JURUSAN TEKNIK INFORMATIKA

FAKULTAS TEKNOLOGI INFORMASI

INSITUT TEKNOLOGI SEPULUH NOPEMBER

# **USULAN TUGAS AKHIR**

# IDENTITAS PENGUSUL

Nama : **Melvin Simon Zepta Marbun**

NRP : **5109100035**

Dosen Wali : **Isye Arieshanti, S.Kom , S.Phil**

# JUDUL TUGAS AKHIR

**Implementasi Metode Steganografi Berbasis Kompresi Teks pada Surat Elektronik**

***Implementation Steganography Method Based on Text Compression in Electronic Mail***

# URAIAN SINGKAT

Dalam tugas akhir ini, kapasitas dan isu keamanan dari steganografi teks memerlukan sebuah pendekatan yang baru. Untuk tujuan ini, maka diusulkan sebuah metode steganografi teks yang menggunakan kompresi data. Karena data yang digunakan berupa teks, maka algoritma kompresi data yang digunakan haruslah *lossless*.

Algoritma data kompresi data Huffmann atau LZW dipilih karena memiliki rasio kompresi yang baik. Metode ini dirancang dengan banyak *stego key* dan menggunakan koding berbasis kombinatorial dengan tujuan untuk meningkatkan keamanan. Pesan rahasia disembunyikan di dalam teks yang dipilih dari teks dasar yang umum dijumpai dalam kehidupan sehari-hari.

Email dipilih sebagai sarana komunikasi antara dua pihak, jadi *stego cover* disusun sedemikian rupa sehingga seperti email dengan beberapa alamat tujuan, yang biasa ditemui saat meneruskan email. Berdasarkan hasil percobaan yang dilakukan, penghitungan kapasitas rata-rata mencapai 7,962% pada pesan rahasia yang berjumlah 300 karakter. Perbandingan antara skema yang diusulkan dengan beberapa metode kontemporer yang sudah ada sebelumnya juga telah dilakukan. Hasil percobaan menunjukan terjadi penambahan kapasitas yang signifikan saat menggunakan skema yang diusulkan.

# PENDAHULUAN

## Latar Belakang

Perkembangan teknologi jaringan dan perangkat digital membuat pengiriman data digital multimedia menjadi cepat dan mudah. Namun pendistribusian data digital melalui jaringan publik seperti internet tidak cukup aman karena sering terjadi pelanggaran hak cipta, pemalsuan, pembajakan, dan penipuan. Oleh karena itu, metode untuk melindungi data digital terutama data yang sensitif sangat diperlukan[[1](#Cha10)].

Meskipun penggunaan dokumen elektronik sudah menjadi hal yang umum, tidak banyak orang dapat mengenali dokumen-dokumen yang berisi data tersembunyi. Dikatakan tersembunyi karena data biasanya terdapat di dalam sebuah berkas, tetapi sulit untuk diidentifikasi keberadaannya. Data tersembunyi dapat diklasifikasikan menjadi dua jenis, yang dibuat secara otomatis oleh aplikasi dan yang dibuat dan disembunyikan oleh seseorang untuk tujuan tertentu [[2](#Par09)].

Pada umumnya, *watermarking* dan steganografi digunakan untuk menyembunyikan informasi [[1](#Cha10)]. *Watermarking* adalah penanda yang disisipkan di dalam sebuah data teks, suara maupun gambar yang digunakan untuk mengidentifikasi kepemilikan atau hak cipta dari data tersebut. Sebaliknya,steganografi umumnya digunakan pada komunikasi rahasia [[1](#Cha10)].

Kebutuhan utama dari sistem penyembunyian data diketahui sebagai berikut: *imperceptibility* (keberadaan pesan tidak mudah dideteksi oleh indera manusia), kehandalan, dan kapasitas. Seperti yang ditunjukan oleh gambar 1, masing-masing kebutuhan adalah sudut dari segitiga dalam sistem penyembunyian data dan selalu ada pertukaran antar kebutuhan-kebutuhan yang kontradiktif ini.

*imperceptibility*

kapasitas kehandalan

Gambar

Kapasitas berkaitan dengan jumlah *bit* data yang dapat disembunyikan pada *cover media*. *Cover media* adalah media yang digunakan sebagai perantara dalam menyisipkan pesan rahasia. *Imperceptibility* berhubungan dengan kemampuan dari penyadap untuk mencari tahu informasi yang tersembunyi dengan mudah. Kehandalan berhubungan dengan ketahanan dari kemungkinan pengubahan atau perusakan data yang tidak kelihatan [[3](#Gut07)]. Dalam steganografi, keberadaan pesan disamarkan. *Cover media* yang digunakan sebagai tempat untuk menyisipkan data dapat berupa berkas suara atau video, email atau gambar digital [[6](#Zak11)].

