

Jurnal Ilmiah

ilmu komputer

ISSN 1693-1629

Edisi 11

Publikasi Hasil Penelitian diterbitkan Departemen Ilmu Komputer FMIPA
Institut Pertanian Bogor

1
Pembandingan Stabilitas Algoritma Seleksi Fitur menggunakan Transformasi Ranking Normal
Taufik Djatna, Yasuhiko Morimoto

7
Pengembangan Data Warehouse Program Tracking Stasiun TV di Indonesia
Wisnu Ananta Kusuma, Hari Agung Adrianto, Arsanda Prawisda

15
Pengembangan SIJELITA sebagai Sistem Keamanan Perumahan berbasis SMS dan PSTN
Heru Sukoco, Sri Wahjuni, Andi Setiadi

24
Perkiraan Biaya Perangkat Lunak menggunakan Logika Fuzzy
Irman Hermadi, Wisnu Ananta Kusuma, Denny Setia Mulyadi

32
Pengenalan Kata Berbahasa Indonesia dengan Hidden Markov Model (HMM) menggunakan Algoritma Baum-Welch
Agus Buono, Arief Ramadhan, Ruvinna

41
Seleksi Fitur menggunakan Fast Correlation Based Filter pada Algoritma Voting Feature Intervals 5
Aziz Kustiyo, Hida Nur Firqiani, Endang Purnama Giri



Vol.6 No.2 / Desember 2008

Sekapur Sirih

Pembaca yang budiman,

Dengan mengucapkan Puji dan Syukur kehadiran Allah SWT, Jurnal Ilmiah Ilmu Komputer Edisi 11/Vol.6 No.2 Desember 2008, akhirnya dapat terbit dan sampai di hadapan pembaca.

Dalam Edisi 11 ini, Redaksi menyajikan enam laporan hasil penelitian yang meliputi bidang Jaringan Komputer, Data Warehouse, Kecerdasan Komputasional, dan Rekayasa Perangkat Lunak.

Kami harapkan materi-materi yang disajikan pada Edisi 11 ini dapat memberikan kontribusi terhadap perkembangan dunia Ilmu Komputer di Indonesia.

Selamat membaca !

Bogor, Desember 2008

Redaksi

Daftar Isi

Sekapur Sirih	i
Daftar Isi	iii
Pembandingan Stabilitas Algoritma Seleksi Fitur menggunakan Transformasi Ranking Normal <i>Taufik Djatna, Yasuhiko Morimoto</i>	1
Pengembangan <i>Data Warehouse Program Tracking</i> Stasiun TV di Indonesia <i>Wisnu Ananta Kusuma, Hari Agung Adrianto, Arsanda Prawisda</i>	7
Pengembangan <i>SIJELITA</i> sebagai Sistem Keamanan Perumahan berbasis <i>SMS</i> dan <i>PSTN</i> <i>Heru Sukoco, Sri Wahjuni, Andi Setiadi</i>	15
Perkiraan Biaya Perangkat Lunak menggunakan Logika <i>Fuzzy</i> <i>Irman Hermadi, Wisnu Ananta Kusuma, Denny Setia Mulyadi</i>	24
Pengenalan Kata Berbahasa Indonesia dengan <i>Hidden Markov Model (HMM)</i> menggunakan Algoritma <i>Baum-Welch</i> <i>Agus Buono, Arief Ramadhan, Ruvinna</i>	32
Seleksi Fitur menggunakan <i>Fast Correlation Based Filter</i> pada Algoritma <i>Voting Feature Intervals 5</i> <i>Aziz Kustiyo, Hida Nur Firqiani, Endang Purnama Giri</i>	41

Pengembangan Data Warehouse Program Tracking Stasiun TV di Indonesia

Wisnu Ananta Kusuma, Hari Agung Adrianto, Arsanda Prawisda

Departemen Ilmu Komputer, FMIPA-IPB

Abstrak

Stasiun TV berusaha untuk meningkatkan rating, share, dan jumlah penonton dengan memperhatikan biaya produksi yang dikeluarkan. Data stasiun TV yang berisi rating, share, jumlah penonton, dan biaya produksi adalah data program tracking. Data program tracking diterima stasiun TV dari perusahaan penyedia data setiap minggu. Data acara tersebut menjadi acuan dalam menganalisis potensi sebuah acara. Untuk memudahkan proses analisis, maka dibuat data warehouse yang merupakan tempat penyimpanan data yang terintegrasi, multidimensi, dan menampilkan data dalam suatu bentuk yang diharapkan akan memudahkan proses analisis dalam pembuatan keputusan.

