

**PENGOPTIMUMAN SKEMA LOGIKAL PANGKALAN DATA HUBUNGAN :  
KAJIAN KES PANGKALAN DATA SISTEM MAKLUMAT HIDROLOGI  
MALAYSIA**

NORAZRIN BINTI KURMIN

Tesis ini dikemukakan  
sebagai memenuhi syarat penganugerahan  
ijazah Sarjana Sains (Sains Komputer)

Fakulti Sains Komputer dan Sistem Maklumat  
Universiti Teknologi Malaysia

MAC 2006

*Dengan Nama ALLAH Yang Maha Pemurah Lagi Maha Penyayang..*

*“Buat mak **Salmiah bt Daris** dan abah **Kurmin b Samsuri** yang tercinta,*

***Abg Azly, Acik, Iwan dan Mira** yang dikasihi;*

*Pengorbanan, jasa dan sokongan kalian amat dihargai.*

*Buat teman-teman seperjuangan yang disayangi dan dikasihi fillah;*

***Zam, Yus, Roha, Kuyu, GG, Wid, Norin, Shima, Syu, Yatie, Naym, Chah, Zoo, E, Ct,***

***Irwan, Shah, Razu, Nita, Sue & kawan-kawan;***

*Persahabatan, bantuan, sokongan dan kenangan bersama kalian akan kuabadikan*

*dilubuk hati dan akan kukenang hingga ke akhir hayat.*

*Buat yang istimewa, **Abang Pie**, yang sentiasa menyuntik semangat serta memberi*

*sokongan dan dorongan tanpa jemu,*

*Terima kasih atas pengorbanan, tunjuk ajar dan kesudian untuk bersama-sama  
mengharungi cabaran arus pendidikan dan kehidupan yang menuntut ketahanan serta*

*kesabaran diri yang terlalu tinggi, daku hargai segala-galanya. Semoga bahagia itu*

*berkekalan selamanya...*

*Mudah-mudahan kita semua akan beroleh keredhaan Allah S.W.T dan berbahagia dunia*

*akhirat...Amin..”*

## PENGHARGAAN

“Dengan Nama Allah Yang Maha Pemurah Lagi Maha Penyayang”

Syukur yang seagungnya buat Dia yang Maha Memiliki segalanya, yang memberi sebelum diri meminta, yang memberi lebih daripada yang diminta, lalu dengan kemurahanNya serta dengan izinNya jua saya berjaya menyiapkan projek sarjana ini. Pertama-tamanya, setinggi penghargaan kepada Dr Harihodin Selamat, En Mohd Shafry Mohd Rahim, En Mohd Taib Wahid selaku penyelia, terima kasih di atas segala bimbingan dan tunjuk ajar yang diberikan semasa sesi perbincangan sepanjang proses menyiapkan tesis ini.

Salam penghargaan buat insan-insan istimewa iaitu mak, abah, abang dan adik-adik serta yang sentiasa menyuntik semangat serta memberi sokongan dan dorongan tanpa jemu. Kepada sahabat-sahabat seperjuangan, samada yang terlibat secara langsung atau tidak, jasa baik dan tunjuk ajar daripada kalian amat dihargai.

Akhir kalam, salam penghargaan dan kasih sayang buat semua yang terlibat dalam membantu sehingga terhasilnya tesis ini. Semoga ada kebaikan dengan izinNya jua.

Norazrin Binti Kurmin  
Fakulti Sains Komputer dan Sistem Maklumat  
Universiti Teknologi Malaysia

## ABSTRAK

Dalam satu pangkalan data, adalah penting untuk memastikan penyimpanan, capaian semula dan manipulasi terhadap data dapat dikendalikan dengan berkesan. Penyelidikan ini memfokuskan kepada masalah capaian data di dalam pangkalan data. Isu utama yang sering mendapat perhatian para penyelidik adalah masa pelaksanaan yang lambat dalam proses capaian data. Faktor yang menyebabkan berlaku masalah tersebut adalah struktur pangkalan data yang tidak teratur. Selain itu, jumlah rekod di dalam pangkalan data adalah terlalu banyak. Oleh itu, penyelidikan ini menghasilkan satu teknik pengoptimuman skema logikal pangkalan data hubungan bagi memperbaiki masalah tersebut. Teknik yang dihasilkan adalah berdasarkan kaedah gabungan kesamaan iaitu dua entiti yang mempunyai atribut yang sama akan digabungkan dan juga pengkelasan tuple. Seterusnya, pernyataan pertanyaan akan dianalisis untuk menghasilkan pelaksanaan yang paling optimum. Pengujian perbandingan juga dilakukan diantara teknik yang telah dihasilkan dan teknik sebelumnya iaitu sistem kiub SMHM. Hasil penyelidikan membuktikan bahawa teknik yang dihasilkan dapat memperbaiki masa pelaksanaan pemprosesan pertanyaan, seterusnya dapat menyelesaikan isu yang dibincangkan.

## ABSTRACT

It is important to ensure that storing, retrieving, and data manipulating could be done efficiently in a database. This research is focused on the data access in Sistem Maklumat Hidrologi Malaysia (SMHM). The main issue that attracts researchers is low execution time in accessing data. The factor that causes the problem is the structure of the database. Besides, there is such a huge amount of data in SMHM database. Hence, this research produces an optimization technique of relational database logical scheme in order to fix the problem. The technique that has been produced is based on equijoin technique in which two entities that hold the same attribute will be joined and also the tuple classification. In addition, query will be analyzed in order to produce a maximum optimal execution. Comparison has been done on the techniques that have been developed and the SMHM cube system which is the previous technique. The results of the research prove that the technique that has been produced could improve the execution time for query processor, and solve the issue that has been discussed.

## KANDUNGAN

<b>BAB</b>	<b>PERKARA</b>	<b>MUKA SURAT</b>
	<b>PENGESAHAN STATUS TESIS</b>	
	<b>PENGESAHAN PENYELIA</b>	
	<b>JUDUL</b>	<b>i</b>
	<b>PENGAKUAN PENULIS</b>	<b>ii</b>
	<b>DEDIKASI</b>	<b>iii</b>
	<b>PENGHARGAAN</b>	<b>iv</b>
	<b>ABSTRAK</b>	<b>v</b>
	<b>ABSTRACT</b>	<b>vi</b>
	<b>KANDUNGAN</b>	<b>vii</b>
	<b>SENARAI JADUAL</b>	<b>xiii</b>
	<b>SENARAI RAJAH</b>	<b>xvi</b>
	<b>SENARAI SINGKATAN</b>	<b>xx</b>
	<b>SENARAI ISTILAH</b>	<b>xxi</b>
	<b>SENARAI SIMBOL</b>	<b>xxiii</b>
	<b>SENARAI LAMPIRAN</b>	<b>xxiv</b>
<b>1</b>	<b>Pengenalan</b>	<b>1</b>
	1.0 Pengenalan	1
	1.1 Latarbelakang Masalah	2
	1.2 Penyataan masalah	7
	1.3 Matlamat Utama	7
	1.4 Objektif	8