Dalam tugas akhir ini, masalah kapasitas dan keamanan diatasi dengan mengusulkan sebuah pendekatan steganography teks berbasis kompresi, dengan tujuan untuk mendapatkan penambahan rasio yang signifikan pada jumlah data rahasia yang akan disembunyikan di dalam *cover media* sembari meningkatkan kompleksitas prosedur ekstraksi dari data rahasia. Karenanya, data rahasia disisipkan pada teks yang dipilih dari teks dasar. Teks dasar adalah teks sudah pernah ada sebelumnya, seperti teks pemberitahuan (notifikasi) dan abstraksi dari artikel yang dapat digunakan untuk diskusi kelompok. Saat proses penyisipan, keaslian teks dijaga dengan menyamarkan bagian informasi rahasia. Email dipilih sebagai sarana komunikasi karena proses komunikasinya terjadi antara dua pihak. Oleh sebab itu, *stego cover* disusun sedemikian rupa sehingga berupa email dengan beberapa alamat tujuan, seperti yang biasa ditemui saat meneruskan email. *Stego cover* adalah teks yang digunakan untuk menyisipkan pesan rahasia. Saat menyusun *stego cover*, digunakan daftar alamat email yang sebelumnya telah disusun. Diagram yang menjelaskan proses ditunjukan oleh Gambar 2.



Gambar . Diagram alir metode yang diusulkan

Untuk menambah kapasitas, akan digunakan teknik kompresi data. Dalam tugas akhir ini akan mengggunakan teknik kompresi data lossless karena data yang digunakan berupa teks. Dan pada penerapan teknik kompresi data lossless, dipilih algoritma LZW atau algoritma *Huffman Coding* karena rasio kompresinya yang baik.

Untuk meningkatkan keamanan, akan digunakan beberapa *stego keys*. *Stego keys* adalah kunci yang digunakan oleh pengirim pada saat melakukan penyisipan pesan dan diperlukan oleh penerima untuk melakukan prosedur ekstraksi. *Stego keys* dapat diklasifikasikan menjadi dua bagian berdasarkan tugas masing-masing. Bagian pertama adalah *stego keys* yang dibuat saat fase penyisipan dan yang lainnya adalah *global stego key* yang bersama-sama digunakan oleh pengirim dan penerima sebelumnya. *Global stego key* adalah kunci yang langsung diketahui oleh penerima saat menerima pesan, tanpa melakukan prosedur apapun untuk mendapatkannya.

## Rumusan Masalah

1. Bagaimana cara menambah ukuran data yang disembunyikan pada *stego cover*.
2. Bagaimana menyamarkan *stego cover* sehingga tidak mudah dikenali oleh pihak-pihak lain yang tidak memiliki kepentingan dalam proses komunikasi yang dilakukan.
3. Bagaimana meningkatkan sekuritas dalam proses steganografi.

## Batasan Masalah

1. Aplikasi yang akan dibuat berdasarkan implementasi terhadap paper milik Esra Satir dan Hakan Isik yang berjudul “A Huffman Compression Based Text Steganography Method”.
2. Algoritma yang digunakan berasal dari algoritma yang ada di dalam jurnal ilmiah yang dijadikan acuan.
3. Aplikasi yang akan dibangun tidak mengurusi mekanisme pengiriman email.

## Tujuan dan Manfaat

1. Proses kompresi dapat menjadi sarana untuk meningkatkan sekuritas dalam proses steganografi dan menambah ukuran data dalam pesan rahasia.
2. Memanfaatkan email sebagai media pengirim pesan secara umum untuk mengirim pesan rahasia.

