

## TABLA DE CONTENIDOS

	página
Dedicatoria	I
Agradecimientos	II
Tabla de Contenidos	IV
Índice de Figuras	VI
Índice de Tablas	VIII
Resumen	IX
<b>1. Introducción</b>	<b>10</b>
1.1. Contexto . . . . .	10
1.1.1. Objetivo General . . . . .	10
1.1.2. Objetivos Específicos . . . . .	11
1.2. Alcances . . . . .	11
1.3. Descripción de los contenidos . . . . .	11
<b>2. Antecedentes</b>	<b>13</b>
2.1. Espacios Métricos . . . . .	13
2.1.1. Consultas de Proximidad . . . . .	14
2.1.2. Maldición de la dimensionalidad . . . . .	15
2.1.3. Ejemplos de espacios métricos . . . . .	17
2.2. Índices . . . . .	19
2.2.1. Estrategias utilizadas en espacios métricos . . . . .	19
2.2.2. Métodos basados en clusters [7] . . . . .	22
2.3. Manejo de Memoria Secundaria . . . . .	23
2.3.1. Sistemas de Manejo de Archivos [10] . . . . .	23
2.4. Trabajo Previo: DLC y DSC . . . . .	29
2.4.1. Dynamic List of Clusters . . . . .	29
2.4.2. Dynamic Set of Clusters . . . . .	31

<b>3. Metodología</b>	<b>33</b>
3.1. Desarrollo y proceso . . . . .	33
3.2. Obtención de los datos . . . . .	33
3.3. Presentación e implementación . . . . .	34
3.4. Pruebas experimentales . . . . .	34
3.5. Análisis de los resultados . . . . .	35
<b>4. Desarrollo del Trabajo</b>	<b>36</b>
4.1. Diseño del Índice . . . . .	36
4.1.1. Estrategia de Ordenamiento . . . . .	37
4.2. Implementación del Índice . . . . .	45
4.2.1. Ambiente de Programación . . . . .	45
4.2.2. Algoritmos . . . . .	46
<b>5. Resultados Experimentales</b>	<b>54</b>
5.1. Descripción de los experimentos . . . . .	54
5.1.1. Experimentos de construcción . . . . .	55
5.1.2. Experimentos de búsqueda por rango . . . . .	55
5.2. Resultados . . . . .	55
5.2.1. Bases de datos utilizadas . . . . .	55
5.2.2. Resultados de experimentos de construcción . . . . .	56
5.2.3. Resultados de experimentos de búsqueda por rango . . . . .	65
<b>6. Conclusiones y Trabajo Futuro</b>	<b>73</b>
<b>Bibliografía</b>	<b>75</b>

## ÍNDICE DE FIGURAS

	página
2.1. Consulta por rango. . . . .	14
2.2. Consulta por los k vecinos cercanos. . . . .	15
2.3. Histogramas de distancias de espacios métricos de baja y alta dimen- sionalidad. . . . .	16
2.4. Ejemplo de indexación y consultas. . . . .	20
4.1. Partes del Índice. . . . .	37
4.2. Clase Elemento. . . . .	39
4.3. Clases Vector y String. . . . .	40
4.4. Clase Pivote. . . . .	40
4.5. Clases Cargador y Consultor. . . . .	41
4.6. Ejemplo de creación de cluster con algoritmo diseñado. . . . .	42
5.1. Comparación interna del índice en cálculo de métrica. . . . .	56
5.2. Comparación interna del índice en escrituras en disco. . . . .	56
5.3. Uso de RAM en construcción E.V. dim. 208 . . . . .	57
5.4. Captura de la ejecución del método de pivotes. . . . .	57
5.5. Construcción: Cálculo de distancia E.V. dim. 20. . . . .	58
5.6. Construcción: Escrituras a disco E.V. dim. 20. . . . .	58
5.7. Construcción: Lecturas de disco E.V. dim. 20. . . . .	59
5.8. Construcción: Mov. de cabezal de disco E.V. dim. 20. . . . .	59
5.9. Construcción: BOLDLC vs Pivotes, cálc. de dist. E.V. dim. 20. . . . .	60
5.10. Construcción: Uso de RAM para E.V. dim. 20. . . . .	60
5.11. Construcción: Cálculos de distancia E.V. dim. 112. . . . .	61
5.12. Contrucción: Escrituras en disco E.V. dim 112. . . . .	61
5.13. Construcción: Lecturas de disco E.V. dim 112. . . . .	62
5.14. Construcción: Mov. de cabezal de disco E.V. dim 112 . . . . .	62
5.15. Construcción: BOLDLC vs Pivotes, cálc. de distancia E.V. dim 112 .	62
5.16. Construcción: Uso de RAM para E.V. dim. 112. . . . .	62
5.17. Construcción: Cálculos de distancia en diccionario. . . . .	63
5.18. Construcción: Escrituras a disco en diccionario. . . . .	63

