



UNIVERSITAS INDONESIA

**PERANCANGAN *DATA WAREHOUSE*
DAN
PENERAPAN *DATA MINING* PADA *SISTEM MONITORING*
JARINGAN GSM
STUDI KASUS:
PT. INDOSAT TBK**

SYAHREZA ZAIN

0606155562

**PROGRAM STUDI MAGISTER TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS INDONESIA
2008**



UNIVERSITAS INDONESIA

**PERANCANGAN *DATA WAREHOUSE*
DAN
PENERAPAN *DATA MINING* PADA SISTEM *MONITORING*
JARINGAN GSM
STUDI KASUS:
PT. INDOSAT TBK**

**Tesis diajukan sebagai salah satu syarat
untuk memperoleh gelar Magister Teknologi Informasi**

**Oleh:
SYAHREZA ZAIN**

0606155562

**PROGRAM STUDI MAGISTER TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS INDONESIA
2008**

LEMBAR PERSETUJUAN

JUDUL TESIS: PERANCANGAN DATA WAREHOUSE DAN PENERAPAN
DATA MINING PADA SISTEM MONITORING JARINGAN
GSM – STUDI KASUS: P.T. INDOSAT TBK

NAMA : SYAHREZA ZAIN

NPM : 0606155562

TESIS INI TELAH DIPERIKSA DAN DISETUJUI

JAKARTA,

YUDHO GIRI SUCAHYO, Ph.D, CISA
PEMBIMBING

YOVA RULDEVIYANI, M.Kom
PEMBIMBING

KATA PENGANTAR

Segala puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa, karena hanya dengan berkat dan kasih karunia-Nya penulis dapat menyelesaikan Tesis yang berjudul “*PERANCANGAN DATA WAREHOUSE DAN PENERAPAN DATA MINING PADA SISTEM MONITORING JARINGAN GSM – STUDI KASUS: P.T. INDOSAT TBK*” ini sesuai dengan yang direncanakan. Tesis ini dibuat guna melengkapi persyaratan kelulusan pada Program Studi Magister Teknologi Informasi, Fakultas Ilmu Komputer, Universitas Indonesia.

Penyusunan Tesis ini tidak akan berhasil tanpa bantuan dan dukungan dari berbagai pihak, baik secara langsung maupun tidak langsung. Oleh karena itu pada kesempatan ini penulis ingin menyampaikan rasa terima kasih secara khusus kepada:

1. Orang tua, kak maya dan soraya yang telah memberikan dukungan dan bantuan dalam setiap tindakan yang penulis lakukan.
2. Ibu Yova Ruldeviyani, M.Kom sebagai Dosen Pembimbing yang telah membimbing penulis selama mengerjakan penelitian Tesis ini.
3. Bapak Yudho Giri Sucahyo, Ph.D, CISA sebagai Dosen Pembimbing Akademis dan Ketua Program Studi Magister Teknologi Informasi, yang telah memberikan banyak nasehat dan petunjuk kepada penulis mulai dari proses pengajuan proposal sampai dengan penyelesaian Tesis ini.
4. Bapak Zainal A. Hasibuan, Ph.D selaku wakil dekan bidang akademik yang telah memberikan banyak nasehat dan petunjuk mengenai metode dan cara melakukan penelitian yang baik.
5. Segenap staf PT. Indosat Tbk, yang telah membantu penulis dalam menyediakan data dan informasi yang dibutuhkan dalam pembuatan tesis ini dan memberikan dukungannya, terutama mba Cut, mas Gofur,

mas Adhit (*metrica team*) dan mba Vira yang banyak membantu dalam pengerjaan tesis ini.

6. Teman-teman MTI UI terutama *Akatsuki Team* (ryan, syahrul, lutfi, fajri) angkatan 2007B, yang telah memberikan dukungan kepada penulis dalam menyelesaikan Tesis ini.
7. Saudara-saudara serta teman-teman penulis yang telah banyak memberikan dukungan moral dan material selama proses penyusunan Tesis ini.
8. Serta segenap pihak lain yang tidak dapat disebutkan satu-persatu disini, yang telah berkontribusi baik secara langsung maupun tidak langsung dalam penyusunan Tesis ini.

