

**IMPLEMENTASI DYNAMIC MARKOV COMPRESSION (DMC)
UNTUK KOMPRESI DATA TEKS UNICODE**

Skripsi



Oleh

ALEXANDER ADY SAPUTRA

71110156

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA

2016

**IMPLEMENTASI DYNAMIC MARKOV COMPRESSION
(DMC) UNTUK KOMPRESI DATA TEKS UNICODE**

Skripsi



Diajukan kepada Program Studi Teknik Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana
Sebagai Salah Satu Syarat dalam Memperoleh Gelar
Sarjana Komputer

Disusun oleh

ALEXANDER ADY SAPUTRA

71110156

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA
2016

PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

IMPLEMENTASI ALGORITMA DYNAMIC MARKOV COMPRESSION (DMC) UNTUK KOMPRESI DATA TEKS UNICODE

yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan Sarjana Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi keserjanaan di lingkungan Universitas Kristen Duta Wacana maupun di Perguruan Tinggi atau instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.

Jika dikemudian hari didapati bahwa hasil skripsi ini adalah hasil plagiasi atau tiruan dari skripsi lain, saya bersedia dikenai sanksi yakni pencabutan gelar keserjanaan saya.

Yogyakarta, 9 Maret 2016



ALEXANDER ADY SAPUTRA
71110156

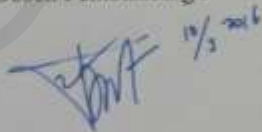
HALAMAN PERSETUJUAN

Judul Skripsi : IMPLEMENTASI ALGORITMA DYNAMIC
MARKOV COMPRESSION (DMC) UNTUK
KOMPRESI DATA TEKS UNICODE

Nama Mahasiswa : ALEXANDER ADY SAPUTRA
NIM : 71110156
Matakuliah : Skripsi (Tugas Akhir)
Kode : TIW276
Semester : Genap
Tahun Akademik : 2015/2016

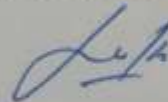
Telah diperiksa dan disetujui di
Yogyakarta,
Pada tanggal 9 Maret 2016

Dosen Pembimbing I



Antonius Rachmat C., S.Kom., M.Cs.

Dosen Pembimbing II



Lukas Chrisantyo, S.Kom., M.Eng.

HALAMAN PENGESAHAN

IMPLEMENTASI ALGORITMA DYNAMIC MARKOV COMPRESSION (DMC) UNTUK KOMPRESI DATA TEKS UNICODE

Oleh: ALEXANDER ADY SAPUTRA / 71110156

Dipertahankan di depan Dewan Penguji Skripsi
Program Studi Teknik Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana - Yogyakarta
Dan dinyatakan diterima untuk memenuhi salah satu syarat memperoleh gelar
Sarjana Komputer
pada tanggal 16 Maret 2016

Yogyakarta, 20 Maret 2016
Mengesahkan,

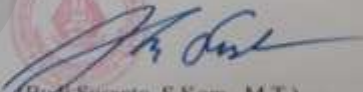
Dewan Penguji:


1. Antonius Rachmat C., S.Kom., M.Cs.
2. Lukas Chrisantyo, S.Kom., M.Eng.
3. Joko Purwadi, M.Kom.
4. Willy Sudarto Raharjo, S.Kom., M.Cs.

DUTA WACANA

Dekan

Ketua Program Studi


(Budi Susanto, S.Kom., M.T.)


(Gloria Virginia, Ph.D.)

UCAPAN TERIMA KASIH

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa yang telah melimpahkan rahmat dan anugerah, sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul Implementasi Dynamic Markov Compression (DMC) Untuk Kompresi Data Teks *Unicode*.

Dalam menyelesaikan pembuatan program dan laporan Tugas Akhir ini, penulis telah banyak menerima bimbingan, saran, dan masukan dari berbagai pihak, baik secara langsung maupun tidak langsung. Untuk itu dengan segala kerendahan hati, pada kesempatan ini penulis menyampaikan ucapan terima kasih kepada :

1. Antonius Rachmat C., S.Kom.,M.Cs selaku dosen pembimbing I yang telah memberikan bimbingannya dengan sabar dan baik kepada penulis.
2. Lukas Chrisantyo, S.Kom., M.Eng selaku dosen pembimbing II atas bimbingannya, petunjuk dan masukan yang diberikan selama pengerjaan tugas akhir ini.
3. Keluarga tercinta yang telah memberi dukungan dan semangat.
4. Dessy, yang memberikan saran dan kritik dalam perancangan dan pembuatan sistem, serta dukungan dan semangat yang selama ini tidak pernah berhenti mengalir.
5. Teman-teman lain yang tidak dapat disebutkan satu per satu, terima kasih atas dukungan dan doa kalian.
6. Pihak lain yang tidak dapat disebutkan satu per satu, sehingga Tugas Akhir ini dapat terselesaikan dengan baik.