1.5	Skop Penyelidikan	8
1.6	Kepentingan Penyelidikan	9
1.7	Sumbangan Ilmiah	10
1.8	Struktur Tesis	11
<b>2</b>	<b>KAJIAN LITERASI</b>	<b>13</b>
2.0	Pengenalan	13
2.1	Kajian kes (Sistem Maklumat Hidrologi Malaysia (SMHM))	14
2.1.1	Data hidrologi SMHM	14
2.1.2	Pangkalan Data SMHM	15
2.1.3	Perbincangan	17
2.2	Pemprosesan Pertanyaan	19
2.2.1	Penguraian Pertanyaan	21
2.2.2	Pengoptimuman Pertanyaan	25
	2.2.2.1 Pendekatan Heuristik untuk Pengoptimuman Pertanyaan	27
	2.2.2.2 Strategi-strategi peraturan pemprosesan	30
2.3	Kaedah Gabungan	31
2.3.1	Operasi gabungan	32
2.3.2	Jenis Gabungan	33
	2.3.2.1 Gabungan Kesamaan	33
	2.3.2.2 Semi Gabungan	34
	2.3.2.3 Gabungan Luaran	35
	2.3.2.4 Perbincangan	35
2.3.3	Implementasi Gabungan	36
	2.3.3.1 Gabungan Gelung Bersarang	37
	2.3.3.2 Gabungan Gelung Susun-gabung	38
	2.3.3.3 Gabungan Cencangan	39
2.4	Perbandingan teknik-teknik Pengoptimuman	40
2.5	Kesimpulan	44

<b>3</b>	<b>METODOLOGI PENYELIDIKAN</b>	<b>46</b>
3.0	Pengenalan	46
3.1	Rangka Kerja Penyelidikan	48
3.2	Analisis Pangkalan Data Semasa (Kajian kes : SMHM)	49
3.3	Pengoptimuman Skema Logikal Pangkalan Data	
	Hubungan	51
3.3.1	Rekabentuk Umum Pengoptimuman Skema	
	Logikal Pangkalan Data Hubungan	52
3.3.2	Perlaksanaan Teknik Ke Atas Pangkalan	
	Data SMHM	54
3.4	Pemprosesan Pertanyaan	55
3.5	Kesimpulan	56
<b>4</b>	<b>TEKNIK PENGOPTIMUMAN SKEMA LOGIKAL</b>	
	<b>PANGKALAN DATA HUBUNGAN BERDASARKAN</b>	
	<b>GABUNGAN-KESAMAAN</b>	<b>57</b>
4.0	Pengenalan	57
4.1	Konsep Pengoptimuman Skema Logikal Pangkalan	
	Data Hubungan Berdasarkan Gabungan Kesamaan	58
4.1.1	Peraturan-peraturan Pengoptimuman Skema	
	Logikal Pangkalan Data Hubungan	60
	4.1.1.1 Menghapuskan atribut yang bertindan	60
	4.1.1.2 Menggabungkan dua entiti menjadi satu	62
	4.1.1.3 Membahagikan entiti kepada entiti lain	
	berdasarkan atribut yang sesuai	63
4.1.2	Algoritma Pengoptimuman Skema Logikal	
	Pangkalan Data Hubungan	65
4.2	Pemprosesan Pertanyaan	66
	4.2.1 Penguraian Pertanyaan	67
	4.2.2 Pengoptimuman Pertanyaan	68
	4.2.3 Penjanaan Kod	69
4.3	Perbincangan	69



<b>5</b>	<b>PERLAKSANAAN PENGOPTIMUMAN GABUNGAN-KESAMAAN</b>	<b>71</b>
5.0	Pengenalan	71
5.1	Implementasi Pengoptimuman Skema Logikal Pangkalan Data Hubungan	72
5.1.1	Gabungan Entiti	72
5.1.2	Klasifikasi Entiti	74
5.2	Implementasi Teknik Pengoptimuman Skema Logikal Pangkalan Data Hubungan Ke Atas SMHM	77
5.2.1	Data SMHM	77
5.2.2	Pangkalan Data SMHM	80
5.3	Pemprosesan Pertanyaan SMHM	86
5.3.1	Perlaksanaan Penguraian Pertanyaan	87
5.3.2	Perlaksanaan Pengoptimuman Pertanyaan	88
5.3.3	Perlaksanaan Penjanaan Kod	89
5.4	Kesimpulan	89
<b>6</b>	<b>PENGUJIAN PERBANDINGAN</b>	<b>91</b>
6.0	Pengenalan	91
6.1	Persekitaran Pengujian	92
6.2	Kriteria Perbandingan	93
6.3	Contoh Penyataan Pertanyaan	94
6.4	Pengujian Perbandingan Pengoptimuman Pertanyaan	96
6.4.1	Pemprosesan Pertanyaan Bagi Jumlah Taburan Hujan	98
6.4.2	Pemprosesan Pertanyaan Bagi Purata Taburan Hujan	103
6.5	Perbandingan Masa Perlaksanaan Pelan	108

6.6	Perbandingan Sebelum dan Selepas Pengoptimuman	111
6.5	Perbincangan	114
<b>7</b>	<b>PERBINCANGAN DAN KESIMPULAN</b>	<b>115</b>
7.0	Pengenalan	115
7.1	Hasil Penyelidikan	117
7.2	Kebaikan dan Kelemahan Teknik Yang Dihasilkan	119
7.3	Cadangan Pembaikan	121
7.4	Kesimpulan	121
	<b>RUJUKAN</b>	<b>123</b>
	Lampiran A	129

## SENARAI JADUAL

NO JADUAL	TAJUK	MUKA SURAT
2.1	Peraturan-peraturan Pengubahan (Chen, 2003)	24
2.2	Hubungan X	32
2.3	Hubungan Y	32
2.4	Hubungan $Z = X \times Y$	32
2.5	Hasil Gabungan Kesamaan	34
2.6	Hasil Gabungan Semi-Join	35
2.7	Perbandingan Teknik-teknik Pengoptimuman	43
3.1	Jadual SMHM	49
3.2	Hubungan DT_NAME	52
3.3	Hubungan DT_REAL	52
3.4	Hubungan DT_STATION (Hasil Gabungan DT_NAME dan DT_REAL)	53
4.1	Operator Perbandingan	62
5.1	Bilangan <i>Tuple</i> Bagi Setiap Entiti (Gabungan)	83