Tesis ini dibuat guna melengkapi persyaratan untuk memperoleh gelar Magister Ilmu Komputer pada Program Magister Teknologi Informasi, Fakultas Ilmu Komputer, Universitas Indonesia. Semoga hasil dari Tesis ini dapat memberikan manfaat bagi para pembaca dan semua pihak yang berkepentingan.

Jakarta, Juli 2008

Penulis

ABSTRACT

Nowadays, world already encountered competition age of data, data became a very valuable asset for a company to win the business competition. Telecommunication Company is one of company that depends to data needs. PT Indosat Tbk, as one of leading Telecommunication Company in Indonesia, already used database to manage the company's own data. Information, which is collect from database, is an actual data. Besides collecting the actual data, there is valuable information which is still hidden. To get this hidden data, there is a need of good data processing, using data mining. In data mining, there is a need of storage media to accommodate large scale data, which is called data warehouse. This research is purposed to give an alternative solution to PT Indosat Tbk by collecting the data to data warehouse, to find the potential hidden data using data mining deployment.

Stages that will be done on this research, i.e.: analyze business and information needs, collecting data by observing to company, recitation of document and literature, and also data analyzing then to be used for dimensioning data warehouse. And on the last stage is implementation of data mining using time series analysis technique to see trend of PT Indosat Tbk network performance. On the designing, analyzing, and implementation stages, data warehouse and data mining will use tool Microsoft SQL Server 2005 and RapidMiner. By implementation data warehouse, it would aid to see information of network performance quicker, more detail and accurate. Implementation data mining using time series analysis is aid the company to make an accurate decision by using the ability of time series analysis to forecast element data of PT Indosat Tbk network. So by collecting all the data to data warehouse and data mining (time series analysis) implementation, would aid PT Indosat Tbk to represent a good quality of service and the stated KPI (Key Performance Indicator) can be reached or even exceeded.

Keywords : *Data Warehouse, Data Mining, Forecasting, Time Series Analysis*
vii+120 pages; 54 figures; 13 tables; 4 attachments; 0 technical documentation
Bibliography: 18 (1996-2008)

ABSTRAK

Saat ini, dunia telah memasuki era persaingan berbasis data. Data menjadi aset yang sangat berharga bagi perusahaan dalam memenangkan persaingan bisnis. Perusahaan telekomunikasi adalah salah satu perusahaan yang sangat bergantung terhadap data. PT Indosat Tbk, sebagai salah satu perusahaan telekomunikasi di Indonesia, telah memanfaatkan *database* untuk mengelola data perusahaan. Informasi yang diperoleh melalui *database* merupakan data aktual. Selain memperoleh informasi melalui data aktual, terdapat juga informasi berharga lain yang masih tersembunyi. Untuk mendapatkannya diperlukan teknik pengolahan data yang baik yaitu menggunakan *data mining*. Dalam melakukan *data mining* diperlukan suatu media penyimpanan untuk menampung data skala besar yaitu *data warehouse*. Penelitian ini bertujuan untuk memberikan solusi alternatif bagi PT Indosat Tbk dengan pengumpulan data ke suatu sumber pengetahuan (*data warehouse*) dalam mencari potensi-potensi data tersembunyi melalui penerapan *data mining*.

Tahapan yang akan dilakukan dalam penelitian ini yaitu penelaahan kebutuhan bisnis dan informasi, pengumpulan data melalui observasi ke lapangan, kajian dokumen atau literatur, menelaah data kemudian merancang *data warehouse* dan pada tahap akhir dilakukan pengimplementasian *data mining* dengan menggunakan teknik *time series analysis* untuk melihat tren kinerja jaringan dari PT Indosat Tbk. Pada tahap perancangan, analisis dan implementasi *data warehouse* dan *data mining* akan digunakan *tools Microsoft SQL Server 2005* dan *RapidMiner*. Pengimplementasian *data warehouse* pada PT. Indosat Tbk dapat membantu melihat informasi kinerja jaringan dengan lebih cepat, detil dan akurat. Sedangkan pengimplementasian *data mining* melalui *time series analysis* dapat membantu perusahaan dalam pengambilan keputusan yang tepat dengan kemampuan *time series analysis* untuk melakukan prediksi (*forecasting*) terhadap data elemen jaringan yang dimiliki oleh PT. Indosat Tbk. Sehingga dengan pengumpulan data ke *data warehouse* dan penerapan *data mining (time series analysis)* dapat membantu PT. Indosat Tbk menghadirkan *quality of service* yang baik dan target KPI (*Key Performance Indicator*) yang ditetapkan dapat tercapai atau bahkan terlampaui.