Penulis menyadari bahwa program dan laporan Tugas Akhir ini masih jauh dari sempurna. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari pembaca sekalian, sehingga suatu saat penulis dapat memberikan karya yang lebih baik lagi.

Akhir kata penulis ingin meminta maaf apabila ada kesalahan baik dalam penyusunan laporan maupun yang pernah penulis lakukan sewaktu membuat program Tugas Akhir.

Yogyakarta, 10 Maret 2016

Penulis

©UKDWN

INTISARI

IMPLEMENTASI DYNAMIC MARKOV COMPRESSION (DMC) UNTUK KOMPRESI DATA TEKS UNICODE

Kompresi data merupakan proses pengubahan sekumpulan data menjadi suatu bentuk kode untuk menghemat kebutuhan tempat penyimpanan dan waktu untuk transmisi data. Keuntungan yang didapat dari proses pengompresian data adalah penghematan penggunaan tempat penyimpanan data seperti *harddisk* dan *flashdisk*, serta penghematan *bandwith* pada proses pengiriman data.

Berdasarkan permasalahan tersebut, penulis mengimplementasikan sistem yang menggunakan algoritma Dynamic Markov Compression (DMC) untuk kompresi data teks *unicode*. Penggunaan data teks *unicode* ini dikarenakan beberapa penelitian yang pernah dilakukan hanya menggunakan data uji teks yang memuat karakter ASCII. Karakter ASCII memanfaatkan 8 bit untuk menyimpan kode-kode atau simbol, sedangkan *unicode* sendiri dalam pengkodeannya memanfaatkan 32 bit, sehingga dapat menyimpan kode-kode unik seperti kode bahasa yang akan digunakan sebagai data uji.

Penelitian yang dilakukan menunjukkan hasil rasio kompresi berkisar 68 – 69%, hal ini disebabkan karena proses penyimpanan menggunakan kamus indeks dari 0 – 1859 (000000000000 - 11101000011). Selain itu kelebihan dari penyimpanan menggunakan kamus indeks adalah ukuran dan jenis file data uji tidak berpengaruh terhadap rasio kompresi, akan tetapi kekurangannya adalah file yang dapat dikompres hanya dari indeks 0 – 1859 (000000000000 - 11101000011).

Kata kunci : *Dynamic Markov Compression*, kompresi

DAFTAR ISI

HALAMAN JUDUL	
PERNYATAAN KEASLIAN TUGAS AKHIR	iii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PENGESAHAN	v
UCAPAN TERIMA KASIH	vi
INTISARI.....	viii
DAFTAR ISI	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Pertanyaan Penelitian	2
1.3 Batasan Sistem	2
1.4 Tujuan Penelitian	2
1.5 Metode Penelitian.....	2
1.6 Sistematika Penulisan	3
BAB 2 TINJAUAN PUSTAKA.....	5
2.1 Tinjauan Pustaka	5
2.2 Landasan Teori.....	8
2.2.1 Kompresi Data	8
2.2.1.1 Pengertian Kompresi Data.....	8
2.2.1.2 Algoritma <i>Dynamic Markov Compression</i> (DMC)	11
2.2.2 <i>Unicode</i>	15

BAB 3 ANALISIS DAN PERANCANGAN SISTEM	16
3.1 Analisis Kebutuhan Sistem	16
3.1.1 Analisis Kebutuhan Sistem Non Fungsional.....	16
3.1.2 Analisis Kebutuhan Sistem Fungsional	16
3.2 Use Case Diagram	17
3.2.1 Use Case Skenario	17
3.3 Flowchart	19
3.4 Struktur Data	21
3.5 Rancangan Antar Muka.....	22
3.6 Rancangan Pengujian Sistem	24
3.7 Data Uji	25
3.8 Contoh Kasus	26
BAB 4 IMPLEMENTASI DAN ANALISIS	28
4.1 Implementasi Sistem	28
4.1.1 Form Halaman Depan	28
4.1.2 Form Utama	28
4.1.3 Form Kompresi	29
4.1.4 Form Dekompresi	30
4.2 Analisis Sistem	31
4.2.1 Analisis Program	31
4.2.2 Pengujian Teks <i>Unicode</i> Lirik Lagu	33
4.2.3 Pengujian Teks <i>Unicode</i> Puisi.....	37
4.2.4 Pengujian Teks <i>Unicode</i> Cerpen	39
4.3 Pengujian Berdasarkan File Asli dan Dekompresi.....	40
BAB 5 KESIMPULAN DAN SARAN	41
5.1 Kesimpulan	41
5.1 Saran.....	41
Daftar Pustaka	42
Lampiran	