5.2	Bilangan Entiti Klasifikasi	83
5.3	Bilangan <i>Tuple</i> Bagi Setiap Entiti (Klasifikasi)	84
6.1	Hubungan SMHM sebelum dan selepas pengoptimuman (Jumlah kapasiti data adalah 100 MB)	93
6.2	Contoh penyataan pertanyaan bagi pangkalan data SMHM	94
6.3	Masa Perlaksanaan Bagi Pertanyaan P1	100
6.4	Masa Perlaksanaan Bagi Pertanyaan P2	101
6.5	Masa Perlaksanaan Bagi Pertanyaan P3	102
6.6	Masa Perlaksanaan Bagi Pertanyaan P4	103
6.7	Masa Perlaksanaan Bagi Pertanyaan P5	104
6.8	Masa Perlaksanaan Bagi Pertanyaan P6	105
6.9	Masa Perlaksanaan Bagi Pertanyaan P7	106
6.10	Masa Perlaksanaan Bagi Pertanyaan P8	107
6.11	Masa Perlaksanaan Pertanyaan P1, P2, P3 dan P4	111
6.12	Perbandingan Masa Perlaksanaan Bagi P1, P2, P3 dan P4	112

6.13	Masa Pelaksanaan Pertanyaan P5, P6, P7 dan P8	113
6.14	Perbandingan Masa Pelaksanaan Bagi P5, P6, P7 dan P8	113

**SENARAI RAJAH**

<b>NO RAJAH</b>	<b>TAJUK</b>	<b>MUKA SURAT</b>
1.1	Sistem Kiub	5
2.1	Data Hujan	15
2.2	Terjemahan Kiub Kepada Bentuk Jadual	17
2.3	Fasa Pemprosesan Pertanyaan (Connoly dan Begg, 1998)	20
2.4	Pepohon Pertanyaan	22
2.5	Algoritma Bagi Gabungan Gelung Bersarang (Priti dan Margaret, 1992)	37
2.6	Algoritma Bagi Gabungan Gelung Susun-Gabung (Priti dan Margaret, 1992)	38
2.7	Algoritma Bagi Gabungan Cencangan (Connoly dan Begg, 1998)	39
2.8	Gabungan PERF bagi $R_1$ dan $R_2$	41
3.1	Rangka Kerja Penyelidikan	47
3.2	Skema Pangkalan Data SMHM	50

3.3	Senibina Sistem	55
4.1	Rekabentuk Keseluruhan Pengoptimuman	59
4.2	Proses Gabungan Entiti	61
4.3	Klasifikasi Entiti a) Entiti Asal b) Entiti Sementara c) Entiti Baru	64
4.4	Algoritma Bagi Pengoptimuman Jadual	65
4.5	Struktur Asas Pepohon Pertanyaan a) Sebelum Pengoptimum b) Selepas Pengoptimuman	68
5.1	Antaramuka Bagi Input Entiti dan Atribut Gabungan	72
5.2	Aturcara Gabungan Entiti	73
5.3	Antaramuka Hasil Gabungan Entiti	73
5.4	Aturcara Mengira Bilangan Tuple Bagi Setiap Entiti	74
5.5	Antaramuka Input Proses Klasifikasi Entiti	75
5.6	Keratan Aturcara Pengelompokan	75
5.7	Keratan Aturcara Pengklasifikasian	76
5.8	Antaramuka Hasil Klasifikasi Entiti	76

5.9	Entiti DT_CHAR	78
5.10	Entiti DT_NAME	79
5.11	Entiti DT_REAL	80
5.12	Implementasi Pengoptimuman Skema Logikal Pangkalan Data Hubungan Ke Atas SMHM	81
5.13	Skema Logikal Pangkalan Data SMHM selepas pengoptimuman	85
5.14	Pepohon Pertanyaan a) Sebelum Pengoptimuman b) Selepas Pengoptimuman	88
5.15	Perlaksanaan Penyataan Pertanyaan	89
6.1	Pelan Pertanyaan	97
6.2	Graf Perbandingan Bagi Pertanyaan P1	100
6.3	Graf Perbandingan Bagi Pertanyaan P2	101
6.4	Graf Perbandingan Bagi Pertanyaan P3	102
6.5	Graf Perbandingan Bagi Pertanyaan P4	103
6.6	Graf Perbandingan Bagi Pertanyaan P5	105
6.7	Graf Perbandingan Bagi Pertanyaan P6	106



6.8	Graf Perbandingan Bagi Pertanyaan P7	107
6.9	Graf Perbandingan Bagi Pertanyaan P8	108
6.10	Graf Perbandingan Masa Pelaksanaan Bagi Penyataan Pertanyaan Jumlah Taburan Hujan	109
6.11	Graf Perbandingan Masa Pelaksanaan Bagi Penyataan Pertanyaan Purata Taburan Hujan	110
6.12	Graf Perbandingan P1, P2, P3, P4	112
6.13	Graf Perbandingan P5, P6, P7, P8	114
7.1	Klasifikasi bagi entiti DT_REAL	119

**SENARAI SINGKATAN**

DBMS	-	<i>Database Management System</i>
SMHM	-	Sistem Maklumat Hidrologi Malaysia
TI	-	<i>Time Index</i>
TID	-	<i>Time Identification</i>
DID	-	<i>Attribute Identification</i>
FID	-	<i>Feature Identification</i>
JPS	-	Jabatan Pengairan dan Saliran Malaysia
SPDM	-	Sistem Pangkalan Data Multimedia
TANGO	-	<i>Temporal Adaptive Next Generation Query Optimizer and Processor</i>
HIS	-	Sistem Maklumat Hidrologi
SQL	-	<i>Structured Query Language</i>
PERF	-	<i>Partially Encoded Record Filters</i>
TIDEDA	-	<i>Time Dependant Data</i>
ASP	-	<i>Active Server Pages</i>
IIS	-	<i>Internet Information Server</i>

**SENARAI ISTILAH**

Sistem Pengurusan Pangkalan Data	-	<i>Database Management System</i>
Capaian semula maklumat	-	<i>Information retrieval</i>
Pertanyaan	-	<i>Query</i>
Semasa	-	<i>Current</i>
Lepas	-	<i>Past</i>
Ruang	-	<i>Spatial</i>
Bukan ruang	-	<i>Non-spatial</i>
Penyinaran	-	<i>Radiation</i>
Bersandarkan masa	-	<i>Temporal</i>
Penolakan	-	<i>Pushing</i>
File Teks	-	<i>Text file</i>
Titik	-	<i>Point</i>
Selang	-	<i>Interval</i>
Gabungan kesamaan	-	<i>Equijoin</i>
Hubungan	-	<i>Relation</i>
Jadual	-	<i>Table</i>
Bahasa Pertanyaan Berstruktur	-	<i>Structured Query Language</i>
Pangkalan pengetahuan	-	<i>Knowledge Base</i>
Penyusunan semula nod dalaman	-	<i>Internal node reordering</i>
Penyusunan semula nod daun	-	<i>Leaf node reordering</i>
Binaan dalaman	-	<i>Built in</i>
Pengumpul Statistik	-	<i>Statistic Collector</i>
Penganggar Kos	-	<i>Cost Estimator</i>
Gabungan	-	<i>Join</i>
Semi gabungan	-	<i>Semi join</i>

