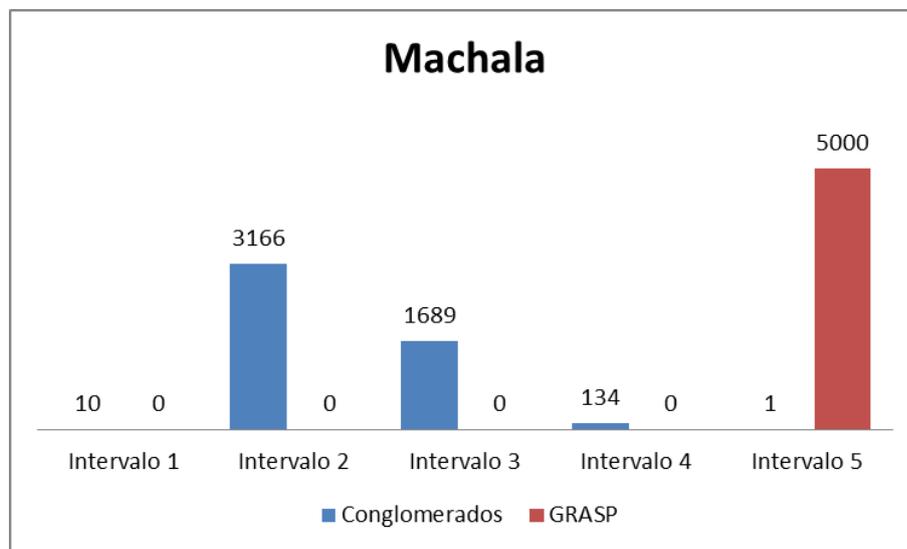


Figura 3.4: Resultados de las comparaciones de la regional Machala.
Fuente: Método de conglomerados y algoritmo GRASP. Elaboración propia.

Donde los intervalos mostrados en la gráfica anterior vienen dados de la siguiente manera:

Intervalo 1	[0, 27; 0, 40]
Intervalo 2	(0, 40; 0, 52]
Intervalo 3	(0, 52; 0, 65]
Intervalo 4	(0, 65; 0, 77]
Intervalo 5	(0, 77; 0, 90]

Cuadro 3.8: Intervalos para la diversidad obtenida.
Fuente: Método de conglomerados y algoritmo GRASP. Elaboración propia.

De los resultados obtenidos se tiene que el 99,64% de los resultados obtenidos con el GRASP alcanzan una diversidad de 0,90, mientras que más del 97,1% de grupos formados utilizando análisis de conglomerados se ubican en los Intervalo 2 y 3.

3.5. Quito

Para el proceso de selección de encuestadores en la regional Quito se presentaron 62 candidatos, cuyas características promedio fueron mostradas en los Cuadros (2.12) y (2.13). Tras conformar 5000 grupos diferentes con el método de conglomerados propuesto y compararlos con los resultados obtenidos al correr el algoritmo GRASP el mismo número de veces se obtuvieron los siguientes resultados.

DIVERSIDAD		
	Conglomerados	GRASP
Mínimo	0,28	0,77
Máximo	0,73	0,79
Promedio	0,51	0,79
Moda	N/A	0,79
Rango	0,45	0,02
Desviación Estándar	0,07	0,00

Cuadro 3.9: Diversidad Quito.

Fuente: Método de conglomerados y algoritmo GRASP. Elaboración propia.

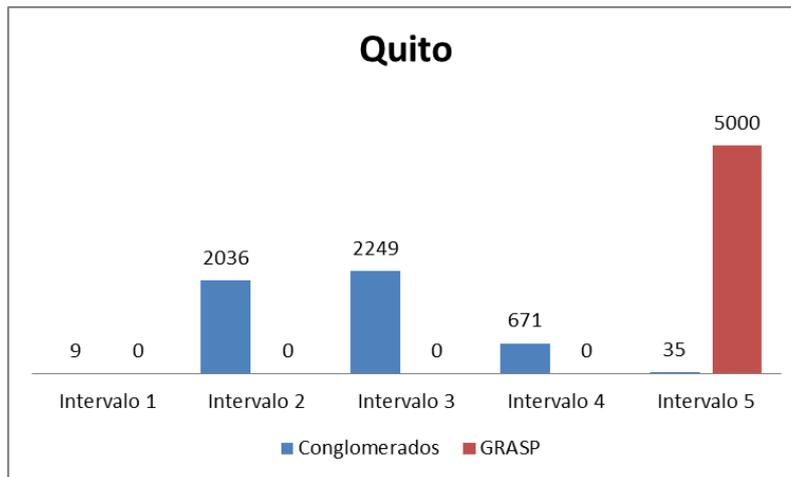


Figura 3.5: Resultados de las comparaciones de la regional Quito.

Fuente: Método de conglomerados y algoritmo GRASP. Elaboración propia.

Donde los intervalos mostrados en la gráfica anterior vienen dados de la siguiente manera:

Intervalo 1	[0, 28; 0, 39]
Intervalo 2	(0, 39; 0, 49]
Intervalo 3	(0, 49; 0, 59]
Intervalo 4	(0, 59; 0, 69]
Intervalo 5	(0, 69; 0, 79]

Cuadro 3.10: Intervalos para la diversidad obtenida.

Fuente: Método de conglomerados y algoritmo GRASP. Elaboración propia.

De los resultados obtenidos se tiene que el 99,29% de los resultados obtenidos con el GRASP alcanzan una diversidad de 0,79, mientras que más del 85,7% de grupos formados utilizando análisis de conglomerados se ubican en los Intervalo 2 y 3.