Hasil dari penelitian ini adalah suatu data warehouse untuk data program tracking dan suatu OLAP browser yang mempunyai fasilitas untuk menambah data yang datang setiap minggunya dan visualisasi berupa tabel pivot dan diagram batang dalam menampilkan data numerik dan tabel relasional untuk menampilkan data kategorik. Visualisasi ini dibuat untuk mempermudah pengguna dalam melihat data dalam proses analisis.

Kata Kunci : Data warehouse, Multidimensi, Online Analytical Processing (OLAP), Skema bintang.

PENDAHULUAN

Latar Belakang

Saat ini persaingan di dunia pertelevisian semakin ketat. Setiap stasiun TV berlomba-lomba memasang acara yang menarik banyak penonton. Dalam hal ini, stasiun TV berusaha untuk meningkatkan rating dan share, dengan memperhatikan biaya produksi yang dikeluarkan. Data stasiun TV yang berisi rating, share, jumlah penonton, dan biaya produksi adalah data program tracking.

Data program tracking diterima stasiun TV dari perusahaan penyedia data setiap minggu. Data acara selama seminggu itu yang menjadi acuan dalam menganalisis potensi sebuah acara. Tetapi data dalam seminggu belum cukup memberikan informasi dalam proses analisis data. Karena itu, stasiun TV harus menganalisis minggu-minggu sebelumnya (*bulan dan tahun sebelumnya*) untuk mengetahui ketepatan penempatan acara, pola penempatan acara di stasiun TV, dan pengambilan keputusan lainnya.

Proses analisis data program tracking masih dilakukan secara manual, sehingga memakan waktu yang cukup lama. Hal tersebut tidak efisien jika melihat data berikutnya akan datang lagi hanya dalam waktu satu minggu. Oleh karena itu, perlu dibuat data warehouse yang merupakan tempat penyimpanan data yang terintegrasi dan dapat digunakan untuk kueri dan analisis.

Kueri dapat dengan mudah dijalankan dan ditampilkan hasilnya dengan menggunakan suatu OLAP browser. OLAP browser yang banyak dikembangkan saat ini masih terbatas pada proses analisis untuk data numerik, sedangkan dalam menganalisis data program tracking, stasiun TV di Indonesia membutuhkan suatu proses analisis untuk data kategorik seperti pada tipe dan asal negara dari suatu acara. Oleh karena itu, perlu dikembangkan suatu OLAP browser yang dapat digunakan untuk proses analisis data kategorik, sesuai dengan kebutuhan stasiun TV di Indonesia tersebut.

Tujuan

Tujuan penelitian ini adalah:

1. Mengembangkan data warehouse program tracking stasiun TV di Indonesia.
2. Mengembangkan suatu OLAP browser untuk memenuhi kebutuhan proses analisis data numerik dan data kategorik pada data warehouse program tracking stasiun TV di Indonesia.

Ruang Lingkup

Penelitian ini dibatasi pada pengembangan data warehouse untuk stasiun TV di Indonesia. Data yang digunakan adalah data program tracking stasiun TV di Indonesia, tertanggal 26 Desember 2004 sampai dengan 6 Agustus 2005.

dibaca saat proses analisis. Selain itu, penamaan atribut juga akan ditransformasi untuk memudahkan pengguna untuk mengingatnya.

c Pembersihan Data

Atribut dengan nilai NULL akan diisikan suatu kata untuk memudahkan pengguna dalam proses analisis, dan pengguna tidak kehilangan informasi.

d Memuat Data ke dalam Basis Data Data Warehouse (loading)

Data yang telah mengalami proses ekstraksi, transformasi, dan pembersihan sudah siap digunakan dalam pengembangan data warehouse.

e Refresh

Data akan disegarkan jika ada data baru yang masuk.

