

Kompresi Data pada Laporan Cuaca Menggunakan Metode *Huffman Coding*

IF 38-01

Aliya Nur Rezkita 1301144161

Febby Febriansyah 1301140371

Rizkiyana Prima Putra 1301140181

Tari Lestari 1301144281

Fakultas Informatika, Telkom University, Bandung

Kontak : 082240203610 (Rizkiyana Prima Putra)

Abstrak—Sebuah instansi pengamat cuaca memiliki beberapa alat yang dapat memprediksi kemungkinan cuaca yang akan terjadi setiap harinya. Kemungkinan cuaca yang dapat diprediksi diantaranya cerah, berawan, hujan, dan berkabut. Hasil dari setiap sejumlah n observasi yang dilakukan akan ditransmisikan ke server utama. Namun, permasalahannya adalah ketika banyaknya alat yang digunakan akan mengakibatkan komunikasi transmisi data menjadi *overload*. Sehingga diperlukanlah suatu metode yang dapat mengompresi data ke dalam ukuran yang lebih kecil. Pada paper ini kami menggunakan metode *Huffman Code* untuk meng-*encoding* 4 nilai probabilitas cuaca menjadi deretan bit yang unik agar proses transmisi data menjadi lebih efisien.

Kata Kunci—Kompresi data, Algoritma Huffman, Laporan Cuaca, Probabilitas.

Abstract—A weather forecaster has several tools that can predict the possible weather that will occur each day. That possible weather conditions are sunny, cloudy, rainy, and foggy. The results of each n observation will be transmitted to the main server. However, the problem is when the number of tools used will make data transmission communication becomes textit overload. So we need a method that can compress data into smaller sizes. In this paper we use the textit Huffman Code method to encode the 4 weather probability values into unique bit sequences to make data transmission more efficient.

Keywords—Data Compression, Huffman Coding, Weather Report, Probability.

I. PENDAHULUAN

Banyak kegiatan atau aktifitas manusia yang bergantung pada faktor cuaca. Faktor cuaca ini terkadang memiliki pengaruh yang sangat besar bagi keberlangsungan kegiatan yang dilakukan. Peran seorang prakiawan cuaca sangat dibutuhkan mengingat cuaca adalah kondisi udara yang berlangsung dalam jangka waktu yang singkat, maka proses prakiraan cuaca jangka pendek harus dilakukan secara cepat dan seakurat mungkin. Kondisi cuaca di suatu tempat ditentukan oleh sejumlah faktor diantaranya temperatur udara, kelembaban udara, arah angin, kecepatan angin dan sebagainya. Dengan melihat faktor-faktor ini, seorang prakiawan cuaca dapat memprediksikan kondisi cuaca yang akan berlangsung pada keesokan harinya.

Berbagai metode dilakukan untuk memperoleh hasil informasi cuaca yang akurat, salah satunya menggunakan pen-

dekatan model cuaca numerik. Perkembangan model cuaca numerik seiring dengan perkembangan kemampuan komputasi dan penambahan jaringan pengamatan telah mencapai akurasi prediksinya yang baik dan sudah banyak digunakan dalam membuat prakiraan cuaca oleh pusat layanan cuaca di banyak negara. Pola cuaca yang berbeda antar wilayah, mengharuskan dilakukan pengujian model cuaca numerik seperti pemilihan skema parameterisasi, syarat awal, waktu spin-up agar mampu menghasilkan prediksi yang terbaik [1].

Namun, penggunaan model numerical melibatkan perhitungan parameter cuaca yang cukup banyak, umumnya prakiraan cuaca yang menggunakan metode ini mempunyai keluaran (file output) yang cukup besar. Besar kecilnya hasil keluaran ditentukan oleh beberapa faktor, antara lain: luas daerah yang diprediksi, panjang prediksi, resolusi spasial yang digunakan, dan parameter cuaca yang dihasilkan [2].

Untuk Kota Bandung yang mempunyai wilayah yang cukup luas, penggunaan metode numerical ini akan memberikan permasalahan yang cukup signifikan dalam hal distribusi dan penyimpanan data-data prediksi cuaca. Ditambah lagi penggunaan dynamical downscaling (perolehan informasi spasial cuaca yang lebih detil/resolusi tinggi) untuk peningkatan akurasi prakiraan yang akan semakin memperbesar ukuran file output. Oleh karena itu, dibutuhkan suatu proses pengubahan sekumpulan data menjadi suatu bentuk kode untuk menghemat kebutuhan tempat penyimpanan dan waktu untuk transmisi data.

Dalam paper ini, pengubahan bentuk data menjadi kode, penulis akan menggunakan Algoritma Huffman yaitu algoritma dimana setiap nilai-nilai probabilitas masing-masing cuaca dikodekan hanya dengan rangkaian beberapa bit, dimana nilai yang sering muncul dikodekan dengan rangkaian bit yang pendek dan nilai yang jarang muncul dikodekan dengan rangkaian bit yang lebih panjang sehingga akan mempermudah untuk menyelesaikan suatu masalah perkiraan cuaca.