Capítulo 4

Conclusiones

El principal objetivo de esta tesis es abordar el problema de la conformación de equipos de trabajo para el levantamiento de campo de la Encuesta Nacional de Empleo y Desempleo Urbano, ENEMDU, realizada trimestralmente por el Instituto Nacional de Estadísticas y Censos, INEC. Con este fin se desarrolló un estudio que parte de las bases del análisis funcional con la noción de distancia o disimilitud interelemento de un conjunto arbitrario. A continuación se abordó la noción de medida, y basados en la disimilitud interelemento se definieron alternativas de medidas de diversidad, todo esto con el afán de definir de manera correcta el problema de optimización denominado problema de la diversidad máxima.

Basados principalmente en el hecho de que para la mayoría de problemas de optimización no existe un algoritmo exacto con complejidad polinómica que encuentre la solución óptima a dicho problema y que la cardinalidad del espacio de búsqueda de estos problemas suele ser muy grande, lo cual hace inviable el uso de algoritmos exactos ya que la cantidad de tiempo que necesitaría para encontrar una solución es inaceptable, se optó en esta tesis por el uso de un algoritmo metaheurístico.

De entre todas las metaheurísticas existentes en esta tesis se utilizó la metaheurística GRASP, denominada así por su acrónimo en inglés Greedy Randomized Adaptive Search Procedure (GRASP), que en castellano se podría traducir como procedimientos de búsqueda voraz, aleatorizados y adaptativos. Ésta es una técnica iterativa, cuyo objetivo principal es resolver problemas difíciles en el campo de la optimización combinatoria, donde, en cada iteración del algoritmo se incluye una fase de construcción de una solución inicial, una de mejoramiento y una fase de actualización de la solución generada en fase inicial, es decir esta técnica construye inicialmente una solución que es mejorada inmediatamente en cada uno de las iteraciones del algoritmo.

Dentro del estudio de la metaheurística se esquematizó el algoritmo de solución para este tipo de problemas de optimización y se mostró su carácter genérico en el sentido de que si un problema puede ser escrito de la forma (1.6), este algoritmo de solución puede ser aplicado sin ningún tipo de adaptación. Además utilizando un ejemplo académico se visualizaron los principios elementales en los que está basado el método, como es el caso de la aleatoriedad, la cual nos permite generar un espectro

más amplio del conjunto de soluciones factibles. En la fase de implementación se pudo constatar la velocidad de convergencia del algoritmo, si consideramos que para el análisis se optó por considerar la mayor aleatoriedad posible

Con el objeto de constatar la fortaleza del método planteado en el contexto del problema en estudio se compararon los resultados obtenidos con la metaheurística GRASP, con una suerte de algoritmo alterno basado en el análisis de conglomerados. Este algoritmo alterno se construyó considerando que el propósito del análisis de conglomerados es el agrupar los sujetos o variables homogéneas en función de los valores observados dentro de un conjunto aparentemente heterogéneo de forma que los datos agrupados dentro de un mismo conglomerado sean homogéneos entre sí (mínima varianza) y que estos conglomerados sean lo más heterogéneos posibles entre ellos (máxima varianza), lo cual garantiza que a elegir miembros de diferentes conglomerados se obtiene una heterogeneidad entre las características de los mismos.

De la experimentación numérica se obtuvo que para el caso de la sucursal Ambato el 99,99 % de los resultados obtenidos con el GRASP alcanzan la diversidad máxima para esta regional, 1,57, mientras que más del 50 % de grupos formados utilizando análisis de conglomerados poseen una diversidad inferior a 1,02. En la sucursal Cuenca se tiene que el 99,99 % de los resultados obtenidos con el GRASP alcanzan la diversidad máxima del ejercicio, 1,40, mientras que más del 94 % de grupos formados utilizando análisis de conglomerados poseen una diversidad menor a 1,06. En la sucursal Guayaquil el 96,72 % de los resultados obtenidos con el GRASP alcanzan la diversidad máxima del ejercicio, 1,60, mientras que más del 94 % de grupos formados utilizando análisis de conglomerados poseen una diversidad menor a 1,17. En la sucursal Machala el 99,64 % de los resultados obtenidos con el GRASP alcanzan la diversidad máxima del ejercicio, 0,90, mientras que más del 97,1 % de grupos formados utilizando análisis de conglomerados poseen una diversidad menor a 0,63. Finalmente en la sucursal Quito el 99,29 % de los resultados obtenidos con el GRASP alcanzan la diversidad máxima del ejercicio, 0,79, mientras que más del 99,3 % de grupos formados utilizando análisis de conglomerados poseen una diversidad menor a 0,69.

Al comparar los resultados obtenidos de 5000 diferentes equipos conformados a través del algoritmo basado en análisis de conglomerados con el resultado obtenido a través de la metaheurística GRASP se evidenció que los resultados generados por la metaheurística fueron en un 97 % superiores a los obtenidos con este algoritmo alterno con lo cual se pudo evidenciar la fuerza y eficacia de la metaheurística GRASP.