Kata kunci : *Data Warehouse, Data Mining, Forecasting, Time Series Analysis*
viii+120 halaman; 54 gambar; 13 tabel; 4 lampiran; 0 dokumentasi teknis
Daftar acuan: 18 (1996-2008)

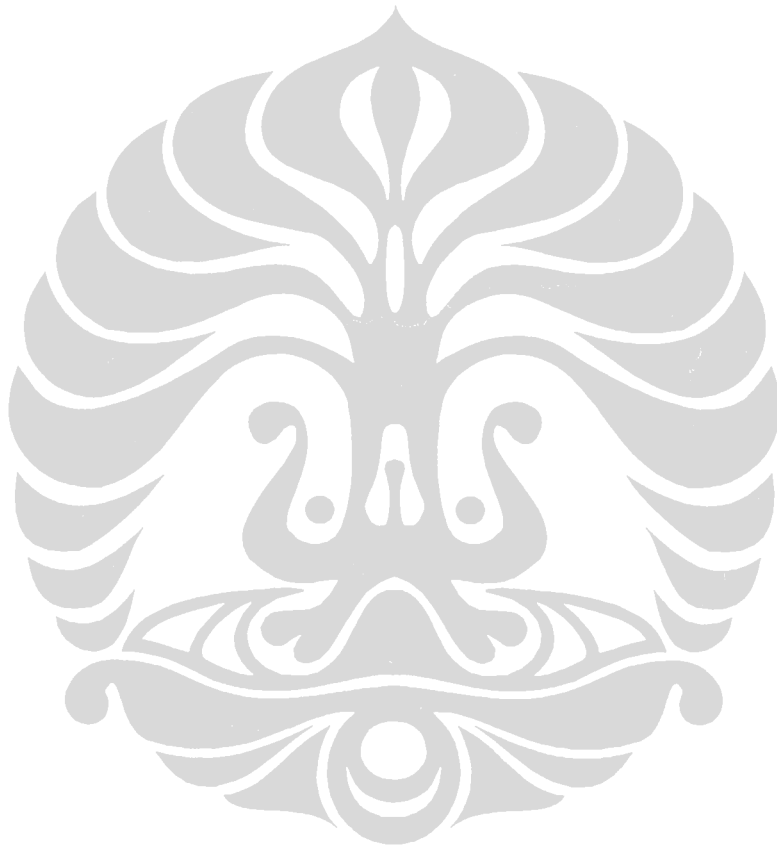
DAFTAR ISI

KATA PENGANTAR.....	i
<i>ABSTRACT</i>	ii
ABSTRAK.....	iii
DAFTAR ISI.....	1
DAFTAR GAMBAR.....	xii
DAFTAR TABEL.....	xv
DAFTAR LAMPIRAN.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Tujuan.....	3
1.3 Manfaat.....	3
1.4 Ruang Lingkup.....	4
BAB II LANDASAN TEORI.....	5
2.1 <i>Data Warehouse</i>	5
2.1.1 Definisi <i>Data Warehouse</i>	5
2.1.2 Keuntungan <i>Data Warehouse</i>	7
2.1.3 Kategori Data pada <i>Data Warehouse</i>	8
2.1.4 Arsitektur <i>Data Warehouse</i>	9
2.1.5 Tahapan <i>Data Warehouse</i>	15
2.1.6 Desain <i>Data Warehouse</i>	18
2.2 <i>Data Mining</i>	21
2.2.1 Definisi <i>Data Mining</i>	21