# TINJUAN PUSTAKA

Kompresi data

Tujuan proses kompresi data adalah untuk mengurangi redundansi dari sebuah deskripsi data [[6](#GGa02)]. Pada umumnya, algoritma kompresi data dibedakan antara *lossless* atau *lossy*. Kompresi data *lossless* adalah transformasi dari representasi himpunan data asli sehingga memungkinkan untuk menghasilkan himpunan data yang sama seperti aslinya dengan melakukan proses dekompresi. Kompresi *lossless* digunakan saat berkas asli dan berkas hasil dekompresi harus sama persis (contoh penggunaan dalam kompresi berkas teks, kode yang *executable*, berkas dari pengolah kata, dll). Kompresi data *lossy* adalah transformasi dari representasi himpunan data asli yang tidak mungkin untuk menghasilkan kembali himpunan data yang sama persis dengan aslinya saat didekompresi, tetapi menghasilkan representasi data yang mendekati aslinya. Kompresi *lossy* digunakan di Internet dan secara khusus pada media *streaming* dan aplikasi telepon.

Steganografi

Steganografi adalah seni menulis data rahasia dengan cara tertentu sehingga tidak seorangpun tahu keberadaan data rahasia tersebut, kecuali penerima yang dituju [[6](#ADe09)]. Pada umumnya, untuk mengekstraksi data rahasia dari perantara yang digunakan diperlukan satu atau beberapa kunci (*key*).

# METODOLOGY

## Fase Penyisipan



Gambar . Diagram alir proses penyisipan pesan

Sebelum menjelaskan prosedur penyisipan , berikut adalah variabel yang akan digunakan :

= Pesan rahasia (Secret Message)

= Teks dasar (text base)

= Teks yang didapat dari teks dasar

= Jarak relatif (relative distance)

= Himpunan dari ekstensi alamat email

= Matriks Jarak relatif (Matrix of relative distance)

= Matriks

= Matriks Jarak relatif yang dibuat lagi (Matrix of reconstructed relative distances)

= Global Stego Key

= Himpunan dari alamat email yang dipilih dan dimodifikasi

S adalah himpunan karakter dari pesan rahasia. *Text* merepresentasikan sebuah teks dalam teks dasar dan mengandung karakter dari teks. *T* adalah representasi matriks *Text* dalam teks dasar. 3000 adalah jumlah karakter maksimum dari *Text* dalam teks dasar. Jika jumlah karakter dalam *Text* kurang dari 3000; elemen yang berhubungan dengan T ditentukan sebagai 0.

Langkah 1

S mengandung karakter dari pesan rahasia dan *Text* menyimpan karakter dari teks. Dicari kondisi dimana a = b. Karenanya , adalah vektor yang elemen-elemennya adalah selisih antara indeks a dan b saat pemetaan karakter. Operasi ini dapat digambarkan dengan ini :

Karena karakter adalah kode ASCII,maka:

Diasumsikan bahwa . Dalam kasus ini, nilai yang kita perlu untuk adalah indeks dari b, anggap saja 2. Dalam langkah kedua dianggap elemen berikutnya dari S dan Text :

Diasumsikan lagi bahwa . Dalam kasus ini , nilai yang kita perlu untuk adalah perbedaan antara indeks sekarang dari b(4) dan indeks sebelumnya dari b(2),berarti 4-2=2. Operasi ini terjadi secara iteratif melalui *Text* sampai akhir dari pesan rahasia, untuk membentuk .

Langkah 2

dibentuk untuk setiap *Text* dalam *T* . Kemudia setiap nilai disimpan untuk membentuk .

Langkah 3

Melakukan pemeriksaan apakah elemen dari D melebihi 26. Jika benar, nilai E dan R ditentukan oleh persamaan berikut;

(1)

(2)

Jika *d* tidak melebihi 26, perhatikan bahwa nilai *e* yang berhubungan akan menjadi 0 dan nilai *r* yang berhubungan menjadi sama dengan *d*.

Langkah 4

Memperkirakan jumlah pengulangan pada pola ganda untuk setiap baris dari *R* yang sudah dibangun dalam langkah sebelumnya dan mendapatkan :

Diambil nilai maksimum *p* dalam *P*. Kemudian, diambil baris dari *E* dan *R* yang berhubungan dengan indeks baris dari penghitungan nilai *p* maksimal di *P*. Baris-baris berikut dinyatakan sebagai (vektor kelebihan untuk yang dibangun ulang) dan ( yang dibangun ulang). Sementara itu, di dalam *T*, *Text* dipilih sebagai teks pembungkus (*T*\*) yang behubungan dengan indeks baris dari nilai *p* maksimal.

Langkah 5

Dalam langkah ini, dikompres dengan *Huffman Coding* atau LZW. Diupayakan hasil kompresi akan berupa bilangan biner.

Langkah 6

Mendapatkan 13 alamat email dari hasil penghitungan yang dilakukan pada langkah 5.

