PENDETEKSIAN STEGANOGRAFI DALAM MEDIA GAMBAR BERFORMAT JPEG BESERTA ANALISISNYA

Ibnu Alam – NIM : 13506024

Program Studi Teknik Informatika, Institut Teknologi Bandung
Jl. Ganesha 10, Bandung
E-mail : if16024@students.if.itb.ac.id

Abstrak

Makalah ini membahas tentang teknik pendeteksian data yang disembunyikan secara steganografi ke dalam file gambar berformat JPEG. Dalam analisisnya, diambil program *StegDetect* sebagai program pendeteksi, berikut penjelasan algoritma diambil dari *source code*nya.

Steganografi (*Steganography*) merupakan sebuah algoritma untuk membuat orang yang tidak berkepentingan tidak dapat mengetahui isi suatu data seperti kriptografi namun keberadaan data atau pesan tersebut tetap ada, tetapi hanya pesan sebenarnya yang disamarkan. Steganografi (*Steganography*) menjadi cukup populer karena data hasil enkripsinya tidak menarik perhatian karena data yang dikirim masih dapat dilihat dan dibaca secara wajar oleh siapapun, namun pesan sebenarnya yang ada dibalik pesan tersebut yang tidak dapat diketahui. Data yang dimaksudkan adalah *Digital Image*. Implementasi *Steganography* dalam makalah ini meliputi tiga sistem operasi yaitu sistem operasi *Jsteg*, *JPHide, dan Outguess*.

Jsteg adalah sistem steganografi yang sudah cukup lama diperkenalkan secara luas di internet. Algoritma *Jsteg* merupakan algoritma penyamaran *Digital Image* yang paling mudah untuk diketahui dan di dekripsikan. Hal ini disebabkan pesan yang disisipkan berada pada awal data.
Outguess merupakan algoritma yang lebih sulit untuk diketahui dan di dekripsikan bila dibandingkan dengan *Jsteg*. Sistem enkripsi *Outguess* tersedia di internet dalam bentuk *Source Code* untuk UNIX. Penyisipan pesan yang menyebar pada seluruh data cukup membuat *Outguess* sulit untuk diketahui.
JPHide merupakan program yang mampu menyembunyikan sebuah file dalam gambar berformat JPEG yang mana penyembunyian file tersebut dibuat sedemikian rupa sehingga mustahil untuk mengatakan bahwa file pembawa mengandung sebuah file yang disembunyikan.

Pembahasan utama dalam makalah ini adalah *steganalisis,* yaitu seni menemukan keberadaan data rahasia. Steganografi bisa dilacak bila sebuah gambar asli yang belum disentuh dibandingkan dengan gambar yang sudah mengalami proses steganografi. Walaupun kemungkinan data yang disembunyikan tidak dapat diperoleh dengan steganalisis, hal ini masih jauh lebih baik daripada tidak mengetahui sama sekali adanya data yang tersembunyi.

Pendahuluan

Informasi adalah suatu hal yang bernilai tinggi, bahkan menjadi sangat krusial dalam kondisi dan situasi yang benar-benar penting. Terkadang informasi itu perlu dilindungi sehingga tidak sampai pada pihak-pihak yang tidak diinginkan, tetapi pada saat yang sama perlu untuk disampaikan kepada pihak yang lain. Timbulah suatu perkara untuk menyampaikan informasi itu sedemikian rupa sehingga hanya pihak yang telah ditentukan yang dapat menerima informasi tersebut. Salah satu cara yang telah digunakan secara efektif, bahkan sudah dari zaman dahulu adalah dengan penyembunyian pesan atau steganografi.
Steganografi adalah ilmu dan seni untuk menyembunyikan informasi. Implementasinya dalam informasi digital adalah dengan menyisipkan sebuah informasi ke dalam sebuah file media lain. Secara kasat mata file yang menyimpan informasi rahasia ini tidak bisa dibedakan dengan file biasa lainnya sehingga dapat luput dari segala bentuk kecurigaan pihak lain.
Di lain pihak, para pihak yang ingin mendapatkan informasi yang dirahasiakan darinya itu juga memiliki usaha untuk mencari file mana yang menyimpan informasi tersebut. Berhubung file yang dipertanyakan disebarluaskan secara bebas, sulit untuk menentukan apakah suatu informasi tersembunyi dalam file tersebut, kecuali dengan memeriksa setiap file satu persatu.
Satu bentuk media yang sering digunakan sebagai media untuk menyembunyikan informasi adalah file gambar. Jenis file gambar yang umum dipakai di