II. METODE PENELITIAN

Tahapan penelitian yang akan dilakukan dalam paper ini adalah menggunakan algoritma Huffman untuk pengubahan sekumpulan data menjadi suatu bentuk kode agar menghemat kebutuhan tempat penyimpanan dan waktu transmisi data.

Algoritma Huffman

Algoritma Huffman adalah salah satu algoritma kompresi. Algoritma Huffman merupakan algoritma yang paling terkenal untuk mengompres data. Terdapat tiga fase dalam menggunakan algoritma Huffman untuk mengompres sebuah data, pertama adalah fase pembentukan pohon Huffman (*tree*), kedua fase encoding dan ketiga fase decoding. Prinsip yang digunakan oleh algoritma Huffman adalah data yang sering muncul di-encoding dengan rangkaian bit yang pendek dan data yang jarang muncul di-encoding dengan rangkaian bit yang lebih panjang. Teknik kompresi algoritma Huffman mampu memberikan penghematan pemakaian memori sampai 30%. Algoritma Huffman mempunyai kompleksitas $O(n \log n)$ untuk himpunan dengan n data [3].

Langkah-langkah pembentukan pohon Huffman adalah sebagai berikut :

Algorithm 1 Algoritma Huffman

- 1: Baca semua nilai di dalam himpunan data cuaca untuk menghitung frekuensi kemunculan setiap nilainya. Setiap nilai dinyatakan sebagai pohon bersimpul tunggal. Setiap simpul di-assign dengan frekuensi kemunculan nilai tersebut.
- 2: Terapkan strategi algoritma sebagai berikut, gabungkan dua buah pohon yang mempunyai frekuensi terkecil pada sebuah akar. Setelah digabungkan, akar tersebut akan mempunyai frekuensi yang merupakan jumlah dua buah pohon-pohon penyusunnya.
- 3: Ulangi langkah 2 sampai hanya tersisa satu buah pohon Huffman. Agar pemilihan dua pohon yang akan digabungkan berlangsung cepat, maka semua pohon yang ada selalu terurut menaik berdasarkan frekuensi.

Proses encoding untuk satu data dimulai dengan membuat pohon Huffman terlebih dahulu. Setelah itu, kode untuk satu data disusun sebagai biner yang dibaca dari akar sampai ke daun pohon Huffman.

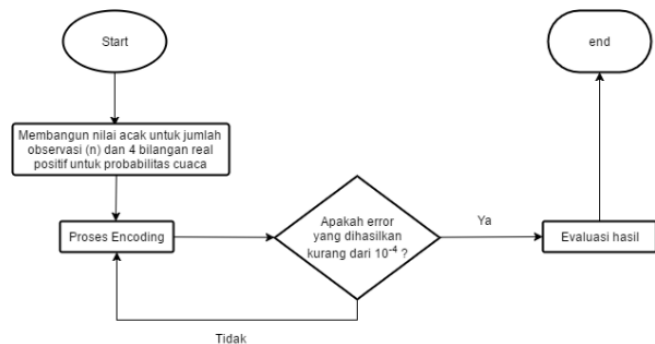
Langkah-langkah untuk melakukan encoding pada suatu data adalah sebagai berikut :

Algorithm 2 Encoding

- 1: Tentukan data yang akan di-encode.
- 2: Mulai dari akar, baca setiap bit yang ada pada cabang yang bersesuaian sampai menemukan daun di mana data itu berada.
- 3: Ulangi langkah 2 sampai seluruh data di-encode.

III. SIMULASI

Flowchart



Gambar 1. Flowchart

Input Simulasi

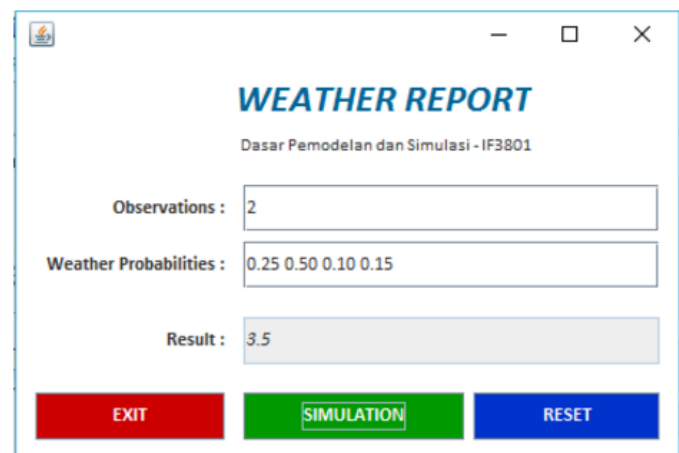
Input simulasi pada kasus Weather Report ini adalah dua baris yang terdiri atas jumlah observasi N dengan interval $1 \leq N \leq 20$ pada baris pertama, lalu 4 angka desimal yang merepresentasikan probabilitas $P_{sunny}, P_{cloudy}, P_{rainy}, P_{frogs}$ yang bila dijumlahkan hasilnya sama dengan 1 pada baris kedua.