2 Proses pengembangan data warehouse ini akan dilakukan menggunakan Microsoft SQL Server 2000. Langkah-langkah tersebut adalah:

a Menentukan Proses

Pada tahap ini ditentukan proses-proses untuk data warehouse program tracking stasiun TV di Indonesia.

b Menentukan Atribut untuk Tabel Fakta

Pada tahap ini ditentukan atribut-atribut pada tabel fakta data warehouse program tracking stasiun TV di Indonesia.

c Menentukan Dimensi

Pada tahap ini ditentukan dimensi-dimensi yang ingin dilihat hubungannya pada data warehouse program tracking stasiun TV di Indonesia. Selain itu, skema yang digunakan untuk membentuk dimensi adalah skema bintang karena masing-masing data berasal dari tabel-tabel terpisah yang tidak saling berhubungan.

d Menentukan Measure

Pada tahap ini ditentukan measure yang ingin dianalisis pada data warehouse program tracking stasiun TV di Indonesia.

e Menyimpan pre-calculation dalam Tabel Fakta

Tahap ini tidak dilakukan karena tidak dilakukan suatu perhitungan tertentu dalam menganalisis data.

f Membuat Keterangan pada Tabel Dimensi

Pada tahap ini dibuat keterangan pada dimensi data warehouse program tracking stasiun TV di Indonesia.

g Menentukan Durasi dari Basis Data

Pada tahap ini ditentukan dimensi waktu pada data warehouse program tracking stasiun TV di Indonesia.

h Melacak Slowly Changing Dimension

Tahap ini tidak dilakukan karena proses diasumsikan tidak berubah.

i Menentukan Kueri

Pada tahap ini ditentukan server digunakan untuk menjalankan kueri penyimpulan pada data warehouse program tracking stasiun TV di Indonesia. Selain itu, ditentukan juga operasi OLAP yang ingin dijalankan untuk menganalisis data tersebut.

Untuk pengembangan OLAP browser akan dilakukan tahap-tahap sebagai berikut:

1 Analisis

Pada tahap ini ditentukan spesifikasi fungsi dari OLAP browser. Spesifikasi fungsi ini berdasar kebutuhan pemakai.

2 Desain

Perancangan OLAP browser didefinisikan sebagai proses dimana kebutuhan OLAP browser telah didefinisikan pada tahap analisis kebutuhan. Perancangan OLAP browser dibagi menjadi perancangan input, proses, dan output.

3 Pengkodean

Implementasi dirancang dan dibangun dengan menggunakan perangkat keras processor Pentium 4 2,6 GHz, memory 256 MB, harddisk 30 GB, dan alat input mouse dan keyboard. Selain itu, perangkat lunak yang digunakan adalah OLAP browser Operasi Windows XP, bahasa pemrograman Microsoft Visual Basic 6.0, DBMS Microsoft SQL Server 2000, dan Microsoft Office Excel 2003

4 Pengujian

Untuk mengetahui apakah OLAP browser dapat berfungsi dengan baik, maka dilakukan pengujian. Tahapan ini menguji kubus data dalam data warehouse yang divisualisasikan dengan OLAP browser tersebut.

5 Pemeliharaan

Tahap ini tidak dilakukan pada pengembangan OLAP browser karena keterbatasan waktu.

HASIL DAN PEMBAHASAN

Pengembangan Data Warehouse

Pengembangan data warehouse telah melalui beberapa proses, yaitu:

1 Praproses yang dilakukan adalah:

a Ekstraksi Data

Atribut yang merupakan pengulangan atau nilainya terdapat pada atribut lain akan dihapus. Pada proses ini didapatkan atribut yang dipakai dari 25 atribut diekstrak menjadi 15 atribut.

b Transformasi Data

Atribut yang ditransformasi adalah:

Atribut Date akan dipisah menjadi empat kolom, yaitu kolom Years, kolom Months, kolom Weeks, dan kolom Dates. Atribut Date dengan format tanggal DD/MM/YY dipisah, dengan bagian YY diisikan ke kolom Years, bagian MM diisikan ke kolom Months, dan bagian DD diisikan ke kolom Dates. Kolom Weeks didapat dengan rumus:

$$\text{Weeks} = \frac{\text{different} - \text{day}}{7} + 1$$

Tabel 4 Deskripsi selang durasi per jam

Selang	Deskripsi Durasi (menit)
1	1-60
2	61-120
3	121-180
4	181-240
5	241-300
6	301-360
7	361-420
8	421-480
9	481-540
10	541-600
11	601-660
12	661-720
13	721-780
14	781-840
15	841-900
16	901-960
17	961-1020
18	1021-1080
19	1081-1140
20	1141-1200
21	1201-1260
22	1261-1320
23	1321-1380
24	1381-1440