jaringan internet adalah JPEG. Mengamati hal tersebut, penulis bermaksud menjelaskan langkah-langkah pendeteksian apakah adanya penyembunyian informasi dalam gambar berformat JPEG. Mengingat JPEG sendiri adalah format kompresi gambar yang memiliki karakteristik tersendiri, penulis juga menyertakan penjelasan dan analisis untuk melengkapi tulisan ini.

Kompresi berformat JPEG

Gambar dalam format JPEG umumnya dikompresi dengan menggunakan JFIF encoding:

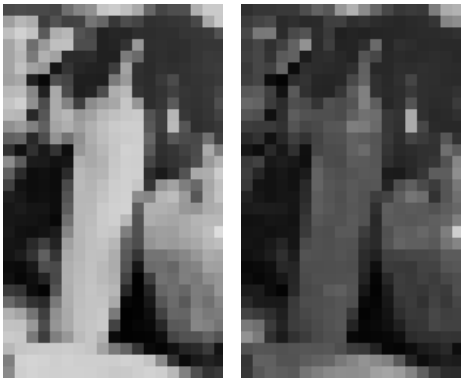
1. Representasi warna dalam gambar diubah dari RGB (Red, Green, Blue) ke YCbCr, yaitu satu komponen *brightness*, luma (Y), dan dua komponen warna, chroma (Cb, Cr).

Ilustrasi:

Gambar dalam format RGB

Diubah ke format YcbCr:

Faktor Y (Intensitas)



Faktor Cb (biru/kuning) dan Cr (merah/hijau)

2. Resolusi data chroma diturunkan (*downsampling*), biasanya dengan faktor pembagian 2. Hal ini dikarenakan mata manusia lebih peka terhadap detail *brightness*daripada detail warna.

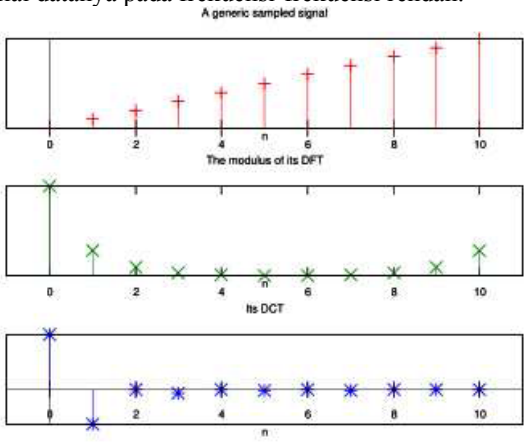
3. Gambar dibagi ke dalam blok-blok 8x8 piksel. Tiap blok akan melalui proses transformasi Discrete Cosine Transform (DCT). DCT menghasilkan spectrum frekuensi spatial dari data Y, Cb, dan Cr. JPEG menggunakan DCT-II dalam prosesnya.

Rumus DCT-II:

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n+\frac{1}{2}\right)k\right]$$

$$k = 0, \ldots, N-1.$$

DCT digunakan untuk menyusutkan data Y, Cb, dan Cr menjadi bentuk yang lebih *compact.* Hal ini disebabkan DCT memiliki sifat menempatkan nilai-nilai datanya pada frekuensi-frekuensi rendah.



Gambar 3.3-1

DCT-II yang bersifat dua dimensi NxN blok (JPEG membagi gambar menjadi blok-blok 8 x 8)

dikomputasi dan hasilnya dikuantisasi dan dikode entropi. DCT-II diaplikasikan pada tiap baris dan kolom dari tiap blok. Hasilnya adalah koefisien transformasi 8 x 8 dimana elemen (0,0) (kiri-atas) adalah komponen DC (*zero frequency*) dan data lain pada indeks vertikal dan horizontal yang lebih besar merepresentasikan frekuensi spatial horizontal dan vertikal yang lebih tinggi.