Capítulo 5

Recomendaciones

Después de probar la metaheurística GRASP y constatar las ventajas de su uso en la conformación de equipos de trabajo, estamos seguros de que esta tesis contribuye a la mejora del proceso de selección del personal, no sólo en el caso particular que abordamos, sino en general; ya que las variables utilizadas son de carácter genérico en el sentido de que pueden ser reemplazadas por cualquier otra sin necesidad de adaptar el algoritmo de solución, por lo tanto se recomienda:

- 1 Ampliar el estudio de esta metaheurística para la conformación de un número arbitrario de grupos, de manera que la diversidad media de los grupos sea la función a maximizar, ya que en la metodología utilizada en esta tesis se pudo observar que la diversidad del primer grupo siempre fue claramente superior ya que por ejemplo para el caso de la sucursal Ambato se tiene que la diferencia entre las diversidades entre el primer y quinto grupos conformados utilizando la Metaheurística GRASP es en promedio del 37%.
- 2 Utilizar el análisis de conglomerados pues se observó que este puede ser un mecanismo idóneo para la generación de la solución inicial y se recomienda comprobar su eficiencia, puesto que este garantiza la máxima diversidad interelemento, lo cual podría optimizar la solución generada por el algoritmo GRASP y reducir sus tiempos de convergencia puesto que se pudo observar en la experimentación numérica desarrollada en el Capítulo 3, se obtuvieron resultados cercanos a la diversidad máxima en cada ejercicio (en un porcentaje bajo) sin ningún tipo de refinamiento posterior, lo cual nos hace pensar que la hipótesis propuesta es válida.

Anexo A

Anexo de bases de datos y resultados

En este capítulo se detalla la información sobre los postulantes a encuestadores y los resultados obtenidos de los diferentes análisis para la conformación de equipos de trabajo de las diferentes regionales del INEC.

Para las diferentes regionales se procederá de la siguiente manera: se inicia con el detalle de la información de cada uno de los postulantes así como su pertenencia a los diferentes conglomerados. A continuación se mostrará uno de los grupos de cinco equipos de trabajo obtenidos con el algoritmo GRASP. Luego se presentarán cinco de los grupos formados con análisis de conglomerados con mayor diversidad encontrados en el muestreo analizado en esta tesis.

A.1. Ambato

A continuación se presenta la información de los 50 candidatos de la regional Ambato.

ID	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información	Liderazgo	Comprensión oral	Comprensión escrita	Expresión Oral	Expresión escrita	Conglomerado
1	8	8	7	5	7	6	5	6	8	4	3
2	7	4	8	8	3	3	5	8	3	7	2

Cuadro A.1: Características postulantes regional Ambato

Fuente: Levantamiento del autor. Elaboración propia.

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

ID	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información	Liderazgo	Comprensión oral	Comprensión escrita	Expresión Oral	Expresión escrita	Conglomerado
3	7	7	5	6	4	4	8	6	7	3	4
4	10	7	5	4	5	3	4	7	8	6	4
5	9	4	7	3	8	6	8	8	3	4	5
6	8	5	6	6	8	7	3	3	5	7	2
7	7	4	5	8	8	3	7	4	4	8	1
8	9	6	8	6	6	3	4	7	7	4	3
9	7	7	7	4	4	4	4	3	5	5	3
10	9	4	8	7	8	8	6	4	6	7	1
11	7	7	5	8	5	4	8	8	5	3	3
12	7	5	3	4	6	5	6	7	4	6	5
13	8	8	4	7	6	7	3	3	7	3	3
14	10	6	4	4	3	3	6	3	6	4	4
15	10	4	4	5	5	7	8	6	3	5	5
16	10	7	5	6	3	6	7	7	8	7	4
17	9	6	4	8	3	7	7	5	6	4	3
18	6	4	7	5	7	6	8	3	8	8	1
19	9	5	4	3	7	5	3	7	6	6	2
20	7	3	7	3	3	7	8	4	3	8	1
21	8	4	4	8	7	7	4	8	5	8	2
22	7	8	4	8	6	6	8	8	7	5	3
23	10	3	5	8	6	7	3	6	4	7	2
24	6	4	3	4	7	6	8	7	7	5	5
25	8	3	5	8	8	4	5	4	3	3	2
26	7	3	8	5	7	5	6	6	4	6	5
27	8	5	4	6	4	3	3	8	4	3	2
28	10	3	3	6	4	7	5	5	4	3	2
29	10	7	3	6	6	8	8	3	7	6	1
30	7	8	6	4	8	6	3	5	4	8	2
31	9	6	5	7	5	5	4	8	7	5	3
32	8	3	7	6	4	7	4	5	8	3	3
33	6	8	7	8	4	7	3	5	5	3	3

Cuadro A.1: Características postulantes regional Ambato

Fuente: Levantamiento del autor. Elaboración propia.

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

ID	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información	Liderazgo	Comprensión oral	Comprensión escrita	Expresión Oral	Expresión escrita	Conglomerado
34	6	8	5	3	7	7	7	4	6	3	3
35	7	4	8	8	6	4	5	8	3	4	2
36	8	4	4	3	8	5	6	8	7	4	5
37	10	4	7	6	5	4	4	3	4	8	2
38	9	8	8	5	6	8	3	7	7	6	3
39	6	7	8	8	7	5	3	5	3	6	2
40	6	3	4	7	8	3	8	7	3	8	5
41	6	8	5	8	3	7	8	4	5	4	3
42	8	4	7	3	8	4	6	4	6	6	1
43	8	3	4	3	3	6	7	8	8	6	4
44	7	5	3	6	4	8	3	4	8	7	3
45	8	3	3	3	3	3	7	7	6	4	4
46	8	7	6	4	3	7	3	7	8	6	3
47	6	6	7	8	5	6	7	4	4	8	1
48	6	7	3	7	5	5	6	3	8	7	1
49	8	8	7	6	5	6	8	4	7	8	1
50	9	8	4	3	7	4	6	6	5	8	4

Cuadro A.1: Características postulantes regional Ambato

Fuente: Levantamiento del autor. Elaboración propia.

Con la información mostrada en el cuadro anterior se obtuvo mediante el algoritmo GRASP el siguiente grupo de cinco equipos.