Pemilihan	-	<i>Select</i>
Kemasukan	-	<i>Insert</i>
Pengemasikinian	-	<i>Update</i>
Penyataan SQL	-	<i>SQL Statement</i>
Peringkat tinggi	-	<i>High level</i>
Peringkat rendah	-	<i>Low level</i>
Selari	-	<i>Parallel</i>
Bentuk normal penghubung	-	<i>Conjunctive normal form</i>
Bentuk normal pemisah	-	<i>Disjunctive normal form</i>
Pemudahan	-	<i>Simplification</i>
Pengunjuran	-	<i>Projection</i>
Dua cara	-	<i>Two ways</i>
Multi cara	-	<i>Multi ways</i>
Hasil darab Karteisan	-	<i>Cartesian Product</i>
Kesamaan	-	<i>Equality</i>
Gabungan gelung bersarang	-	<i>Nested loop join</i>
Gabungan gelung susun gabung	-	<i>Sort merge join</i>
Gabungan cencangan	-	<i>Hash join</i>
Pengurang	-	<i>Reducer</i>
Hubungan Dalaman	-	<i>Inner relation</i>
Hubungan Luaran	-	<i>Outer relation</i>
Gabungan biasa	-	<i>Natural Join</i>
Gabungan luaran	-	<i>Outer join</i>
Semigabungan dua cara	-	<i>Two ways semi join</i>
Gabungan unjuran	-	<i>project join</i>
Kod penerang	-	<i>qualifier code</i>
Kod kaedah	-	<i>method code</i>
Status pengesahan	-	<i>validation status code</i>
Pemilihan	-	<i>selection</i>
Penapisan	-	<i>filtering</i>

**SENARAI SIMBOL**

$\sigma$	-	Pemilihan
$\bowtie$	-	Gabungan
$\wedge$	-	Operator AND
$\times$	-	Hasil Darab Karteisan
$\amalg$	-	Hasil Tambah
$\cup$	-	Kesatuan
$\in$	-	Subset
$\cap$	-	Persilangan

**SENARAI LAMPIRAN**

<b>NO LAMPIRAN</b>	<b>TAJUK</b>	<b>MUKA SURAT</b>
A	Contoh laporan SMHM	129

## **BAB 1**

### **PENGENALAN**

#### **1.0 Pengenalan**

Kini, kaedah yang berkesan dalam capaian semula maklumat dan analisis menjadi semakin penting. Ini sejajar dengan perkembangan sumber dan pengguna maklumat yang semakin meningkat. Dengan perkembangan yang semakin pesat, keperluan untuk mendapatkan maklum balas dalam sebarang urusan dengan cepat dan pantas menjadi keutamaan bagi pengguna. Ia mendorong para penyelidik untuk meningkatkan kaedah yang dapat mempercepatkan masa tindak balas pertanyaan dalam pangkalan data. Sehubungan itu, tesis ini adalah kajian terhadap proses capaian data dalam pangkalan data.

Pangkalan data adalah merujuk kepada himpunan atau koleksi data yang berkaitan dihubungkan secara logikal. Ia diwujudkan untuk memenuhi keperluan sesuatu organisasi. Manakala sistem pengurusan pangkalan data adalah satu sistem yang membolehkan pengguna untuk menaakrif, membina dan menyenggara pangkalan data serta menyediakan kawalan capaian terhadap pangkalan data tersebut (Connolly dan Begg, 1998). Bagi sesetengah aplikasi, ia memerlukan pengurusan terhadap data yang dipengaruhi oleh unsur masa. Data yang disimpan dalam pangkalan data bukan hanya data semasa, ia termasuklah juga data lepas.

Sistem Maklumat Hidrologi Malaysia (SMHM) (SMHB ITP, 1999) merupakan salah satu contoh aplikasi yang menyimpan data lepas. Data tersebut akan digunakan untuk tujuan analisis dan ramalan. SMHM adalah satu sistem maklumat yang menyimpan, memproses, menganalisis dan memaparkan data hidrologi untuk tujuan kajian penilaian sumber air dan maklumat hidrologi. Ia juga digunakan bagi tujuan pembangunan dan pengurusan sumber air serta operasi kawalan semasa banjir atau kemarau. Masalah yang sering timbul dalam pengurusan data tersebut adalah masa proses capaian yang panjang dalam menganalisis dan memaparkan semula data. Faktor-faktor yang dikenal pasti menyebabkan masalah berlaku adalah struktur data dalam pangkalan data yang kurang berkesan dari segi proses masukan dan capaian data. Tambahan lagi, jumlah rekod dalam pangkalan data juga adalah terlalu besar dan banyak. Bagi meningkatkan masa pelaksanaan proses capaian data, satu teknik yang efektif adalah perlu untuk menyelesaikan masalah tersebut.

## **1.9 Latar Belakang Masalah**

Pangkalan data merupakan komponen asas dalam satu sistem maklumat yang digunakan untuk menyimpan data yang berkaitan. Antara operasi utama pangkalan data termasuklah menyenggara dan capaian data dalam pangkalan data. Sistem maklumat hidrologi merupakan salah satu contoh sistem maklumat yang menyimpan dan menguruskan data yang dipengaruhi unsur masa di dalam pangkalan data.

Terdapat dua jenis data yang disimpan dalam pangkalan data hidrologi iaitu data ruang dan data bukan ruang (John, 2000). Data ruang ialah data yang mewakili ruang dan lokasi bagi data geografi seterusnya dipersembahkan dalam bentuk grafik. Manakala maklumat tambahan yang menerangkan data ruang dari sudut sifat-sifat geografi, fenomena alam dan juga aktiviti manusia disebut data bukan ruang. Sebagai contoh, ia terdiri daripada nilai fenomena (hujan, kedalaman air, kualiti, kadar penyejatan dan pemendapan, dan sebagainya), sifat dan ciri-ciri geografi (guna



tanah, kawasan tinggi), nama kawasan, tarikh dan masa dan maklumat-maklumat lain yang menyokong data ruang. Secara asasnya, data ini dipengaruhi oleh unsur masa yang merupakan faktor utama membawa kepada perubahan data (Chen, 2002). Oleh itu, setiap perubahan yang berlaku perlu direkodkan dan disimpan dalam pangkalan data. Sebagai contoh, data taburan hujan iaitu jumlah taburan hujan akan direkodkan bagi setiap jam setiap hari dan begitulah seterusnya. Dengan itu, data yang disimpan dalam pangkalan data adalah menjangkau puluhan tahun. Ini bermaksud, data direkodkan secara berterusan dan menjadikan bilangan rekod di dalam pangkalan data sentiasa bertambah. Data tersebut akan diproses dan dianalisa, seterusnya dipaparkan dalam bentuk laporan untuk memudahkan fungsi-fungsi tertentu dilaksanakan.