4. Amplitudo dari frekuensi komponen-komponen tesebut dikuantisasi. Mata manusia lebih sensitif terhadap variasi kecil warna atau *brightness* dalam lingkup area yang luas daripada variasi *brightness* padafrekuensi tinggi. Oleh karena itu, nilai dari komponen yang berfrekuensi tinggi disimpan dalam akurasi yang lebih rendah daripada komponen yang berfrekuensi rendah. Dalam kasus *encoding* dengan *settings* kualitas yang sangat rendah, komponen frekuensi tinggi akan dibuang seluruhnya.

5. Hasil dari setiap blok 8x8 tersebut akan dikompresi lebih lanjut dengan algoritma *loss-less* yang merupakan variasi dari Huffman *encoding.*

Steganografi

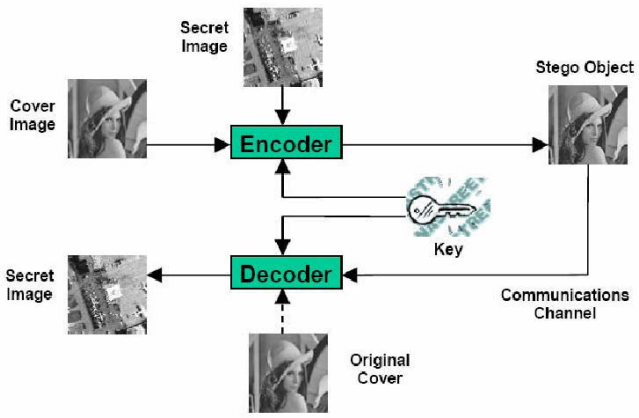


Diagram Steganografi Sederhana

Steganografi adalah ilmu dan seni untuk mengirimkan pesan sedemikian rupa sehingga orang yang tidak berkepentingan tidak mengetahui adanya pengiriman pesan tersebut. Berbeda dengan kriptografi yang tidak menyembunyikan penyampaian pesan tetapi isi pesannya dibuat tidak bisa dimengerti, dalam steganografi, pesan disampaikan dalam bentuk lain seperti gambar, artikel, daftar, atau lainnya. Tentunya pesan rahasianya tidak terlihat. Kelebihannya dibanding kriptografi adalah steganografi tidak mengundang kecurigaan kepada pengirim, penerima, dan perantaranya. Pesan yang disandikan dalam kriptografi, sebagaimanapun sulit dipecahkannya, akan menimbulkan kecurigaan, bahkan bisa membawa dampak lebih, misalnya pada negara yang melarang praktek kriptografi.

Manfaat Steganografi

Sebagai *tools* keamanan, steganografi dapat digunakan untuk berbagai tujuan. Beberapa hal baik termasuk untuk pemberian watermark sebagai perlindungan hak cipta. Steganografi juga bisa digunakan sebagai generator hash satu arah (memilih input dengan panjang variabel dan membuat output panjang statis untuk mengecek apakah input tidak pernah diubah). Steganografi juga bisa digunakan untuk menyisipkan catatan pada gambar-gambar di internet. Lalu, ia bisa menjaga kerahasiaan data-data berharga, melindungi dari kemungkinan sabotase, pencurian, atau akses terlarang. Sayangnya steganografi juga bisa digunakan untuk hal-hal yang tidak baik. Misalnya jika ada orang yang ingin mencuri data, dia bisa menyisipkannya ke file lain dan mengirimkannya lewat email biasa.

Tentunya guna awal sebagai pengirim pesan rahasia masih berlaku. Steganografi digunakandalam komunikasi elektronik sebagai kode dalam transport layer, misal UDP dan untuk mp3.

Pesan steganografi (plaintext) biasanya dienkripsi lalu file pembawa (covertext) dimodifikasi untuk menampung pesan tersebut, menjadi stegotext. Pihak penerima harus tahu metode yang digunakan untuk memperoleh kembali isi pesannya dengan cara mendekripsikannya.