Equipo	Elementos					Diversidad	Tamaño LC
Ambato 1	2	13	20	25	45	1,57157871	50
Ambato 2	5	33	40	43	44	1,37620596	45
Ambato 3	7	27	30	32	34	1,19280656	40
Ambato 4	35	36	37	41	46	1,13835708	35
Ambato 5	11	18	28	39	50	0,99031044	30

Cuadro A.2: Grupo de cinco equipos de trabajo formado utilizando el GRASP.

Fuente: Algoritmo GRASP. Elaboración propia.

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

En el Cuadro (1) se muestra la pertenencia de los distintos candidatos a los conglomerados formados. En el siguiente Cuadro se muestran los cinco equipos, formados a través de análisis de conglomerados, elegidos de forma aleatoria del total de combinaciones de equipos posibles.

Equipo	Elementos					Diversidad
Equipo 1	2	5	13	20	45	1,46876416
Equipo 2	5	13	20	25	43	1,46819011
Equipo 3	2	5	13	20	43	1,44596878
Equipo 4	5	20	25	44	45	1,41788133
Equipo 5	5	13	20	39	43	1,41683279

Cuadro A.3: Equipos formados con análisis de conglomerados.

Fuente: Método de conglomerados. Elaboración propia.

Se debe recalcar la similitud entre el equipo Ambato 1 y los equipos formados con análisis de conglomerados conforme su diversidad se acerca a la obtenida en el primer equipo de la sucursal Ambato.

A.2. Cuenca

En esta sección se muestra la información de los 54 candidatos de la regional Cuenca, así como su pertenencia a los diferentes conglomerados formados.

ID	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información	Liderazgo	Comprensión oral	Comprensión escrita	Expresión Oral	Expresión escrita	Conglomerado
1	8	8	6	3	5	5	6	3	8	3	1
2	6	8	3	5	6	5	8	4	6	4	5
3	6	7	7	6	5	7	6	7	3	5	2

Cuadro A.4: Características postulantes regional Ambato

Fuente: Levantamiento del autor. Elaboración propia.

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

ID	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información	Liderazgo	Comprensión oral	Comprensión escrita	Expresión Oral	Expresión escrita	Conglomerado
4	9	7	5	4	4	3	7	6	3	8	4
5	7	5	5	8	5	3	7	7	4	5	5
6	8	4	6	5	7	3	8	5	7	4	5
7	6	5	5	8	4	4	7	5	6	3	5
8	8	6	7	6	8	8	8	8	7	3	2
9	7	5	7	3	4	8	4	7	3	4	2
10	8	8	3	6	8	3	7	6	8	3	5
11	7	6	7	4	6	7	3	4	8	7	1
12	7	6	4	3	8	4	6	5	8	6	1
13	8	3	3	8	6	5	6	7	4	3	2
14	9	8	6	5	6	8	6	7	6	5	2
15	9	7	8	8	8	4	7	4	3	8	4
16	6	4	7	6	7	8	3	4	6	5	1
17	7	5	8	3	5	8	8	5	7	3	2
18	10	3	4	5	6	7	8	5	7	4	2
19	9	6	7	8	7	7	4	3	8	3	1
20	8	3	3	8	4	5	8	3	7	7	3
21	10	5	5	6	3	4	7	5	6	4	5
22	7	6	8	3	6	3	8	4	4	8	4
23	9	7	7	4	8	4	5	6	7	6	1
24	8	4	3	3	8	8	7	4	8	6	2
25	6	6	7	5	5	6	3	5	4	5	1
26	8	8	6	3	3	4	5	4	3	7	4
27	6	8	7	7	7	8	4	7	6	6	1
28	7	4	3	4	7	7	7	6	4	7	2
29	7	5	7	3	5	3	6	4	6	3	5
30	7	3	4	6	4	5	5	7	4	7	3
31	6	7	6	4	4	7	4	6	3	8	4
32	6	8	8	5	8	4	3	7	4	7	4
33	8	8	6	7	3	3	8	4	5	3	5
34	9	7	3	7	3	7	8	8	8	8	3
35	9	4	6	4	3	4	8	7	5	7	3
36	9	6	7	8	3	3	7	6	5	3	5
37	7	5	8	7	6	6	7	7	3	4	2
38	8	8	5	3	4	3	8	3	6	7	5

Cuadro A.4: Características postulantes regional Ambato

Fuente: Levantamiento del autor. Elaboración propia.

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

ID	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información	Liderazgo	Comprensión oral	Comprensión escrita	Expresión Oral	Expresión escrita	Conglomerado
39	8	6	3	7	5	7	3	7	5	5	1
40	6	8	8	3	6	7	4	5	6	4	1
41	9	7	6	3	5	5	7	7	3	3	5
42	8	5	7	6	5	6	5	6	7	5	1
43	9	8	6	6	5	3	6	7	7	4	5
44	7	6	7	7	5	5	7	6	4	5	5
45	7	5	4	5	6	5	5	5	6	6	1
46	6	4	5	5	6	7	4	7	6	3	2
47	7	6	7	6	7	7	7	6	8	5	2
48	7	8	7	6	6	6	4	5	8	3	1
49	9	8	8	8	4	3	8	7	7	7	5
50	8	4	6	3	5	3	6	7	4	6	4
51	6	8	6	3	6	3	3	6	5	7	4
52	6	5	7	4	8	4	4	5	3	5	4
53	8	3	8	3	5	8	7	7	4	6	2
54	10	3	7	3	3	4	4	5	4	4	4

Cuadro A.4: Características postulantes regional Cuenca

Fuente: Levantamiento del autor. Elaboración propia.