5.19. Construcción: Lecturas de disco en diccionario . . . . .	64
5.20. Construcción: Mov. de cabezal de disco en diccionario . . . . .	64
5.21. Construcción: BOLDLC vs Pivotes, cálc. dist. en diccionario. . . . .	64
5.22. Construcción: Uso de RAM sobre diccionario. . . . .	64
5.23. Búsqueda: Lecturas de disco sobre E.V. dim. 20 . . . . .	66
5.24. Búsqueda: Cálc. de dist. promedio sobre E.V. dim. 20 . . . . .	66
5.25. Búsqueda: Mov. de cab. sobre E.V. dim. 20 . . . . .	67
5.26. Búsqueda: Tiempo prom. sobre E.V. dim. 20 . . . . .	67
5.27. Búsqueda: BOLDLC vs Pivotes, Cálc. de dist. E.V. dim 20. . . . .	67
5.28. Búsqueda: BOLDLC vs Pivotes, Tiempo E.V. dim. 20. . . . .	67
5.29. Búsqueda: Lecturas de disco sobre E.V. dim. 112. . . . .	68
5.30. Búsqueda: Cálc. de dist. promedio sobre E.V. dim. 112. . . . .	68
5.31. Búsqueda: Mov. de cab. sobre E.V. dim. 112 . . . . .	69
5.32. Búsqueda: Tiempo prom. sobre E.V. dim. 112. . . . .	69
5.33. Búsqueda: BOLDLC vs Pivotes, Cálc. de dist. E.V. dim 112. . . . .	69
5.34. Búsqueda: BOLDLC vs Pivotes, Tiempo E.V. dim. 112. . . . .	69
5.35. Búsqueda: Lecturas sobre diccionario . . . . .	70
5.36. Búsqueda: Cálc. de dist. sobre diccionario . . . . .	70
5.37. Búsqueda: Mov. de cab. en diccionario. . . . .	71
5.38. Búsqueda: Tiempo prom. en diccionario. . . . .	71
5.39. Búsqueda: BOLDLC vs Pivotes, Cálc. de dist. en diccionario. . . . .	71
5.40. Búsqueda: BOLDLC vs Pivotes, Tiempo en diccionario. . . . .	71

## ÍNDICE DE TABLAS

	página
5.1. Construcción: Cálculo de distancia E.V. dim. 20. . . . .	58
5.2. Construcción: Escrituras a disco. E.V. dim. 20. . . . .	58
5.3. Contrucción: Cálculos de distancia E.V. dim 112 . . . . .	61
5.4. Construcción: Escrituras a disco E.V. dim 112. . . . .	61
5.5. Construcción: Cálculo de distancia, diccionario inglés. . . . .	63
5.6. Construcción: Escrituras a disco, diccionario inglés. . . . .	63
5.7. Búsqueda: BOLDLC sobre B.D. de dim. 208. . . . .	66