Selang	Deskripsi Durasi (menit)
28	811-840
29	841-870
30	871-900
31	901-930
32	931-960
33	961-990
34	991-1020
35	1021-1050
36	1051-1080
37	1081-1110
38	1111-1140
39	1141-1170
40	1171-1200
41	1201-1230
42	1231-1260
43	1261-1290
44	1291-1320
45	1321-1350
46	1351-1380
47	1381-1410
48	1411-1440

Tabel 5 Deskripsi selang durasi per setengah jam

Selang	Deskripsi Durasi (menit)
1	1-30
2	31-60
3	61-90
4	91-120
5	121-150
6	151-180
7	181-210
8	211-240
9	241-270
10	271-300
11	301-330
12	331-360
13	361-390
14	391-420
15	421-450
16	451-480
17	481-510
18	511-540
19	541-570
20	571-600
21	601-630
22	631-660
23	661-690
24	691-720
25	721-750
26	751-780
27	781-810

- Nama atribut akan dikonsistenkan dengan urutan *Id_program*, *Market*, *Program_name*, *Program_type*, *Program_source*, *Class*, *Channel*, *Years*, *Months*, *Weeks*, *Dates*, *Days*, *Time_start_halfhour*, *Time_start_onehour*, *Time_start_periode*, *Duration_halfhour*, *Duration_onehour*, *Cost*, *TVR*, *Audience*, dan *Share*. Untuk penamaan atribut *Years*, *Months*, *Weeks*, *Dates*, dan *Days* dibuat jamak untuk menghindari kerancuan dengan fungsi *Year()*, *Month()*, dan *Day()* pada Microsoft SQL Server 2000.

c Pembersihan Data

- Atribut *Market* yang NULL pada data *program tracking* stasiun TV menggambarkan program acara yang ditayangkan di seluruh kota, sehingga akan diisikan kata *all* di *cell* tersebut.
- Atribut *Class* adalah suatu atribut yang berisikan kode-kode rahasia pihak penyedia data, sehingga pihak TV tidak mengetahui artinya. Atribut *Class* yang bernilai NULL akan diisikan kata *unknown* di *cell* tersebut.
- Seluruh *row* dihapus jika atributnya (selain *Market*, *Class*, *TVR*, *Share*, *Audience*, dan *Cost*) ada yang bernilai NULL.

d Memuat Data ke dalam Basis Data Data Warehouse (loading)

Data yang telah mengalami proses ekstraksi, transformasi, dan pembersihan sudah siap digunakan dalam pengembangan *data warehouse*.

adalah data *program tracking* dalam bentuk *file spreadsheet* (Excel). Setelah itu data ini akan diproses. Hasil dari fungsi ini akan mengalami pemrosesan kubus data. Input pada OLAP *browser* adalah data dimensi yang akan dilihat hubungannya dan *measure* yang ingin dianalisis nilainya.

- **Perancangan Proses**

Proses-proses pada OLAP *browser* ini adalah data yang diimpor akan mengalami pra-proses data, *update*-an tabel dimensi dan tabel fakta, dan pemrosesan kubus data. Operasi OLAP akan dilakukan berdasarkan dimensi dan *measure* yang diinputkan. Operasi OLAP yang dijalankan adalah *roll-up*, *drill-down*, *slice*, dan *dice*. Hasil dari operasi ini akan ditampilkan pada *window*.

- **Perancangan Output**

Output ditampilkan dalam tabel *pivot* dan diagram batang untuk menganalisis data numerik dan operasi OLAP *roll-up*, *drill-down*, *slice*, dan *dice*, dengan fungsi penyimpulan *summary* atau *maximal*. Selain itu, output ditampilkan juga dalam bentuk tabel relasional untuk menganalisis data kategorik dengan fungsi *modus*.