Con la información mostrada en el cuadro anterior se obtuvo mediante el algoritmo GRASP el siguiente grupo de cinco equipos.

Equipo	Elementos					Diversidad	Tamaño LC
Cuenca 1	9	20	24	33	51	1,40499497	54
Cuenca 2	13	19	26	38	53	1,31293224	49
Cuenca 3	10	17	22	31	39	1,14920768	44
Cuenca 4	1	4	16	28	36	1,12993926	39
Cuenca 5	15	18	32	34	40	1,01170592	34

Cuadro A.5: Grupo de cinco equipos de trabajo formado utilizando el GRASP.

Fuente: Algoritmo GRASP. Elaboración propia.

En el Cuadro (4) se muestra la pertenencia de los distintos candidatos a los conglomerados formados. En el siguiente Cuadro se muestran los cinco equipos,

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

formados a través de análisis de conglomerados, elegidos de forma aleatoria del total de combinaciones de equipos posibles.

Equipo	Elementos					Diversidad
Equipo 1	19	9	20	51	2	1,24720289
Equipo 2	19	9	20	22	2	1,23235665
Equipo 3	19	9	20	26	2	1,23164542
Equipo 4	19	9	20	32	2	1,22405254
Equipo 5	19	53	20	51	2	1,21681638

Cuadro A.6: Equipos formados con análisis de conglomerados.

Fuente: Método de conglomerados. Elaboración propia.

A.3. Guayaquil

A continuación se presenta la información de los 68 candidatos de la regional Guayaquil y su pertenencia a los conglomerados formados.

ID	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información	Liderazgo	Comprensión oral	Comprensión escrita	Expresión Oral	Expresión escrita	Conglomerado
1	7	8	8	2	3	2	6	8	3	7	3
2	7	4	6	8	8	2	7	4	7	6	2
3	7	7	5	7	5	2	7	7	6	7	2
4	10	7	6	2	4	5	7	5	6	7	3
5	8	6	5	3	3	4	4	3	5	6	3
6	6	5	7	8	7	3	8	4	4	7	2
7	7	7	6	3	6	4	8	3	4	7	3
8	6	8	6	8	7	7	8	8	5	8	4
9	9	6	8	8	3	3	4	5	7	3	1
10	9	8	6	4	8	8	8	4	8	5	4

Cuadro A.7: Características postulantes regional Guayaquil

Fuente: Levantamiento del autor. Elaboración propia.

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

ID	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información	Liderazgo	Comprensión oral	Comprensión escrita	Expresión Oral	Expresión escrita	Conglomerado
11	8	7	7	3	7	8	4	5	5	6	4
12	7	4	7	2	5	6	8	5	5	7	3
13	8	8	4	8	2	2	6	6	4	6	3
14	6	5	7	7	5	5	4	4	7	3	1
15	6	6	4	4	5	3	8	6	5	7	3
16	10	7	7	7	8	8	8	6	7	8	4
17	8	6	4	2	3	3	7	4	4	5	3
18	6	7	8	3	3	6	4	7	8	5	3
19	8	7	5	7	8	8	7	8	3	8	4
20	8	6	8	6	8	4	8	7	5	7	2
21	9	8	6	5	8	5	8	7	8	3	4
22	7	5	4	2	2	8	8	6	8	7	3
23	7	7	4	4	7	7	7	4	7	7	4
24	8	6	5	8	2	7	4	8	4	6	5
25	10	7	4	8	8	8	8	5	7	5	4
26	9	7	7	8	8	7	4	7	8	7	4
27	6	8	4	7	7	2	8	8	6	7	2
28	10	7	8	3	2	2	8	7	5	7	3
29	10	6	4	3	7	6	8	8	7	8	4
30	8	5	8	7	8	4	6	8	4	3	1
31	9	5	4	8	7	7	3	7	7	5	1
32	10	8	6	2	4	3	4	4	7	8	3
33	8	7	5	5	2	2	7	6	4	6	3
34	9	4	4	8	4	2	3	7	4	4	1
35	9	7	5	2	3	4	6	8	6	8	3
36	8	7	7	5	8	7	3	3	3	7	1
37	8	7	4	6	2	7	6	5	8	5	5
38	7	5	6	7	6	8	6	8	5	7	5
39	6	6	8	2	5	3	7	5	6	7	3
40	6	6	4	2	8	7	8	8	3	6	4
41	9	6	7	8	8	4	4	8	8	7	2
42	6	5	8	8	8	2	7	5	6	7	2
43	10	5	8	6	2	7	3	7	8	8	5
44	10	8	6	8	7	8	8	8	7	3	4

Cuadro A.7: Características postulantes regional Guayaquil

Fuente: Levantamiento del autor. Elaboración propia.