Langkah 7

Hasil dari ke dalam kelompok masing-masing berisi 12 *bit*. Dalam setiap grup, 9 *bit* pertama() akan digunakan untuk membuat *userside* dari alamat email. Sisa 3 *bit* akan digunakan untuk memodifikasi ekstensi alamat email (sebagai contoh gmail.com). Dengan melakukan operasi ini ,didapat dua bilangan bulat:

(3)

(4)

Bilangan bulat ini kemudian digunakan untuk memilih alamat email dari . ; *global stego key* adalah sebuah himpunan yang mengandung daftar email yang dihasilkan sebelumnya. dan dikonversikan menjadi huruf dengan menggunakan *Latin Square*. Kemudian kedua huruf ini akan dipetakan ke satu alamat email dengan menggunakan . Bagian 3 *bit* terakhir dari masing-masing kelompok dinamakan .

(5)

Variabel akan digunakan untuk memodifikasi ekstensi alamat email dengan menggunakan *A*. Ekstensi alamat email ditentukan dengan menggunakan indeks biner dari elemen dalam *A* (Modifikasi ini digunakan sebagai *stego key* yang adalah bagian dari ).

Langkah 8

Modifikasi alamat email terpilih dengan tujuan untuk melengkapi pembangunan dari himpunan dengan menggunakan . Pemodifikasian ini dilakukan dengan menambahkan kelebihan angka pada jumlah alamat email terpilih sebelum karakter “@” dan juga digunakan sebagai stego key. Karena tidak ada satupun aturan atau pembatasan dalam membentuk alamat email, angka-angka yang berlebih terlihat seperti bagian alami dari alamat email. Jadi adalah himpunan yang mengandung alamat email yang terpilih dan termodifikasi. Banyak elemen dari himpunan disusun sebagai *stego key*. Disamping itu, diasumsikan juga bahwa .

Langkah 9

*Stego cover* dibangun dengan menggunakan *T*\* sebagai teks perantara dan himpunan . Media untuk menyembunyikan pesan rahasia disusun sedemikian rupa sehingga seperti email dengan beberapa alamat tujuan, yang biasa ditemui saat meneruskan email. Perhatikan bahwa alamat email yang pasti digunakan sebagai *stego key* menurut angka yang berlebih. Dan dengan 3 *bit* terakhir (*z*) yang menentukan ekstensi alamat email. Bagaimanapun, tidak mungkin untuk mengetahui yang mana yang merupakan *stego key* tanpa memiliki . hanya diketahui oleh pengirim dan penerima sebelumnya.

## Pembuatan dan penggunaan *stego key*

Metode yang diusulkan menggunakan *stego key* untuk meningkatkan pengamanan. Berdasarkan tugasnya, *stego key* dapat diklasifikasikan ke dalam dua jenis. Yang pertama adalah himpunan alamat email terpilih dan termodifikasi;. Hal ini dilakukan dengan menyisipkan informasi yang *overflow* sebelum karakter “@” dan memilih ekstensi alamat email menurut *z* melalui *A*. Pembuatan dilakukan pada fase penyisipan. digunakan untuk menyisipkan informasi yang menunjukan posisi yang benar dari karakter tersembunyi. *Stego key* terlihat sebagai bagian alami dari *stego cover* yang sudah disusun sedemikian rupa sehingga seperti email dengan beberapa alamat tujuan, yang biasa ditemui saat meneruskan email. Surat tidak akan terkirim pada alamat email yang terpilih dan ditentukan, namun hanya dikirim pada penerima utama saja. *Stego key* yang lainnya adalah *global stego key*;. Tujuan dari penggunaan global stego key adalah untuk mengetahui posisi pesan rahasia yang benar yang disisipkan di alamat email yang dimodifikasi.

## Fase ekstraksi

Langkah 1

Setiap elemen dalam *stego cover* dilakukan perbandingan antara elemen dari dengan elemen . Tujuan dari operasi ini adalah untuk mencari tahu apakah ada perbedaan antara alamat-alamat email yang dibandingkan.

Langkah 2

Memeriksa elemen dari dengan dimulai pada elemen ke-14 (karena elemen 13 pertama termasuk nilai sebagai hasil kompresi). Dua karakter pertama dari masing-masing alamat email diubah menjadi angka dengan menggunakan *Latin Square*. Lalu didapatkan nilai dan . Nilai *z* didapat dengan memeriksa ekstensi alamat email dengan menggunakan . Kemudian dilakukan penghitungan nilai dan untuk masing-masing kelompok dari 12 *bit* dengan menggunakan persamaan berikut :

(6)

(7)

Dengan menggabungkan nilai dan , didapat dalam bentuk *bit stream* dan siap untuk didekompresi.