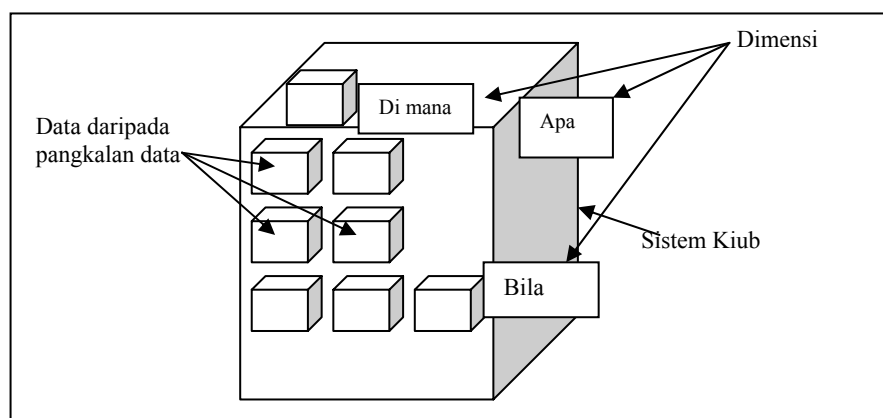
Berdasarkan beberapa kajian yang telah dijalankan, masalah utama yang sering dihadapi dalam menguruskan data yang bersandarkan kepada masa adalah kesukaran dalam mengendalikan data yang kompleks (Lorentzos, 1994; Chen, 2002; Min *et al*, 2001; Kafer dan Schoning, 1992; Mesru *et al*, 2004). Selain itu, masa pelaksanaan dalam capaian data juga menjadi isu utama para penyelidik (Kouramajin, 1994; Lubna, 2002; Tao, 2002; Chen, 2003; SMHB. ITP, 1999). Perbincangan seterusnya hanya akan menerangkan beberapa penyelidikan lepas yang berkaitan dengan fokus utama penyelidikan ini.

Dalam penyelidikan yang dijalankan oleh Chen (2002), beliau telah membangunkan satu sistem maklumat hidrologi masa nyata secara atas talian. Menurut beliau, proses mengumpul dan menganalisa maklumat hidrologi memerlukan masa yang terlalu lama. Oleh itu, sistem tersebut dibangunkan untuk menyelesaikan masalah itu. Antara data hidrologi yang disimpan adalah data taburan hujan, data aliran sungai, data penyinaran dan data salji. Data yang dikumpul dan direkodkan merangkumi kawasan tadahan seluas 695km<sup>2</sup>. Komunikasi radio dan satelit digunakan untuk mengumpul data taburan hujan dan salji. Imej satelit adalah antara sumber data kompleks yang perlu untuk disimpan dan diuruskan dalam pangkalan data. Pembangunan sistem ini hanya memfokuskan kepada pengendalian data yang kompleks. Ia tidak mengambil kira pengurusan dan manipulasi data bagi bilangan rekod yang terlalu banyak. Oleh itu, masalah yang dibincangkan masih tidak dapat diatasi sepenuhnya.

Kouramajin (1994) menerangkan satu teknik pengindeksan data berdasarkan masa iaitu Time Index (TI) sebagai salah satu teknik yang dapat memperbaiki masa pelaksanaan pertanyaan. Teknik ini digunakan untuk mencapai satu versi objek dalam satu titik masa tertentu atau dalam satu selang masa tertentu. Data terdiri dari satu kumpulan versi objek yang didefinisikan menggunakan pangkalan data hubungan berdasarkan masa. Terdapat pelbagai operator carian yang dapat disokong oleh TI, antara operator asas adalah berasaskan titik dan berasaskan selang. Proses capaian bagi jumlah data yang kecil dapat dilaksanakan dengan efisien menggunakan operator-operator tersebut. Walau bagaimanapun, masalah timbul bagi capaian data dalam satu jumlah rekod yang besar dan banyak iaitu memerlukan masa yang lama dalam proses tersebut.

Selain itu, teknik pengoptimuman jadual juga merupakan salah satu teknik yang dapat memperbaiki masalah yang dikenal pasti. Walau bagaimanapun, ia masih lagi mendapat perhatian para penyelidik walaupun terdapat pelbagai teknik pengoptimuman telah dicadangkan. Ini kerana masih timbul masalah dalam pengendalian data dalam pangkalan data. Sebagai contoh, Chen (2003) mencadangkan satu teknik pengoptimuman menggunakan kaedah berasaskan histogram. Beliau menggunakan kaedah gabungan-kesamaan untuk menggabungkan dua hubungan yang mempunyai atribut yang sama dan seterusnya mengimplementasikan algoritma berasaskan histogram iaitu *Averaged Rectangular Attribute Cardinality (R-ACM)* dan *Trapezoidal Attribute Cardinality (T-ACM)* untuk memperbaiki proses pelaksanaan pertanyaan.

SMHB (1999) memperkenalkan Sistem Kiub yang menerangkan bagaimana data hidrologi disimpan dan dicapai dalam pangkalan data hubungan. Berdasarkan penerangan sebelum ini, setiap data hidrologi akan menggabungkan tiga asas utama iaitu data bukan ruang, data ruang, dan masa. Oleh itu, Sistem Kiub diperkenalkan berdasarkan gabungan 3 asas tersebut. Rajah 1.1 menunjukkan konsep sistem kiub yang merujuk kepada data hidrologi yang hendak dicapai mesti mempunyai 3 ciri penting iaitu Di mana?, Apa?, dan Bila?



**Rajah 1.1 : Sistem Kiub**

Berikut adalah penerangan terperinci bagi 3 ciri penting tersebut :

- i) FID (*Feature Identification*) – kekunci ini merupakan kekunci utama yang menjadi rujukan bagi setiap ciri geografi yang terdapat di atas permukaan bumi. Sebagai contoh, data yang akan direkodkan adalah data taburan hujan .
- ii) TID (*Time Identification*) – kekunci ini digunakan untuk rujukan dari masa ke masa apabila data direkodkan. Kegunaannya adalah lebih kepada untuk menguruskan data yang berubah. Data akan direkod berdasarkan peristiwa yang berlaku. Sebagai contoh, data taburan hujan akan direkodkan berdasarkan sela masa tertentu.
- iii) DID (*Attribute Identification*) – Kekunci ini merujuk kepada keterangan bagi data yang diambil. Sebagai contoh, jika data tersebut adalah data hujan, ia akan merujuk kepada nama stesen dan maklumat terperinci tentang stesen tersebut.