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

ID	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información	Liderazgo	Comprensión oral	Comprensión escrita	Expresión Oral	Expresión escrita	Conglomerado
45	10	8	8	3	2	7	7	7	8	3	3
46	8	6	8	3	8	4	6	4	8	8	2
47	8	7	8	7	7	6	3	4	3	3	1
48	7	6	6	2	6	7	3	8	5	3	1
49	9	7	6	6	7	4	4	7	5	8	2
50	9	7	8	2	3	4	3	8	6	4	3
51	10	7	5	8	3	7	7	4	8	7	5
52	7	7	8	3	3	2	4	7	7	4	3
53	7	7	5	8	5	7	4	8	6	7	5
54	9	8	7	7	8	3	6	4	8	8	2
55	7	8	6	3	3	2	3	7	5	8	3
56	8	8	5	7	3	7	5	8	3	8	5
57	10	8	7	3	8	6	6	8	7	6	4
58	6	6	7	2	2	7	7	5	3	5	3
59	8	4	4	3	2	2	6	7	4	6	3
60	6	8	5	7	8	8	7	7	8	7	4
61	6	6	4	8	2	3	8	4	7	7	2
62	10	5	8	7	6	2	7	6	3	7	2
63	10	7	6	2	7	4	3	7	5	7	3
64	7	8	5	5	7	7	8	6	5	4	4
65	7	8	7	6	8	7	6	8	8	8	4
66	9	8	5	4	5	5	7	7	6	7	3
67	7	5	8	7	2	7	5	6	7	8	5
68	8	5	7	5	8	2	6	5	8	7	2

Cuadro A.7: Características postulantes regional Guayaquil

Fuente: Levantamiento del autor. Elaboración propia.

Con la información mostrada en el cuadro anterior se obtuvo mediante el algoritmo GRASP el siguiente grupo de cinco equipos.

Equipo	Elementos					Diversidad	Tamaño LC
Guayaquil 1	1	2	22	34	47	1,59752359	68
Guayaquil 2	9	36	40	50	61	1,49364561	63
Guayaquil 3	13	30	31	32	58	1,33222698	58
Guayaquil 4	6	42	45	48	55	1,24458863	53
Guayaquil 5	24	25	28	46	52	1,14715225	48

Cuadro A.8: Grupo de cinco equipos de trabajo formado utilizando el GRASP.

Fuente: Algoritmo GRASP. Elaboración propia.

En el Cuadro (7) se muestra la pertenencia de los distintos candidatos a los conglomerados formados. En el siguiente Cuadro se muestran los cinco equipos, formados a través de análisis de conglomerados, elegidos de forma aleatoria del total de combinaciones de equipos posibles.

Equipo	Elementos					Diversidad
Equipo 1	2	24	40	47	52	1,37211743
Equipo 2	2	9	24	40	58	1,37710986
Equipo 3	2	24	36	40	50	1,37796526
Equipo 4	2	24	36	40	45	1,38032729
Equipo 5	1	2	24	40	47	1,4007511

Cuadro A.9: Equipos formados con análisis de conglomerados.

Fuente: Método de conglomerados. Elaboración propia.

Se debe recalcar la similitud entre el equipo Guayaquil 1 y los equipos formados con análisis de conglomerados conforme su diversidad se acerca a la obtenida en el primer equipo de la sucursal Guayaquil.

A.4. Machala

A continuación se presenta la información de los 48 candidatos de la regional Machala y su pertenencia a los conglomerados formados.

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

ID	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información	Liderazgo	Comprensión oral	Comprensión escrita	Expresión Oral	Expresión escrita	Conglomerado
1	6	8	8	8	4	7	4	4	8	6	2
2	9	5	4	5	6	4	5	8	6	6	5
3	6	7	8	7	5	8	8	4	7	5	2
4	8	8	6	8	5	6	4	5	8	6	2
5	8	4	4	4	4	5	8	8	5	5	5
6	9	5	8	8	8	5	4	5	8	7	4
7	6	8	5	6	7	8	4	5	6	7	3
8	7	4	4	7	7	8	6	5	6	4	3
9	9	4	6	6	6	5	6	5	5	8	4
10	7	6	4	6	8	4	4	6	7	7	4
11	6	5	5	8	5	4	5	4	4	7	2
12	8	6	7	5	6	6	5	4	8	8	2
13	6	7	8	8	5	4	8	7	8	6	1
14	7	6	8	6	6	8	8	6	6	8	2
15	10	8	7	4	5	5	6	8	5	8	1
16	6	6	6	4	5	5	5	4	4	5	2
17	6	6	7	7	7	4	7	6	6	6	4
18	6	8	4	8	5	7	8	7	6	6	1
19	7	5	6	6	7	4	6	4	7	7	4
20	8	5	6	6	4	5	6	8	8	8	1
21	6	5	7	5	8	6	4	6	5	5	3
22	6	5	5	5	6	6	4	7	5	4	5
23	7	5	4	4	5	7	5	6	5	6	5
24	8	8	6	7	8	8	4	6	5	6	3
25	6	4	4	6	8	8	8	4	5	6	3
26	7	4	5	8	8	4	7	6	7	6	4
27	6	4	8	7	8	4	5	5	5	8	4
28	7	6	8	7	4	6	6	4	6	5	2
29	6	4	7	6	6	5	4	8	8	7	4

Cuadro A.10: Características postulantes regional Machala

Fuente: Levantamiento del autor. Elaboración propia.

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

ID	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información	Liderazgo	Comprensión oral	Comprensión escrita	Expresión Oral	Expresión escrita	Conglomerado
30	7	7	8	6	6	5	6	6	4	7	2
31	9	8	5	8	6	7	7	8	4	6	1
32	6	7	5	4	8	6	5	4	6	4	3
33	6	7	4	4	6	5	4	4	6	5	3
34	6	7	5	7	8	4	6	7	5	5	3
35	6	7	8	6	4	8	5	6	8	7	2
36	7	5	4	4	6	5	5	5	8	6	4
37	8	7	6	6	7	7	5	8	6	7	3
38	7	6	5	7	5	4	4	7	4	8	5
39	7	7	7	7	5	6	8	4	7	4	2
40	6	6	7	4	4	8	7	4	4	7	2
41	6	4	7	4	8	4	7	5	8	4	4
42	7	6	7	7	5	6	8	6	4	4	2
43	6	8	4	6	4	5	6	7	7	8	1
44	10	5	4	5	8	4	4	5	4	6	3
45	7	7	5	7	5	6	8	4	5	8	2
46	8	5	6	5	6	6	4	8	6	7	5
47	8	8	6	4	8	8	8	5	6	5	3
48	8	7	6	5	4	4	7	8	7	4	1

Cuadro A.10: Características postulantes regional Machala

Fuente: Levantamiento del autor. Elaboración propia.