DAFTAR TABEL

Tabel 2.1 Hasil Kompresi dengan Algoritma Huffman.....	6
Tabel 2.2 Hasil Kompresi dengan Algoritma DMC.....	7
Tabel 3.1 Indeks Teks <i>Unicode</i>	21
Tabel 3.2 Data Uji Berdasarkan Jenis dan Bahasa	25
Tabel 4.1 Hasil Pengujian Lirik Lagu <i>Unicode</i>	34
Tabel 4.2 Hasil Pengujian Puisi <i>Unicode</i>	37
Tabel 4.3 Hasil Pengujian Cerpen <i>Unicode</i>	38
Tabel 4.4 Hasil Pengujian File Asli dan Dekompresi	40

DAFTAR GAMBAR

Gambar 2.1 Cara Menghitung Rasio Kompresi	10
Gambar 2.2 Model Awal DMC	11
Gambar 2.3 Grafik Perbandingan Rasio Kompresi Huffman, LZW, dan DMC	12
Gambar 2.4 Sebuah Model yang Diciptakan Metode Oleh DMC	12
Gambar 2.5 Model Markov Sebelum dan Setelah Cloning	13
Gambar 3.1 Use Case Diagram Sistem	17
Gambar 3.2 Flowchart Kompresi <i>Dynamic Markov</i>	19
Gambar 3.3 Flowchart Dekompresi <i>Dynamic Markov</i>	20
Gambar 3.4 Nodes Next 1 dan 0	21
Gambar 3.5 Tampilan Halaman Awal	22
Gambar 3.6 Tampilan Halaman Utama	23
Gambar 3.7 Tampilan Halaman Kompresi	23
Gambar 3.8 Tampilan Halaman Dekompresi	24
Gambar 4.1 Tampilan Form Home	28
Gambar 4.2 Tampilan Form Utama	29
Gambar 4.3 Tampilan Form Kompresi	30
Gambar 4.4 Tampilan Form Dekompresi	31
Gambar 4.5 Hasil Kompresi dan Dekompresi Lirik Lagu	33

DAFTAR LAMPIRAN

Lampiran A : Tabel Hasil Pengujian Lampiran 1

Lampiran B : Source Code Lampiran 2

Lampiran C : Sumber Data Uji Lampiran 3

©UKDW

INTISARI

IMPLEMENTASI DYNAMIC MARKOV COMPRESSION (DMC) UNTUK KOMPRESI DATA TEKS UNICODE

Kompresi data merupakan proses pengubahan sekumpulan data menjadi suatu bentuk kode untuk menghemat kebutuhan tempat penyimpanan dan waktu untuk transmisi data. Keuntungan yang didapat dari proses pengompresian data adalah penghematan penggunaan tempat penyimpanan data seperti *harddisk* dan *flashdisk*, serta penghematan *bandwith* pada proses pengiriman data.

Berdasarkan permasalahan tersebut, penulis mengimplementasikan sistem yang menggunakan algoritma Dynamic Markov Compression (DMC) untuk kompresi data teks *unicode*. Penggunaan data teks *unicode* ini dikarenakan beberapa penelitian yang pernah dilakukan hanya menggunakan data uji teks yang memuat karakter ASCII. Karakter ASCII memanfaatkan 8 bit untuk menyimpan kode-kode atau simbol, sedangkan *unicode* sendiri dalam pengkodeannya memanfaatkan 32 bit, sehingga dapat menyimpan kode-kode unik seperti kode bahasa yang akan digunakan sebagai data uji.