Suatu capaian yang hendak dilaksanakan hanya akan merujuk kepada 3 ciri penting tersebut. Hal ini dapat mengurangkan bebanan semasa carian. Ia hanya merujuk tiga perkara berbanding banyak hubungan yang wujud dalam pangkalan data. Walau bagaimanapun, masalah ini masih lagi dibincangkan dalam capaian data dari satu rekod yang besar dan banyak. Masalah ini tidak dapat diatasi sepenuhnya menggunakan sistem kiub. Ia juga turut menjadi isu besar di dalam menyimpan data hidrologi kerana data atau rekod yang disimpan adalah menjangkau puluhan tahun.

Di dalam tesis ini, SMHM dijadikan sebagai kajian kes penyelidikan. SMHM adalah satu sistem maklumat yang menyimpan data hidrologi termasuk data hujan, data paras air, data penyejatan, data pemendapan dan data kualiti air. Ia telah dibangunkan oleh Institut Teknologi Perisian, Universiti Teknologi Malaysia dengan kerjasama Syed Muhammad Sdn. Bhd. (SMHB ITP, 1999). Seperti yang diterangkan, data yang disimpan akan digunakan untuk tujuan kajian penilaian sumber air dan maklumat hidrologi. Selain itu, ia juga digunakan untuk tujuan pembangunan dan pengurusan. Untuk memudahkan pihak yang berkenaan melaksanakan tugas tersebut, data yang disimpan akan dianalisis dan dipaparkan dalam bentuk laporan berdasarkan keperluan yang dikehendaki. Berdasarkan kajian terhadap keperluan pengguna SMHM, masalah capaian data yang lama adalah dalam penghasilan laporan tersebut.

Berikut adalah antara laporan-laporan yang terlibat secara langsung dalam pembangunan projek SMHM (Moore, 1996):

- i) Laporan harian bagi taburan hujan dengan jumlah, nilai maksimum dan nilai minimum.
- ii) Laporan harian bagi taburan hujan dengan pilihan untuk mendapatkan nilai maksimum bagi satu jangka masa dari 15 minit hingga 30 hari.
- iii) Laporan bulanan dan tahunan bagi taburan hujan dengan jumlah, nilai maksimum dan nilai minimum.
- iv) Laporan taburan hujan bagi 5 tahun dengan jumlah, nilai maksimum dan nilai minimum.
- v) Laporan bagi stesen yang menerima jumlah taburan hujan paling minimum atau maksimum pada satu masa tertentu.

Rujuk Lampiran A untuk contoh laporan yang dihasilkan. Berdasarkan kajian yang dijalankan terhadap pangkalan data SMHM, faktor-faktor yang dikenal pasti mempengaruhi masa perlaksanaan capaian data dalam pangkalan data adalah struktur jadual dan hubungan dalam pangkalan data. Tambahan lagi, jumlah rekod di dalam pangkalan data adalah terlalu banyak dan besar. Oleh itu, penulis merumuskan bahawa perlu satu teknik yang dapat menyelenggara dan menguruskan

data tersebut dengan baik dan berkesan supaya dapat menyelesaikan masalah yang telah dibincangkan.

### **1.10 Penyataan Masalah**

Berdasarkan kepada masalah yang dihadapi oleh SMHM, perlu satu teknik yang dapat mempercepatkan masa pelaksanaan pertanyaan dalam capaian data di dalam pangkalan data. Berikut adalah beberapa persoalan yang perlu diuraikan :

- Bagaimanakah cara untuk mengoptimumkan masa tindak balas suatu pertanyaan.
- Bagaimanakah penganalisaan terhadap suatu pertanyaan dijalankan untuk menentukan cara pelaksanaan yang paling berkesan.

### **1.11 Matlamat Utama**

Matlamat utama penyelidikan ini adalah untuk menghasilkan satu teknik yang dapat mempercepatkan masa pelaksanaan pertanyaan dalam capaian data di dalam pangkalan data hidrologi.

## 1.12 Objektif

Objektif penyelidikan ini adalah seperti berikut :

- i. Menghasilkan satu teknik pengoptimuman skema logikal pangkalan data hubungan yang dapat meminimakan masa pelaksanaan pertanyaan dalam proses capaian data di dalam pangkalan data.
- ii. Menghasilkan dan menganalisis penyataan pertanyaan dan cara pelaksanaan berdasarkan teknik pengoptimuman yang telah dihasilkan.
- iii. Melaksanakan pengujian keberkesanan terhadap teknik yang telah dihasilkan iaitu teknik pengoptimuman jadual menggunakan data hidrologi SMHM.

## 1.13 Skop Penyelidikan

Skop penyelidikan adalah tertumpu kepada :

- i. Teknik yang dihasilkan melibatkan penstrukturan semula jadual dan hubungan antara jadual dalam pangkalan data. Seterusnya, pengoptimuman pertanyaan dilaksanakan semasa pelaksanaan capaian data.
- ii. Berdasarkan kajian terhadap keperluan pengguna SMHM, pertanyaan untuk penghasilan laporan memerlukan masa yang panjang. Oleh itu, di dalam penyelidikan ini, hanya melibatkan pertanyaan untuk

penghasilan laporan tersebut. Beberapa laporan yang digunakan untuk tujuan analisis dan perbandingan adalah laporan harian, bulanan dan tahunan bagi data taburan hujan.

- iii. Penyelidikan ini hanya mengambilkira masa pelaksanaan terhadap capaian data, manakala ruang storan yang digunakan untuk menyimpan data tersebut tidak diambilkira.
- iv. Teknik pengoptimuman skema logikal pangkalan data hubungan yang dihasilkan adalah menggunakan model pangkalan data hubungan.
- v. Dalam penyelidikan ini, data hidrologi digunakan untuk tujuan analisis. Data yang disimpan adalah data mentah yang diperolehi dari pihak JPS. Ia termasuk data taburan hujan, data kualiti air, data paras air, dan data penyejatan yang merangkumi data untuk 15 tahun (1988-2002).
- vi. Pengujian dilakukan untuk melihat keberkesanan teknik yang dicadangkan berbanding sistem SMHM berdasarkan masa tindak balas pelaksanaan suatu pertanyaan.