3 Pengkodean

Fungsi yang tersedia pada OLAP *browser* ini adalah:

- **Update Data**, fungsi untuk menambah data yang terdiri dari fungsi:

a **Ekstraksi, pembersihan, dan transformasi data.**

Data baru (*data pada minggu terbaru*) yang berbentuk *file spreadsheet* (Excel) diimpor ke Microsoft SQL Sever 2000 dan dimasukkan ke dalam tabel sementara pertama (*tabel Exceltemp*). Lalu pada tabel sementara pertama ini dilakukan pembersihan dan transformasi data. Hasilnya dimasukkan ke dalam tabel sementara kedua (*tabel Tabel_asli_temp*).

b **Update anggota dimensi dan tabel fakta.**

Pada tabel sementara kedua dicari nilai-nilai anggota dimensi yang belum tercatat pada masing-masing tabel dimensi. Nilai dimensi yang belum ada ditambahkan sebagai anggota baru pada tabel dimensi tersebut. Setelah itu, tabel-tabel dimensi dihubungkan dengan tabel sementara kedua untuk meng-*update* tabel fakta dengan data terbaru (*minggu terbaru*). Data pada tabel sementara kedua juga ditambahkan ke tabel asli.

- **Process Cube**

Fungsi ini merupakan fungsi untuk memproses kembali kubus data pada Microsoft Analysis Service SQL Server 2000 setelah data di-*update*.

- **Pengoperasian OLAP**

Operasi OLAP yang diterapkan pada kubus data adalah:

a **Operasi Roll-up**

Operasi *roll-up* dilakukan untuk:

- i Mengetahui biaya produksi per bulan untuk tipe

acara dengan *source* acara atau *channel* tertentu.

- 2 Mengetahui *rating* dan *share maximal* per bulan untuk tipe atau *source* acara pada *channel* tertentu.

b **Drill-down**

Operasi *drill-down* dilakukan untuk:

- 1 Mengetahui *rating*, *share*, jumlah penonton, dan biaya produksi pada masing-masing *channel* untuk hari tertentu dan asal negara (*source*) atau tipe acara tertentu.

- 2 Mengetahui *rating*, *share*, jumlah penonton, dan biaya produksi tiap acara (*atau suatu tipe acara*) pada waktu tayang per setengah jam untuk hari atau *channel* tertentu.

- 3 Mengetahui *rating*, *share*, jumlah penonton, dan biaya produksi tiap acara (*atau suatu tipe acara*) pada durasi per setengah jam untuk hari atau *channel* tertentu.

c **Slice and dice**

Operasi *dice* dan *slice* adalah operasi yang paling sering dilakukan stasiun TV untuk melihat *rating*, *share*, jumlah penonton, dan biaya produksi. Semua kombinasi dimensi dapat dilakukan dalam operasi ini. Hasil dari operasi OLAP divisualisasikan dengan:

- **Pivot table** dan diagram batang

Fungsi ini merupakan fungsi untuk visualisasi data numerik. Data diambil dari kubus data. Sumbu Y menggambarkan *measure*, sedangkan sumbu X dan warna pada batang menunjukkan dimensi yang ingin dilihat hubungannya.

- **Visualisasi Kategorik**

Fungsi ini merupakan visualisasi untuk data yang merupakan dimensi (*bersifat kategorik*). Visualisasi berupa tabel, dengan data yang diambil langsung dari basis data relasional. Fungsi agregasi yang dibuat pada fungsi ini adalah *modus*.

4 Pengujian

Pengujian pada tahap pertama, yaitu melakukan pengecekan nilai data yang ditampilkan untuk masing-masing *measures* di kubus data. Pengecekan ini menunjukkan bahwa nilai-nilai data adalah sama dengan nilai data yang terdapat pada *file* asli.

Pengujian tahap kedua adalah melakukan pengecekan terhadap fungsi-fungsi yang terdapat pada OLAP *browser*. Pengecekan ini menunjukkan bahwa kubus data dapat diakses dengan baik dan fungsi-fungsi yang terdapat pada *browser* tersebut berjalan dengan baik. Namun, terdapat kelemahan pada *browser* ini jika data yang ditampilkan terlalu banyak, fasilitas visualisasi tidak dapat berfungsi dengan baik dan prosesnya agak lama. Skala diagram batang terlalu kecil sehingga data sulit untuk dilihat.

Hasil pengujian menunjukkan bahwa kubus data dapat merepresentasikan nilai-nilai data dengan tepat sehingga OLAP *browser* dapat memvisualisasikan data tersebut secara tepat, baik dalam bentuk tabel *pivot*, diagram batang, maupun tabel relasional.