Penelitian yang dilakukan menunjukkan hasil rasio kompresi berkisar 68 – 69%, hal ini disebabkan karena proses penyimpanan menggunakan kamus indeks dari 0 – 1859 (000000000000 - 11101000011). Selain itu kelebihan dari penyimpanan menggunakan kamus indeks adalah ukuran dan jenis file data uji tidak berpengaruh terhadap rasio kompresi, akan tetapi kekurangannya adalah file yang dapat dikompres hanya dari indeks 0 – 1859 (000000000000 - 11101000011).

Kata kunci : *Dynamic Markov Compression*, kompresi

BAB 1

PENDAHULUAN

1. 1 Latar Belakang

Kompresi data saat ini bukanlah hal baru lagi dalam dunia komputer. Kompresi data merupakan proses perubahan sekumpulan data menjadi suatu bentuk kode untuk menghemat kebutuhan tempat penyimpanan dan waktu untuk transmisi data. Keuntungan yang didapat dari proses pengompresian data adalah penghematan penggunaan tempat penyimpanan data seperti *harddisk* dan *flashdisk*, serta penghematan *bandwith* pada proses pengiriman data. Metode kompresi dapat dibagi menjadi dua kelompok, yaitu *lossy compression* dan *lossless compression*.

Metode yang biasa digunakan untuk kompresi data adalah *lossless compression*, yaitu kompresi data yang hasil dekompresinya sama dengan data asli sebelum dikompresi (tidak ada yang dihilangkan selama proses kompresi). Contoh algoritma yang menerapkan metode ini antara lain Huffman Coding, Arithmetic Coding, LZW, dan DMC (Dynamic Markov Compression).

Beberapa penelitian yang telah dilakukan untuk membandingkan algoritma-algoritma tersebut menunjukkan bahwa algoritma DMC memiliki rasio hasil kompresi yang lebih baik dibandingkan dengan algoritma lainnya, namun waktu yang dibutuhkan untuk kompresi lebih lama. Hal ini disebabkan karena proses kompresi DMC dilakukan per bit, tidak seperti metode lain yang melakukan proses kompresi per byte. Lamanya waktu kompresi inilah yang membuat algoritma DMC tidak banyak digunakan. Selain menunjukkan perbandingan mengenai rasio hasil kompresi dan waktu kompresi DMC dengan algoritma lainnya, penelitian tersebut juga menunjukkan bahwa DMC sangat baik digunakan untuk melakukan kompresi *file* teks, source code, dan *file* aplikasi (Linawati & Panggabean, 2004, hlm. 9).

Berdasarkan permasalahan tersebut, penulis ingin mengimplementasikan sistem yang menggunakan algoritma DMC untuk kompresi data teks *unicode*. Penggunaan data teks *unicode* ini dikarenakan beberapa penelitian yang pernah

dilakukan hanya menggunakan data uji teks yang memuat karakter ASCII. Karakter ASCII memanfaatkan 8 bit untuk menyimpan kode-kode atau simbol, sedangkan *unicode* sendiri dalam pengkodeannya memanfaatkan 32 bit, sehingga dapat menyimpan kode-kode unik seperti kode bahasa yang akan digunakan sebagai data uji.

1.2 Pertanyaan Penelitian

Pertanyaan penelitian yang menjadi dasar penulisan tugas akhir ini yaitu :

1. Bagaimana cara mengimplementasikan algoritma Dynamic Markov untuk kompresi data teks *unicode* ?
2. Bagaimana pengaruh variasi teks *unicode* (ukuran dan jenis) terhadap rasio kompresi ?

1.3 Batasan Sistem

Agar tulisan ini tidak menyimpang dari ruang lingkup pembahasan, diperlukan batasan masalah sebagai berikut :

1. Data yang dikompresi berupa *file .txt*
2. *Unicode* yang digunakan adalah simbol bahasa yang terdaftar dalam *Unicode Consortium* versi 8.0 dan yang dapat diproses adalah indeks 0 – 1859 (00000000000 - 11101000011).

Sumber *Unicode Consortium* : <http://www.unicode.org/charts/>

1.4 Tujuan Penelitian

Tujuan dari penulisan tugas akhir ini adalah mengimplementasikan algoritma Dynamic Markov Compression pada data teks *unicode* dan menghasilkan sistem yang dapat melakukan kompresi dan dekompresi data teks *unicode*.

1.5 Metode Penelitian

Tahapan yang dilakukan dalam penelitian ini adalah :

1. Mengumpulkan bahan-bahan referensi

Mengumpulkan dan mempelajari bahan-bahan referensi yang berhubungan dengan kompresi data Dynamic Markov Compression.