#### **1.14 Kepentingan Penyelidikan**

Penyelidikan yang dijalankan adalah berdasarkan masalah yang timbul daripada isu capaian data yang perlahan. Penyelidikan ini memfokuskan kepada pengurusan data hidrologi. Data hidrologi sentiasa berubah mengikut masa dan setiap perubahan adalah penting untuk direkodkan. Ini kerana data diperlukan untuk proses analisis atau ramalan pada bila-bila masa. Disebabkan faktor tersebut jumlah data dalam pangkalan data sentiasa bertambah. Masalah yang timbul adalah proses itu memerlukan masa pelaksanaan yang lama dalam proses capaian data. Oleh itu,

penyelidikan dijalankan untuk menghasilkan satu teknik yang dapat memperbaiki masa capaian data dalam pelaksanaan suatu pernyataan pertanyaan. Dalam perkembangan penyelidikan pada masa ini, keperluan bagi mengatasi masalah tersebut adalah penting. Tambahan pula, ia sangat berguna dalam meningkatkan keberkesanan proses capaian bagi memenuhi keperluan dan kehendak pengguna. Selain itu, teknik yang dihasilkan juga dapat membantu pihak JPS untuk meningkatkan lagi keberkesanan SMHM dalam menguruskan data hidrologi. Di samping itu, ia juga dapat memudahkan pihak JPS dalam melaksanakan tugas seterusnya dapat meningkatkan produktiviti kerja di kalangan kakitangan yang terlibat.

### **1.15 Sumbangan Ilmiah**

Hasil utama penyelidikan adalah satu teknik pengoptimuman skema logikal pangkalan data hubungan yang dihasilkan menggunakan kaedah gabungan-kesamaan. Idea ini adalah berdasarkan isu penyelidikan yang memerlukan satu teknik yang dapat memperbaiki masa pelaksanaan pertanyaan dalam proses capaian data dalam pangkalan data. Teknik ini dapat memperbaiki hubungan antara jadual (entiti) di dalam pangkalan data seterusnya dapat mengatasi masalah yang telah dibincangkan dalam penyelidikan ini. Kaedah yang dihasilkan membawa kepada penyelesaian masalah SMHM iaitu masa capaian data yang lama. Ini dapat dibuktikan dengan pengujian perbandingan yang dilaksanakan antara sistem kajian kes iaitu SMHM dan teknik yang dicadangkan iaitu Pengoptimuman Skema Logikal Pangkalan Data Hubungan berdasarkan kaedah gabungan kesamaan.

Bagi suatu pernyataan pertanyaan, terdapat beberapa pelaksanaan pertanyaan akan dihasilkan berdasarkan teknik pengoptimuman jadual yang telah dihasilkan di atas. Pertanyaan tersebut akan dianalisis menggunakan pengoptimum pertanyaan untuk memilih dan menentukan pertanyaan yang dapat meningkatkan keberkesanan pelaksanaan pertanyaan.



## 1.16 Struktur Tesis

Tesis ini membincangkan penyelidikan yang telah dijalankan dan mengandungi tujuh (7) bab.

Bab 1 adalah pengenalan kepada penyelidikan yang telah dijalankan. Ia merangkumi perbincangan mengenai latar belakang masalah, pernyataan masalah dan matlamat serta objektif penyelidikan. Perbincangan berkenaan sumbangan ilmiah turut dimuatkan dalam bab ini.

Bab 2 adalah perbincangan mengenai kajian literasi yang telah dijalankan. Ia merangkumi kajian terhadap pangkalan data semasa SMHM dan perbincangan berkenaan teknik-teknik berkaitan yang telah dibangunkan. Di samping itu, terdapat juga perbandingan dan rumusan yang dibuat dalam menentukan kaedah atau teknik yang sesuai untuk menyelesaikan masalah yang telah dibincangkan.

Bab 3 menerangkan mengenai metodologi penyelidikan. Dalam bab ini, penulis membincangkan kaedah-kaedah yang terlibat dalam penyelidikan ini. Selain itu, penerangan secara ringkas berkaitan teknik yang dihasilkan juga turut dibincangkan.

Bab 4 menerangkan secara terperinci teknik yang telah dihasilkan iaitu teknik Pengoptimuman Jadual berdasarkan kaedah Gabungan-Kesamaan.

Bab 5 membincangkan proses implementasi bagi teknik yang dihasilkan. Ia diimplementasikan menggunakan pangkalan data hubungan SMHM.

Bab 6 membincangkan berkaitan pengujian yang dilaksanakan dalam penyelidikan ini. Terdapat beberapa pertanyaan yang telah dihasilkan dan proses ini juga telah dibincangkan dalam bahagian ini. Selain itu, graf perbandingan antara teknik yang dibandingkan turut dimuatkan dan dibincangkan.

Bab 7 adalah perbincangan berkenaan hasil-hasil pengujian dan perbandingan antara sistem kajian kes iaitu SMHM dan teknik yang dicadangkan. Seterusnya penulis membincangkan hasil yang diperolehi daripada penyelidikan, kerja-kerja peningkatan untuk masa hadapan dan kesimpulan.

**RUJUKAN**

- Ab. Razak C H (2000). *Pangkalan Data*. Fakulti Sains Komputer dan Sistem Maklumat : Universiti Teknologi Malaysia.
- Abdullah M.Z (2004), *Discrete Mathematics*, Open University Malaysia.
- Alan R.H, Wu O Q and Yao S.B (1985), Query Optimization on Local Area Networks. *ACM Transaction on Office Information*, 3(1): 35-62.
- Bennet K, Ferris M C dan Ioannidis Y E (1991), A Genetic Algorithm for Database Query Optimization. In Proceedings of the Fourth International Conference on Genetic Algorithm pp 400-407.
- Chaudhuri S (1988) . An Overview of Query Optimization in Relational Systems. *Proceedings of the seventeenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems*. Seattle, Washington, US : 34-43.
- Chaudari S and Shim K (1995). An Overview of Cost-based Optimization of Queries with Aggregates. *IEEE Data Engineering Bulletin*. 18(3): 3-9.
- Chen D (2002). Real-time online hydrological information and modelling system using object-oriented approach and relational database for flood defence. *Flood Defence '2002*. New York: Science Press.
- Chen J, B.CS (2003). *On Utilizing New Histogram-Based Methods for Query Optimization*. Carleton University : Master Thesis.
- Christian M (2002). A Survey of Database Query Optimization and Genetic Algorithms, *ACM Transactions on Database Systems*, 4 (2) : 133-155.