Teknik-Teknik Steganografi

Makalah ini membahas analisis dari tiga algoritma steganografi yang direlisasikan dalam program-program yang sudah banyak diaplikasikan di internet, yaitu JPHide & JPSeek, Outguess, dan JSteg.

a. JPHide& JPSeek

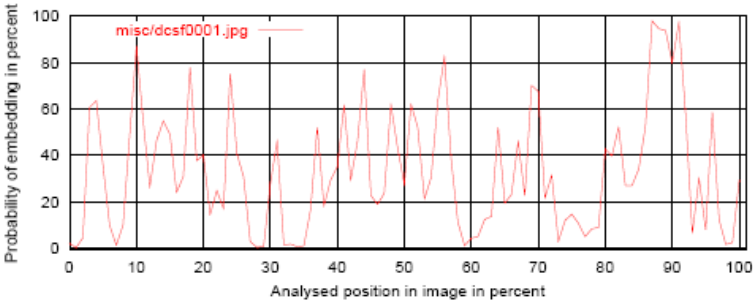
Program ini merupakan program yang mampu menyembunyikan sebuah file dalam gambar berformat JPEG. Tentunya banyak program yang serupa, tetapi program ini memiliki kelebihan, yaitu penyembunyian file tersebut dibuat sedemikian rupa

sehingga mustahil untuk mengatakan bahwa file pembawa mengandung sebuah file yang disembunyikan.

Misalnya sebuah gambar biasa dimasukkan sebuah file dengan rasio rendah(di bawah 5%) dengan catatan file asli tidak disebarkan (mencegah analis membuat perbandingan file), tidak mungkin menyimpulkan dengan cukup yakin bahwa file tersebut mengandung data tersembunyi. Tetapi, dengan menaikkan rasio data yang disembunyikan, nilai-nilai koefisien pada JPEG akan bergeser dari daerah normal sehingga dapat menimbulkan kecurigaan. Di atas 15%, perubahan mulai terlihat kasat mata. Oleh karena itu, file pembawa berupa gambar dengan banyak detail lebih baik daripada gambar yang lebih kosong. Perubahan pada gambar langit cerah di atas salju putih lebih mudah terlihat daripada gambar air terjun di tengah hutan.

b. Outguess

Outguess adalah perangkat steganografi universal yang memungkinkan penyisipan informasi rahasia ke dalam sumber data apapun (tidak harus gambar) yang memiliki bit redundan. Program ini memerlukan handler yang spesifik terhadap data yang akan mengekstrak bit-bit redundan dan menuliskannya kembali setelah diadakan modifikasi.



Pseudocode :

Input: Image I, Message ~m, Key k
Output: Image J
Seed PRNG with k
for each bit b of ~m
  c := pseudo-random DCT coefficient from I
  while c = 0 or c = 1,
    c := pseudo-random DCT coefficient from I
  end while
  c := c mod 2 + b
  replace coefficient in I by c
end for

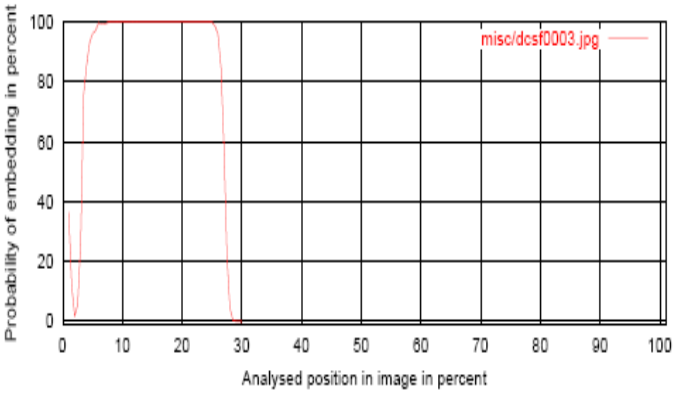
Steganografi menyembunyikan adanya komunikasi. Sistem steganografi klasik bergantung pada menjaga kerahasiaan sistem enkoding, sedangkan steganografi modern hanya dapat dideteksi jika informasi rahasia berupa kunci rahasia tertangkap. Karena sifat sistemnya yang menyusup ke dalam, steganografi meninggalkan jejak dalam karakteristik dari medium penyimpanannya. Jejak tersebut dalam file JPEG adalah distribusi bit dalam daerah Terdeteksinya jejak tersebut memberikan peluang bagi pihak ketiga untuk menemukan file mana yang sudah diubah, menandakan bahwa ada komunikasi rahasia yang sedang berlangsung. Walaupun isinya tetap rahasia, tapi sifat tersembunyi tersebut jadi diketahui.