2.2.2	Teknik <i>Data Mining</i>	23
2.2.3	<i>Predictive Modeling</i>	24
2.2.4	<i>Classification</i>	24
2.2.5	<i>Clustering</i>	26
2.2.6	<i>Link Analysis</i>	27
2.2.7	<i>Deviation Detection</i>	28
2.2.8	<i>Time Series Analysis</i>	29
2.2.9	<i>Time Series Analysis Pattern</i>	30
2.2.10	Algoritma <i>Microsoft Time Series</i>	32
2.2.11	<i>RapidMiner</i>	36
2.2.12	Tahapan <i>Data Mining</i>	37
2.3	Penelitian di Bidang <i>Data Warehouse</i> dan <i>Data Mining</i>	42
BAB III METODOLOGI PENELITIAN		44
3.1	Pengumpulan Data	44
3.2	Menelaah Kebutuhan Bisnis dan Informasi	45
3.3	Menelaah Data dan Perancangan <i>Data Warehouse</i>	45
3.4	Implementasi <i>Data Mining</i>	46
BAB IV PERANCANGAN <i>DATA WAREHOUSE</i>		49
4.1	Profil Perusahaan	49
4.1.1	Latar Belakang Perusahaan.....	50
4.1.2	Visi dan Misi Perusahaan.....	51
4.1.3	Struktur Organisasi Perusahaan	51
4.1.4	Proses Bisnis	53
4.2	Perancangan Arsitektur Perusahaan.....	54

4.2.1	Arsitektur <i>Logical</i>	55
4.2.2	Arsitektur Fisik	56
4.3	Sumber Data.....	57
4.4	Data <i>Staging</i>	59
4.5	Proses ETL.....	60
4.5.1	Proses <i>Extract</i>	60
4.5.2	Proses <i>Cleansing</i>	68
4.5.3	Proses Transformasi.....	69
4.5.4	Proses <i>Loading</i>	72
4.5.5	Manajemen Data	74
4.6	Skema Bintang	74
4.6.1	Tabel Fakta.....	75
4.6.2	Tabel Dimensi.....	76
4.6.2.1	Dimensi <i>Cell</i>	76
4.6.2.2	Dimensi <i>BTS</i>	77
4.6.2.3	Dimensi <i>BSC</i>	77
4.6.2.4	Dimensi <i>Vendor</i>	77
4.6.2.5	Dimensi <i>Time</i>	78
4.6.2.6	Dimensi <i>Geography</i>	78
BAB V <u>I</u> MPLEMENTASI <i>DATA WAREHOUSE</i>		79
5.1	Presentasi <i>Data Warehouse</i>	79
5.1.1	Informasi <i>Traffic</i>	83
5.1.2	Informasi <i>Alert</i>	87
BAB VI <u>I</u> MPLEMENTASI <i>DATA MINING</i>		91

BAB VII_KESIMPULAN DAN SARAN.....	116
DAFTAR PUSTAKA	119
LAMPIRAN-LAMPIRAN	L-ii



DAFTAR GAMBAR

Gambar 2.1 – Arsitektur <i>data warehouse</i>	9
Gambar 2.2 – Contoh transformasi <i>single-field</i>	17
Gambar 2.3 – Contoh transformasi <i>multi-field</i>	17
Gambar 2.4 – <i>Star schema</i> (Connolly and Begg, 2005)	19
Gambar 2.5 – <i>Snow schema</i> (Connolly and Begg, 2005)	20
Gambar 2.6 – <i>Starflake schema</i> (Connolly and Begg, 2005).....	21
Gambar 2.7 - Contoh <i>classification</i> menggunakan <i>tree induction</i>	25
Gambar 2.8 – Contoh <i>classification</i> menggunakan <i>neural induction</i>	25
Gambar 2.9 – Contoh <i>database segmentation</i> menggunakan <i>scatterplot</i>	27
Gambar 2.10 – Contoh visualisasi dari data pada Gambar 2.9	28
Gambar 2.11 – <i>Time series plots</i>	29
Gambar 2.12 – <i>Microsoft Stock Value</i>	31
Gambar 2.13 – <i>Case Transformation</i>	33
Gambar 2.14 – Matrix koefisien	34
Gambar 2.15 – <i>Regression tree</i> pada data <i>time series</i>	35
Gambar 2.16 – Tahapan <i>Data Mining</i>	37
Gambar 2.17 – Tiga Tahap <i>Data Mining</i>	41
Gambar 3.1 – Alur Pengerjaan.....	47
Gambar 4.1 – Struktur Organisasi Tingkat Direktorat.....	52
Gambar 4.2 – Struktur Organisasi di Group <i>Network Quality & Surveillance</i>	53
Gambar 4.3 – Arsitektur <i>Logical</i>	56