Con la información mostrada en el cuadro anterior se obtuvo mediante el algoritmo GRASP el siguiente grupo de cinco equipos.

Equipo	Elementos					Diversidad	Tamaño LC
Machala 1	1	5	25	38	41	0,89736621	48
Machala 2	8	10	27	40	48	0,77122667	43
Machala 3	11	29	32	42	43	0,64393599	38
Machala 4	3	15	26	35	44	0,61055737	33
Machala 5	6	20	31	39	47	0,58984975	28

Cuadro A.11: Grupo de cinco equipos de trabajo formado utilizando el GRASP.

Fuente: Algoritmo GRASP. Elaboración propia.

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

En el Cuadro (10) se muestra la pertenencia de los distintos candidatos a los conglomerados formados. En el siguiente Cuadro se muestran los cinco equipos, formados a través de análisis de conglomerados, elegidos de forma aleatoria del total de combinaciones de equipos posibles.

Equipo	Elementos					Diversidad
Equipo 1	1	2	15	25	41	0,78264163
Equipo 2	1	2	25	41	43	0,76374143
Equipo 3	1	2	8	15	41	0,73998863
Equipo 4	1	2	18	25	41	0,73277238
Equipo 5	2	25	40	41	43	0,73092956

Cuadro A.12: Equipos formados con análisis de conglomerados.

Fuente: Método de conglomerados. Elaboración propia.

Fuente: Método de conglomerados. Elaboración propia.

Se debe recalcar la similitud entre el equipo Machala 1 y los equipos formados con análisis de conglomerados conforme su diversidad se acerca a la obtenida en el primer equipo de la sucursal Machala.

A.5. Quito

En esta sección presenta la información de los 62 candidatos de la regional Quito y su pertenencia a los conglomerados formados.

ID	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información	Liderazgo	Comprensión oral	Comprensión escrita	Expresión Oral	Expresión escrita	Conglomerado
1	10	4	7	6	5	6	7	7	6	4	3
2	9	7	4	4	6	8	8	7	6	8	2
3	8	7	6	5	5	8	6	7	7	5	2
4	7	5	8	7	4	7	7	8	6	4	3
5	7	5	4	4	7	4	7	7	8	6	5

Cuadro A.13: Características postulantes regional Machala

Fuente: Levantamiento del autor. Elaboración propia.

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

ID	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información	Liderazgo	Comprensión oral	Comprensión escrita	Expresión Oral	Expresión escrita	Conglomerado
6	8	5	7	8	4	7	4	8	6	5	4
7	8	5	6	5	5	4	7	7	6	6	5
8	10	8	4	4	7	5	7	4	6	7	2
9	7	8	7	4	5	5	4	8	7	4	4
10	8	5	6	7	8	4	4	6	7	4	3
11	6	4	5	6	5	7	7	6	6	4	3
12	7	4	5	5	7	8	6	5	7	8	2
13	7	4	7	4	7	4	7	4	4	6	5
14	8	5	7	7	7	4	7	4	5	4	3
15	9	7	4	7	7	7	7	5	7	4	3
16	7	7	4	4	4	6	8	8	5	6	2
17	7	5	7	6	6	7	7	6	8	7	2
18	8	7	4	6	7	8	4	4	6	8	1
19	8	6	6	5	7	7	8	5	6	7	2
20	8	6	4	7	6	8	7	6	7	7	2
21	8	4	4	4	6	5	8	7	6	4	5
22	9	8	6	5	6	6	6	6	8	8	2
23	8	8	4	7	6	7	8	8	7	4	3
24	8	4	4	7	7	7	7	6	7	6	3
25	9	8	5	4	5	7	7	4	5	4	2
26	10	6	4	4	8	6	6	8	5	4	5
27	7	7	6	7	4	4	4	6	7	5	4
28	8	4	8	4	6	5	6	6	4	4	5
29	6	7	5	5	4	4	7	7	4	6	5
30	7	8	4	5	4	6	6	7	7	7	2
31	6	7	7	6	8	5	4	4	4	7	1
32	9	6	8	4	7	5	8	8	4	8	5
33	10	8	6	8	6	4	6	4	8	7	1
34	8	7	4	7	7	7	7	6	7	5	3
35	6	6	6	4	4	7	5	4	6	6	4
36	9	5	8	4	7	4	4	6	7	4	4

Cuadro A.13: Características postulantes regional Machala

Fuente: Levantamiento del autor. Elaboración propia.