2. Analisis Masalah dan Perancangan Sistem

Melakukan analisis masalah yang dimulai dengan identifikasi masalah, memahami kerja sistem yang akan dibuat, menganalisis dan membuat laporan tentang hasil analisis, serta membuat rancangan dan *interface* sistem.

3. Implementasi Sistem

Perancangan sistem diimplementasikan dalam bentuk kode program (*coding*).

4. Pengujian Sistem

Pengujian dilakukan terhadap program kompresi *file* teks *unicode* dan perhitungan rasio kompresi.

5. Dokumentasi Sistem

Penyusunan laporan tugas akhir lengkap dengan analisis yang didapatkan.

1.6 Sistematika Penulisan

Sistematika penulisan skripsi ini terdiri dari 5 bab yaitu :

BAB 1 PENDAHULUAN

Bab ini menjelaskan latar belakang, pertanyaan penelitian, batasan sistem, tujuan penelitian, metode penelitian dan sistematika penulisan.

BAB 2 LANDASAN TEORI

Bab ini menjelaskan landasan teori dari penelitian yang dilakukan. Dalam penelitian ini, teori yang digunakan adalah kompresi data, algoritma kompresi DMC (Dynamic Markov Compression), serta data *unicode* yang menjadi penelitian.

BAB 3 ANALISIS DAN PERANCANGAN SISTEM

Pada bab ini akan dijelaskan mengenai analisis pembuatan sistem, yaitu implementasi DMC yang dimulai dengan identifikasi masalah, penyusunan alur kerja sistem (menggunakan diagram alir), melakukan analisis dan pembuatan laporan, serta pembuatan *interface* (antar muka) sistem.

BAB 4 IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini akan menguraikan hasil pengujian algoritma DMC untuk data teks *unicode* serta rasio hasil kompresi yang didapatkan untuk setiap data uji berisi teks *unicode*.

BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan yang didapatkan dari hasil pengujian terhadap algoritma DMC, serta saran-saran yang diberikan untuk penelitian selanjutnya.

©UKDW

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan implementasi dan analisis yang telah dibuat pada bab sebelumnya, maka dapat disimpulkan sebagai berikut :

1. Ukuran file data uji dan jenis data uji tidak berpengaruh terhadap rasio kompresi.
2. Kelebihan cara penyimpanan menggunakan indeks adalah rasio kompresi hanya sekitar 68 – 69%, Selain itu kekurangannya adalah tidak bisa mengkompresi file diluar indeks yang sudah ditentukan dan membutuhkan waktu yang cukup lama jika memproses file berukuran besar.

5.2 Saran

Penggunaan indeks dalam program Dynamic Markov Compression ini belum optimal karena indeks yang digunakan masih terbatas. Untuk penelitian selanjutnya, saran yang diberikan penulis adalah pengembangan cara penyimpanan menggunakan indeks atau dengan cara penyimpanan lainnya yang lebih efektif sehingga dapat menghasilkan rasio kompresi yang lebih baik lagi.

Daftar Pustaka

- Anggriani, M. (2011). *Perbandingan Metode Kompresi Huffman dan Dynamic Markov Compression (DMC)*. Bandung: Universitas Pendidikan Indonesia.
- Azizah, N. (2006). *Kompresi Format Audio File WAV dengan Menggunakan Metode Lossless*. Bandung: Universitas Komputer Indonesia.
- Haryanto, R.I. (2008). *Kompresi Data dengan Algoritma Huffman dan Perbandingannya dengan Algoritma LZW dan DMC*. Bandung: Institut Teknologi Bandung.
- Linawati & Panggabean, H.P. (2004). *Perbandingan Kinerja Algoritma Kompresi Huffman, LZW, dan DMC Pada Berbagai Tipe File*. Bandung: Universitas Katolik Parahyangan.
- Salomon, D. *Data Compression The Complete Reference*. 4th ed. New York : Springer, 2007.
- Saputro, N.C, Suwarno, S., & Santosa, R.G. (2009). *Implementasi dan Analisis Perbandingan Antara Pengkodean LZ78 dan Shannon Fano Pada Kompresi Data Teks*. Yogyakarta: Universitas Kristen Duta Wacana.
- Sayood, Khalid. (2006). *Introduction to Data Compression*. San Francisco : Elsevier
- Taryani, Y. (2008). *Kompresi Citra Menggunakan Algoritma Dynamic Markov Compression (DMC)*. Bandung : Universitas Komputer Indonesia