Gambar 4.4 – Arsitektur Fisik	56
Gambar 4.5 – Contoh <i>Sheet File</i>	59
Gambar 4.6 – <i>Field Level–Single-Field</i>	70
Gambar 4.7 – <i>Field Level – Multiple-Field</i>	71
Gambar 4.8 – <i>Record Field</i>	72
Gambar 4.9 – Proses <i>Loading</i>	73
Gambar 4.10 – <i>Star Schema Data Measurement</i>	75
Gambar 5.1 – Halaman Utama <i>Data Warehouse</i>	80
Gambar 5.2 – <i>Menu Analysis</i>	81
Gambar 5.3 – Contoh <i>File</i> pada <i>Menu Analysis</i>	82
Gambar 5.4 – <i>Menu Reporting</i>	83
Gambar 5.5 – Laporan <i>Traffic</i> per-vendor	84
Gambar 5.6 – <i>Traffic</i> per-wilayah	85
Gambar 5.7 – <i>Traffic</i> per-bulan	86
Gambar 5.8 – <i>Top 5 Worst Call Drop</i>	88
Gambar 5.9 – <i>KPI Alert</i>	90
Gambar 6.1 – Seleksi Data	93
Gambar 6.2 – <i>Training Data</i>	97
Gambar 6.3 – <i>Select the Data Mining Technique</i>	98
Gambar 6.4 – <i>Mining Models</i>	99
Gambar 6.5 – <i>Actual Data (2007) & Predicted Data (2008)</i>	101
Gambar 6.6 – Pembagian Regional Jaringan PT. Indosat Tbk	102
Gambar 6.7 – Data Aktual Januari – Juni 2008	103
Gambar 6.8 – Perbandingan Prediksi Data dengan Data Aktual	104

Gambar 6.9 – Perbandingan Prediksi Data dengan Data Aktual	105
(Sumatera Bagian Utara).....	105
Gambar 6.10 – Perbandingan Prediksi Data dengan Data Aktual	106
(Sumatera Bagian Selatan).....	106
Gambar 6.11 – Perbandingan Prediksi Data dengan Data Aktual	107
(Jawa Bagian Barat).....	107
Gambar 6.12 – Perbandingan Prediksi Data dengan Data Aktual	108
(Jawa Bagian Tengah).....	108
Gambar 6.13 – Perbandingan Prediksi Data dengan Data Aktual	109
(Jawa Bagian Timur, Bali & Nusa Tenggara).....	109
Gambar 6.14 – Perbandingan Prediksi Data dengan Data Aktual	110
(Jabotabek).....	110
Gambar 6.15 – Perbandingan Prediksi Data dengan Data Aktual	111
(Kalimantan)	111
Gambar 6.16 – Perbandingan Prediksi Data dengan Data Aktual	112
(Sulampapua).....	112
Gambar 6.17 – Perbandingan Prediksi Data Total <i>traffic</i> dengan	113
Data Aktual Total <i>Traffic</i>	113

DAFTAR TABEL

Tabel 2.1 – <i>Data mining operations and associated techniques</i>	23
Tabel 3.1 – Alur Pengerjaan	48
Tabel 4.1 – <i>Metrica Database</i>	58
Tabel 4.2 – <i>Data Sheet File</i>	58
Tabel 4.3 – Hasil Ekstraksi	61
Tabel 4.4 – Contoh <i>field</i> pada proses ekstraksi	62
Tabel 4.5 – Perbedaan Format Data	67
Tabel 4.6 – Perbedaan Penamaan <i>Field</i>	67
Tabel 4.7 – Penyeragaman Penamaan <i>Field</i>	68
Tabel 4.8 – Ketidak-konsistenan Data	69
Tabel 5.1 – <i>Benchmark Report</i>	90
Tabel 6.1 – Tabel Awal	95
Tabel 6.2 – Tabel Hasil Transformasi	96

DAFTAR LAMPIRAN

Lampiran 1 – Daftar Tabel Metrica (Sumber Data).....	L-ii
Lampiran 2 – Ringkasan Tabel DWTemp dan ISATDW	L-ii
Lampiran 3 – Daftar Tabel DWTemp Beserta Atributnya	L-ii
Lampiran 4 – Daftar Tabel ISATDW Beserta Atributnya.....	L-iv