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

ID	Pruebas técnicas	Capacidad para resolver problemas lógicos	Iniciativa	Capacidad para seguir órdenes	Organización de información	Liderazgo	Comprensión oral	Comprensión escrita	Expresión Oral	Expresión escrita	Conglomerado
37	9	7	5	7	6	4	6	6	7	7	3
38	9	7	7	7	5	4	8	7	6	6	3
39	7	7	5	5	7	7	4	7	7	8	2
40	9	7	7	7	4	5	8	6	7	7	3
41	9	4	6	5	8	7	6	6	7	6	5
42	7	6	6	4	4	7	4	5	5	6	4
43	7	7	5	8	6	7	4	4	7	6	1
44	9	4	6	4	7	4	6	7	7	5	5
45	7	7	7	5	7	6	7	8	6	4	3
46	8	8	4	4	5	5	4	4	7	6	1
47	8	4	5	6	8	7	6	7	7	7	5
48	7	7	4	6	7	7	8	4	5	4	3
49	7	4	6	4	5	4	6	7	7	8	5
50	7	5	5	8	7	4	8	6	6	4	3
51	7	4	5	6	4	4	6	5	5	6	5
52	8	7	4	4	4	4	5	5	5	6	4
53	9	6	4	4	6	7	4	4	4	5	1
54	7	5	8	8	5	7	4	4	6	6	1
55	7	8	6	7	5	4	8	5	5	4	3
56	8	7	6	4	7	4	6	7	5	6	5
57	7	5	6	4	4	5	4	6	6	4	4
58	7	7	5	7	6	4	7	6	6	4	3
59	10	8	5	5	7	4	5	5	4	4	1
60	7	5	4	4	7	4	6	6	6	7	5
61	8	6	7	5	6	4	7	6	7	4	3
62	10	6	7	6	8	6	6	4	7	7	1

Cuadro A.13: Características postulantes regional Quito

Fuente: Levantamiento del autor. Elaboración propia.

Con la información mostrada en el cuadro anterior se obtuvo mediante el algoritmo GRASP el siguiente grupo de cinco equipos.

DISEÑO E IMPLEMENTACIÓN DE UN MODELO MATEMÁTICO DE OPTIMIZACIÓN PARA LA CONFORMACIÓN DE EQUIPOS DE TRABAJO

Equipo	Elementos					Diversidad	Tamaño LC
Quito 1	6	9	13	18	21	0,78773676	62
Quito 2	16	36	48	49	54	0,76091943	57
Quito 3	4	26	31	46	50	0,70000467	52
Quito 4	10	12	25	27	32	0,68257405	47
Quito 5	2	5	28	43	55	0,6542697	42

Cuadro A.14: Grupo de cinco equipos de trabajo formado utilizando el GRASP.

Fuente: Algoritmo GRASP. Elaboración propia.

En el Cuadro (13) se muestra la pertenencia de los distintos candidatos a los conglomerados formados. En el siguiente Cuadro se muestran los cinco equipos, formados a través de análisis de conglomerados, elegidos de forma aleatoria del total de combinaciones de equipos posibles.

Equipo	Elementos					Diversidad
Equipo 1	5	6	16	31	48	0,7314184
Equipo 2	5	9	14	16	54	0,72962661
Equipo 3	4	5	6	8	31	0,72685727
Equipo 4	5	9	14	16	18	0,72380535
Equipo 5	5	6	16	48	54	0,7216881

Cuadro A.15: Equipos formados con análisis de conglomerados.

Fuente: Método de conglomerados. Elaboración propia.

Bibliografía

- [1] CUADRAS, C. M. *Nuevos métodos de análisis multivariante*. CMC Editions, 2014.
- [2] DUARTE, A., AND MARTÍ, R. Tabu search and grasp for the maximum diversity problems. *Tabu search and grasp for the maximum diversity problema 178* (2007), 71–84.
- [3] DUARTE MUNOZ, A. Metaheurísticas. *Librería-Editorial Dykinson* (May 8, 2007), pp. 1–5.
- [4] GHOSH, J. Computational aspects of the maximum diversity problem. *Operations research letters 19* (1996), 175–181.
- [5] GLOVER, F., KUO, C., AND DHIR, K. Heuristic algorithms for the maximum diversity problem. *Journal of information and optimization sciences 19*, 1 (1998), 109–132.
- [6] KOCHENBERGER, G., AND GLOVER, F. Diversity data mining, Working Paper. *The University of Mississippi* (1999).
- [7] LAGUNA, M., AND MARTÍ, R. GRASP and path relinking for 2-layer straight line crossing minimization. *INFORMS Journal on Computing* (1999), 44–52.
- [8] MARTÍ, R., AND SANDOYA, F. GRASP and path relinking for the equitable dispersion problem. *Computers and Operations Research Volume 40* (2013), Pages 3091–3099.
- [9] NEWELL, A., AND SIMON, H. Human Problem Solving. *Prentice - Hall, Englewood Cliffs* (1972).
- [10] POLZER, J., MILTON, L., AND SWANN JR., W. . *Admin. Sci. Q. Volume 47* (2002), pp. 296–327.
- [11] REINOSO, R. *Computers and Operations Research Volume 40, Issue 12* (2013), pp. 3091–3099.
- [12] RESENDE, M., MARTÍ, R., GALLEGU, M., AND DUARTE, A. GRASP with past relinking for the max-min diversity problem. *Computers and Operations Research 37* (2010), 498–508.

- [13] RESENDE, M., R. C. GRASP with path-relinking: recent advances and applications. *Metaheuristics: Progress as Real Problem Solvers (T. Ibaraki et al. editors)* (2005), 29–63.
- [14] RODRÍGUEZ ORTIZ, C. *Algoritmos heurísticos y metaheurísticos para el problema de localización de regeneradores*. Proyecto fin de carrera, Universidad Rey Juan Carlos, 2010.
- [15] SANDOYA, F., AND ACEVES, R. Grasp and Path Relinking to Solve the Problem of Selecting Efficient Work Teams, Recent Advances on Meta-Heuristics and Their Application to Real Scenarios. *InTech* (2013).
- [16] SANDOYA, F., AND MARTI, R. Algoritmo GRASP híbrido para resolver una nueva variante del problema de la diversidad máxima. *Congreso Latinoamericano de Investigación de operaciones* (2012).