- Connolly T and Begg C (1998). *Database Systems : A Practical Approach to Design, Implementation, and Management*. 2<sup>nd</sup> Ed. UK. Addison Wesley.
- Dean T (1989). Using Temporal Hierarchies to Efficiently Maintain Large Temporal Databases. *Journal of the Association for Computing Machinery*. 36(4) : 687-718.
- Dumas M, Fauvet M dan Scholl P (2004). TEMPOS : A Platform for Developing Temporal Applications on Top of Object DBMS. *IEEE Transactions on Knowledge and Data Engineering*. 16(3): 354-374.
- Fegaras L (1998). A New Heuristic for Optimizing Large Queries. *Proceedings of the 9<sup>th</sup> International Conference on Database and Expert System Applications*. London, UK: Springer-Verlag, 726-735.
- Finkelstein S, Schkolnick M dan Tiberio P (1988). Physical Database Design for Relational Databases. *ACM Transactions on Database Systems*. 13(1) : 91-128.
- Godfrey P and Gryz J (2001). Partial Evaluation of Views. *Journal of Intelligent Information Systems*. 16(1): 21-39.
- Graefe G (1993). Query Evaluation Techniques for Large Databases. *ACM Computing Surveys*. 25(2): 73-169.
- Haraty R and Fany R (2000). Distributed Query Optimization Using PERF Join. *Proceedings of the 2000 ACM Symposium on Applied Computing*. March 19-21. Como, Italy: ACM. 2000: 284-288.
- Jarke. M and Koch J (1984). Query Optimization in Database Systems. *ACM Computing Surveys*. 16(2): 111-152.

- Jensen, C. S (2000). Introduction to temporal database. *IEEE Trans. Knowl. Data Eng.* 11(1): 36-44
- John L (2000). Using Spatial Analysis for Non-spatial Data. *Proceedings, ESRI International Users' Conference*. July 2000. San Diego, CA.
- Kafer W and Schoning H (1992). Realizing a Temporal Complex-Object Data Model. *Proc. SIGMOD 1992*. New York, USA : ACM Press, 266-275.
- Kouramajin V(1994). *Temporal database: Access Structures, Search Methods, Migration Strategies and Declustering Techniques*. The University of Texas at Arlington : PhD Dissertation.
- Kouramajin V, Kamel I, Elmasri R dan Waheed S (1994). The Time Index+: An Incremental Access Structure for Temporal Databases. *Proceedings of the Third International Conference on Information and Knowledge Management (CIKM'94)*. Nov 29 - Dec 2. Gaithersburg, Maryland : ACM, 296-303.
- Lorentzos, N.A (1994). DBMS support for nonmetric measurement systems. *Knowledge and Data Engineering, IEEE Transactions*. 6(6) :945-953
- Lubna S (2002). *Dynamic Technique in Distributed Query Optimization*. University of Windsor: Master Thesis.
- Mackret L.F and Lohman G.M. (1986). R\* Optimizer Validation and Performance Evaluation for Local Queries. *Proceedings of the Twelfth International Conference on Very Large Databases*. San Francisco : Morgan Kaufman Publishers Inc., 149-159.
- McMahan B J (2004). *Structral Heuristics For Query Optimization*. Rice University, MSc Thesis.

- Mesru K, Kesim N, Cicekli dan Yazici A (2004). Spatio-Temporal Querying in Video Databases. *Information Sciences-Informatic and Computer Science:An International Journal*. 160(1-4) : 131-152.
- Min J.S, Kim D H dan Ryu K H (2001). A spatiotemporal data and indexing. *Proceedings of IEEE Region 10 International Conference*. Aug 19-22. TENCON, 110-113.
- Mohd Rahim S (2002). *Spatial and Non Spatial Enhancement Database for Hydrological Information System (HIS)*. Universiti Teknologi Malaysia : Master Thesis.
- Moore R.V (1996), MHIS Development – Reports High Level Design.
- Navin Kabra (1999). *Query Optimization for Object Relational Database System*. University of Wisconsin Madison : Ph. D Thesis.
- ORACLE Company (2002). Query Optimization in Oracle 9i. *An Oracle White Paper*.
- Priti M and Margaret H.E (1992). Join Processing in Relational Databases. *ACM Computing Surveys*, 24(1):63-113
- Richard R. M and Cliff L (1990). Query Processing for Temporal Databases. *Proceedings of the Sixth International Conference on Data Engineering*. Washington, DC : IEEE Computer Society, 200-208.
- Riise A (1999). The Storage, management and use of spatio-temporal data in the DYNAMAP Project. *DYNAMAP White Paper*.

Selinger et al (1979). Access Path Selection in a Relational Database Management System. *Proceedings of the 1979 ACM SIGMOD International Conference on The Management of Data*. Boston, Mass : 23-24.

Shamkant B.N and Rafi Ahmed. Temporal Extensions to the Relational Model and SQL. In : Tansel et al .*Temporal Databases: Theory, Design and Implementation, eds*. California : The Benjamin/Cummings Publishing Company, Inc.. 92-109.

Slivinkas G, Jensen C. S dan Snodgrass R. T (2001). Adaptable Query Optimization and Evaluation in Temporal Middleware. *International Conference on Management of Data*. May 21-24: ACM SIGMOD, 127-138.

SMHB, ITP (1999), "Establishment Of Hydrological Information System For Water Resources Planning, Development and Management, Jabatan Pengairan Dan Saliran Malaysia, Kuala Lumpur, Technical Report, Volume 1,2,3,4.

Stillger M and Freytag J.C (1985). Testing the Quality of a Query Optimizer. *IEEE Data Engineering Bulletin*. 18(3): 41-48.

Sybase Inc. (1995). SQL Anywhere 5.x Cost-Based Query Optimizer. *Sybase, Workplace Database Division Sybase, Inc*. Volume I and II.

Swami A (1989). Optimization of Large Join Queries : Combining Heuristics and Combinatorial Techniques. *Proceedings of the 1989 ACM SIGMOD International Conference on Management of Data*. Portland, Oregon, United States : ACM SIGMOD, 367-376.

Tao Y (2002). *Indexing and Query Processing of Spatio-Temporal Data*. The Hong Kong University of Science and Technology. PhD Dissertation

- Thiyagarajah M (1999). *Attribute Cardinality Maps: New Query Result Size Estimation Techniques for Database Systems*. Carleton University, Ottawa, Canada: Ph. D Thesis.
- Timos K. S (1988). Multiple Query Optimization. *ACM Transaction on Database Systems*. 13 (1): 23-52.
- Unel G et al (2004). An efficient Query Optimization Strategy for Spatio-Temporal Queries in Video Database. *Journal of Systems and Software*. 73(1) : 113-131
- Valdueriz P and Gardarin G (1984). Join and Semi Join Algorithm for a Multi-Processor Database Machine. In *ACM Trans. Database System*, 9(1): 133-161
- Yannis E.I (1996). Query Optimization. *ACM Computing Surveys, Symposium Issue on The 50<sup>th</sup> Anniversary of ACM*, 28(1):21-123.