Langkah 3

didekompresi untuk mendapat . Nilai didapat dengan cara mengubah dua karakter pertama dari masing-masing alamat email ke bentuk angka melalui *Latin Square* dan dilakukan pada setiap 13 elemen.

Langkah 4

Memperkirakan awal dari dengan menggunakan dan . Jika elemen dan dinyatakan dalam dan (elemen dari dinotasikan sebagai ), maka :

Langkah 5

Dengan menggunakan elemen dalam , dapat diambil elemen dari melalui \* di dalam *stego cover*. Dengan memperhatikan dalam elemen \*, dapat dideteksi nomer indeks saat pemetaan karakter dibentuk (dimana ). Lalu lakukan proses ekstraksi elemen dari untuk setiap elemen di dan gabungkan elemen yang diekstraksi untuk mendapatkan .

# JADWAL KEGIATAN TUGAS AKHIR

Tugas Akhir ini diharapkan dapat dikerjakan menurut jadwal sebagai berikut:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Kegiatan** | **Bulan** | | | | | | | | | | | | | | | | | | | | |
| **Maret 2013** | | | | | **April 2013** | | | | **Mei 2013** | | | | **Juni 2013** | | | | **Juli 2013** | | | |
| 1. | Penyusunan Proposal Tugas Akhir |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2. | Studi Literatur |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3. | Analisa dan Perancangan |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4. | Implementasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5. | Pengujian dan Evaluasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 6. | Penyusunan Buku Tugas Akhir |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

DAFTAR\_PUSTAKAx

|  |  |
| --- | --- |
| [1] | C. Chang and T.D Kieu, "A reversible data hiding scheme using complementary embedding strategy," *Information Science 180*, no. 16, pp. 3045-3058, 2010. |
| [2] | J. Park and S. Lee, "Forensic investigation of Microsoft PowerPoint files," *Digital Investigation*, vol. 6, no. 1-2, pp. 16-24, 2009. |
| [3] | A. Gutub and M. Fattani, "A novel arabic text steganography method using letter points and extensions," in *WASET International Conference on Computer, Information and System Science and Engineering (ICCISSE)*, Viena, Austria, 2007, pp. 28-31. |
| [4] | N Zaker and A Hamzeh, "A novel steganalysis for TPVD steganographic method based on differences of pixel difference histogram," *Multimedia Tool Appl*, no. 10, pp. 1007-1104, 2011. |
| [5] | Jun L, Tong W, and Daxin l, "Research on ordinal properties in combinatorics coding method," *J Comput*, vol. VI, no. 1, pp. 51-58, 2011. |
| [6] | Desoky A, "Listega : list-based steganography metodology," *Int J Inf Secur*, vol. VIII, no. 4, pp. 247-261, 2009. |
| [7] | Easton T and Gary Parker R, "On completing latin squares," *Discrete Appl Math 113*, vol. II, no. 3, pp. 167-181, 2001. |
| [8] | Galambos G and Bekesi J, *Data Compression: Theory and Techniques*, Teacher's Training College Department of Informatics, Ed. USA: Elsevier Science, 2002, vol. I. |
| [9] | H. Sajedi and M. Jamzad, "BSS:boosted steganography scheme with cover image preprocesssing," *Expert System with Applications*, vol. 37, no. 12, pp. 7703-7710, 2010. |
| [10] | F. Al-Haidari, A. Gutub, K. Al-Kahsah, and J. Hamodi, "Improving security and capacity for arabic text steganography using "Kashida" extensions," in *The 7th ACS/IEEE International Conference on Computer Systems and Applications (AICSSA - 2009)*, Rabat, Morroco, 2009, pp. 396-399. |
| [11] | M.A. Aabed, S.M. Awaideh, M.E. Abdul-Rahman, and A. Gutub, "Arabic diacritics based steganography," in *IEEE International Conference on Signal Processing and Communications (ICSPC 2007)*, Dubai, UAE, 2007, pp. 756-759. |
| [12] | Bailey K and Curran K, "An evaluation of image based steganography method using visual inspection and automated detection techniques," *Multimed Tool Appl*, no. 1, pp. 55-58, 2006. |

x