Untuk gambar berformat JPEG, Outguess menjaga statistik berdasarkan hitungan frekuensi. Oleh karena itu tes statistik tidak akan bisa menentukan ada atau tidaknya penyembunyian data. Sebelum memasukkan data ke gambar, Outguess dapat mengukur berapa besar pesan yang dapat ia sembunyikan selagi mempertahankan statistik tersebut.

OutGuess menggunakan objek iterator umum untuk memilh bit mana yang harus dimodifikasi. Sebuah seed bisa dimanfaatkan untuk mengatur perilaku iterator. Perubahan disimpan di dalam data bersama dengan sisa dari pesan tersebut. Dengan mengubah seednya, OutGuess mencoba mencari sebuah urutan bit yang memiliki banyak perubahan minimal.

c. JSteg

JPEG Group's JPEG Software library oleh Derek Upham. Koefisien *DCT* dimodifikasi secara kontinu dari awal suatu gambar digital. *Jsteg* tidak



mendukung enkripsi dan tidak ada pemilihan bit acak. Data dari suatu pesan di gabung di awal dengan ukuran header yang bervariasi. Lima bit pertama dari header menyatakan ukuran *field* dalam bit. Lima bit selanjutnya menyatakan ukuran *field* yang menyatakan ukuran data yang disisipkan.

Gambar digital berisi pesan rahasia dengan sistem *Jsteg* menunjukkan kemungkinan terbesar disisipkan di awal gambar. Hasil menunjukkan 0 ketika test mencapai bagian yang tidak dimodifikasi dari koefisien DCT.

Pseudocode:

```
Input: Image I, Message ~m
Output: Image J
for each bit b of ~m
   c := next DCT coefficient from I
   while c = 0 or c = 1,
      c := next DCT coefficient from I
   end while
   c := c mod 2 + b
   replace coefficient in I by c
end for
```

Deteksi Steganografi

Pendeteksian data yang disembunyikan dengan steganografi disebut steganalisis. Cara termudah untuk melakukannya adalah dengan membandingkan gambar asli dengan gambar yang telah dimodifikasi. Sebuah contoh adalah untuk memeriksa pemindahan informasi melalui gambar-gambar di *website*, seorang steganalis yang menyimpan file-file gambar aslinya dapat membandingkannya dengan gambar-gambar bersesuaian yang sedang dipampang di *website* tersebut. Perbedaan yang didapat dari perbandingan memberikan gambaran tentang data yang disembunyikan.

Steganografi adalah ilmu dan seni penyembunyian terhadap adanya komunikasi. Sistem steganografi klasik bergantung pada menjaga kerahasiaan sistem enkoding, sedangkan steganografi modern hanya dapat dideteksi jika informasi rahasia berupa kunci rahasia tertangkap. Karena sifat sistemnya yang menyusup ke dalam, steganografi meninggalkan jejak dalam karakteristik dari medium penyimpanannya. Terdeteksinya jejak tersebut memberikan peluang bagi pihak ketiga untuk menemukan file mana yang sudah diubah, menandakan bahwa ada komunikasi rahasia yang sedang berlangsung. Walaupun isinya tetap rahasia, tapi sifat tersembunyi tersebut jadi diketahui,

Secara umum, dalam gambar berkompresi tinggi, steganografi menjadi sulit tetapi tetap mungkin. Di suatu sisi kesalahan kompresi dapat menjadi ruang untuk menyisipkan data steganografi. Tetapi di lain pihak, kompresi tinggi menyebabkan data yang dapat disimpan menjadi lebih sedikit, meningkatkan kepadatan enkode, dan mempermudah deteksi steganografi, bahkan sampai tahap kasat mata.