Output Simulasi

Output pada simulasi kasus Weather Report ini adalah jumlah bit minimum dari proses encoding dari input yang diberikan dengan menghasilkan relative error pada 10^{-4} .

IV. HASIL SIMULASI

Berikut hasil simulasi berupa *running* program yang telah kami buat :



Gambar 2. Program Weather Report

Program *Weather Report* diatas dibuat menggunakan bahasa java dengan tampilan GUI (*Graphical User Interface*), dimana pada baris pertama merupakan jumlah observasi, lalu pada baris ke dua merupakan 4 nilai probabilitas cuaca dengan spasi sebagai pemisahannya. Nilai-nilai tersebut haruslah berjumlah sama dengan 1. Dan baris ke tiga sebagai output hasil

simulasi.

Dari program tersebut kami melakukan 2 metode percobaan, pertama kami buat jumlah observasinya tetap namun nilai probabilitasnya berbeda-beda (pada letaknya), dan yang kedua jumlah observasinya berbeda-beda namun nilai probabilitasnya sama. Hasilnya dapat dilihat pada tabel berikut :

Tabel I
HASIL SIMULASI DIMANA JUMLAH OBSERVASI TETAP DAN NILAI
PROBABILITASNYA BERBEDA-BEDA

Observasi	Probabilitas Cuaca	Output Simulasi
5	0.15 0.20 0.15 0.50	8.9574225
5	0.50 0.20 0.15 0.15	8.9574225
5	0.20 0.50 0.15 0.15	8.9574225

Tabel II
HASIL SIMULASI DIMANA JUMLAH OBSERVASI BERBEDA-BEDA DAN NILAI
PROBABILITASNYA SAMA

Observasi	Probabilitas Cuaca	Output Simulasi
5	0.15 0.20 0.15 0.50	8.9574225
10	0.15 0.20 0.15 0.50	17.883808339171875
15	0.15 0.20 0.15 0.50	26.810855613654084

Berdasarkan kedua tabel diatas, dapat dilihat perbandingan bahwa ketika jumlah observasi yang sama namun nilai probabilitas-probabilitasnya berbeda (pada letak penginputannya) maka akan menghasilkan output yang sama. Artinya, nilai-nilai probabilitas pada setiap perkiraan cuaca yang dibangkitkan tidak terlalu berpengaruh pada jumlah bit yang dihasilkan pada proses *encoding*. Namun ketika jumlah observasinya berbeda-beda dan nilai probabilitasnya sama maka output jumlah bit yang dihasilkan akan berbeda-beda pada proses *encoding*-nya.

V. KESIMPULAN

Kesimpulan yang didapat dari hasil analisis dan simulasi kompresi data pada laporan cuaca menggunakan metode Huffman Coding yang kami lakukan yaitu :

- 1) Dengan algoritma huffman ini mampu memberikan penghematan pemakaian memori sampai 30%, nilai ini didapat karena kompleksitas algoritma Huffman yaitu $O(n \log n)$ untuk himpunan dengan n data.
- 2) Banyaknya jumlah observasi dan nilai probabilitas pada masing-masing cuaca akan menghasilkan jumlah bit data dari proses *encoding* yang berbeda-beda.
- 3) Pada algoritma pembentukan *tree* Huffman ini, frekuensi kemunculan nilai probabilitas pada setiap observasi akan mempengaruhi panjang/jumlah bit yang dihasilkan. Semakin sering frekuensi kemunculan nilai probabilitas pada sejumlah n observasi akan menghasilkan panjang/jumlah bit yang relatif sedikit. Sedangkan apabila frekuensi kemunculan nilai probabilitasnya semakin jarang, maka akan menghasilkan panjang/jumlah bit yang lebih banyak. Hal ini akan berpengaruh terhadap pemakaian memori dan efisiensi transmisi data yang dikirimkan.

DAFTAR PUSTAKA

- [1] Gustari, Indra, Tri Wahyuni Hadi, Safwan Hadi, dan Findy Renggono. *Jurnal Meteorologi dan Geofisika* Vol. 13 No. 2: 119-130. 2012
- [2] Hanggoro, Wido, Iis Widya Harmoko, Setyawan Widyarto. *Pendistribusian Data NWP dengan GrADS Data Server*. Seminar Nasional Informatika. 2012.
- [3] Rinaldi Munir. *Kompresi Teks Menggunakan Algoritma dan Pohon Huffman*. Informatika STEI ITB. 2007.