Program yang digunakan untuk pendeteksian adalah *Stagdetect. Stegdetect* dapat menganalisis gambar digital yang memiliki data tersembunyi yang disisipkan dengan sistem JSteg, JPHide, dan OutGuess. Untuk setiap sistem yang mau dideteksi, kita memilih koefisien DCT dengan tujuan dimodifikasi dan melakukan tes χ2-test. Hasil dari *Stegdetect* menampilkan list *steganographic system* ang ditemukan di setiam gambar digital, atau "negative" jika tidak ada data tersembunyi yang disisipkan dalam gambar digital yang terdeteksi oleh *Stegdetect*.

Prinsip deteksi Outguess dimulai dengan mencari kuantitas makroskopis *S(m,q)* yang mengira-ngira perubahan panjang pesan rahasia *m* dan parameter *q* yang berasal dari nilai ekstrim S, misal S(0) dan S_{max}.

$$S(m,q)$$

$$S(m) = S_{stego}\ m$$

Dimana S_{stego} adalah nilai S dari gambar stego (gambar pembawa) yang sedang diamati.

$$B = \sum_{i=1}^{\lfloor M-1/8 \rfloor} \sum_{j=1}^{N} \left| g_{8i,j} - g_{8i+1,j} \right| + \sum_{j=1}^{\lfloor N-1/8 \rfloor} \sum_{i=1}^{M} \left| g_{i,8j} - g_{i,8j+1} \right|$$

Blockiness B adalah penjumlahan (sum) dari diskontinuitas spatial dari pinggiran blok 8x8 JPEG.

B bertambah secar linear

$$\mathbf{S = B(m_{max}) - B(0)}$$

Langkah Pendeteksian:

1. Tentukan *blockiness* $B_S(0)$ gambar stego yang telah didekompresi
2. Gunakan OutGuess untuk memasukkan sebuah pesan dengan panjang maksimal dan hitung *blockiness* $B_S(1)$.
3. *Crop* gambar stego dengan mengurangi 4 piksel untuk membuat gambar kedua. Kompresi gambar ini dengan kuantisasi JPEG yang sama. Hitung blockiness $B(0)$.
4. Lakukan langkah 2 untuk gambar kedua, hitung *blockiness* $B(1)$.
5. Masukkan sebuah pesan dengan panjang maksimum ke gambar stego dari langkah 4 dan hitung *blockiness* $B1(1)$.
6. Kemiringan $S_0 = B(1) - B(0)$ berkorespondensi dengan gambar stego asli, sedangkan $S_1 = B1(1) - (1)$ untuk gambar yang disisipi pesan dengan panjang maksimum.

Kemiringan $S = B_S(1) - B_S(0)$ berada diantara S_0 and S_1. Panjang pesan tersembunyi P dihitung sebagai:

$$p = \frac{S_0 - S}{S_0 - S_1},$$

$p = 0$ berkorespondensi terhadap gambar stego covertext dan $p = 1$ adalah gambar dengan sisipan maksimum. Untuk mengantisipasi random dalam OutGuess, perhitungan dilakukan berkali-kali dan diambil p rata-ratanya sebagai hasil akhir.

Referensi

[1] Munir, Rinaldi, *Diktat Kuliah IF5054 Kriptografi*, Penerbit ITB 2006.

[2] Provos, Niels dan Peter Honeyman. "Hide and Seek: An Introduction to Steganography". *IEEE SECURITY & PRIVACY* edisi Mei/Juni 2003.

[3] Provos, Niels dan Peter Honeyman. "Detecting Steganographic Content on the Internet"

[4] Ankit, Jain. "Steganography : A Solution for Data Hiding"

[5] Schaathun, Georg Hans. "Steganography in JPEG". 2007. University of Surrey.