

Teknik Penyembunyian Data Rahasia pada Berkas Gambar Digital Menggunakan Steganografi *Least Significant Bit Variable-Size*

A Technique of Hiding Secret Data in Digital Images Using Variable-Sized Least Significant Bit Steganography

SHELVIE NIDYA NEYMAN^{1*}, LINDAYATI¹, SUGI GURITMAN²

Abstrak

Salah satu teknik pengamanan data dengan cara penyisipan informasi rahasia pada berkas gambar digital sebagai media penampung ialah steganografi berbasis *least significant bit* (LSB) dengan penambahan kemampuan penyisipan yang bersifat *variable-size*. Operasi teknik steganografi ini menggunakan tiga parameter, yaitu *capacity evaluation* untuk menentukan kapasitas maksimum LSB dari masing-masing *pixel* media penampung, *minimum error replacement* untuk memperkecil tingkat kesalahan saat penyisipan, dan *improved grayscale compensation* untuk memisahkan kesalahan penempelan agar tidak berdekatan pada tempat *pixel* bekerja. Penelitian ini bertujuan mengimplementasikan teknik tersebut untuk menyembunyikan data rahasia pada berkas gambar digital sebagai media penampung dan menghasilkan *stego-image* dengan kesamaan tampilan dan ukuran berkas. Berkas hasil penyisipan (*stego-image*) yang diperoleh diharapkan tidak menimbulkan kecurigaan pihak 'lawan', serta mampu memaksimalkan kapasitas penyisipan data rahasia, dengan memungkinkan penyisipan mencapai bit ke-5 dari LSB media penampung. Hasil penelitian menunjukkan bahwa *stego-image* yang diperoleh secara visual memiliki tampilan hampir sama dengan media penampungnya dan kapasitas penyisipannya mencapai lebih dari 50% ukuran media penampungnya. Kapasitas ini diperoleh tanpa mempertimbangkan *border* dan jumlah bit data pada setiap *pixel* minimal berjumlah empat.

Kata kunci: *capacity evaluation, improved grayscale compensation, least significant bit variabel-size, steganografi,*

Abstract

One of the techniques for securing data by insertion of confidential information in a digital image at a container media is steganography based on least significant bit (LSB) with additional ability of variable-sized insertion. The operation of this steganographic technique used three parameters: the capacity evaluation to determine the maximum capacity of the LSB of each pixel of container media, the minimum error replacement to minimize the error rate during insertion, and improved grayscale compensation for separating the errors of attachment so as not to be adjacent on the place where the pixels are working. This study aims to implement these techniques in hiding secret data in digital image as container media, to generate the stego-image with the similarity in appearance and size of the file. The stego-image is expected to be able to maximize the capacity of confidential data insertion, by allowing insertion up to the 5th bit of the least significant bits of pixel media placeholders. The results show that the stego image obtained is visually similar to the container media with capacity of more than 50% from the total size of the container media. This capacity does not include the border and the number of data bits in each pixel is at least four.

Keywords: capacity evaluation, improved grayscale compensation, least significant bit variable-size, steganography

¹Departemen Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor, Bogor 16680;

²Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor, Bogor 16680

*Penulis Korespondensi: Tel/Faks: 0251-8625584; Surel: shelvie@ipb.ac.id

PENDAHULUAN

Seiring dengan perkembangan ilmu bidang teknologi komunikasi dan informasi, jaringan komputer memunculkan kebutuhan akan perlindungan keamanan data dengan cara penyisipan informasi tertentu yang bersifat rahasia pada suatu berkas digital. Steganografi merupakan teknik keamanan data yang dapat digunakan untuk melayani kebutuhan tersebut (Cox *et al.* 2008).

Salah satu metode steganografi yang sering diimplementasikan ialah *least significant bits* (LSB), yang menyembunyikan data rahasia langsung pada media penampung (berkas gambar) dengan cara menggantikan bit LSB dengan bit pada data rahasia untuk setiap *pixel* gambar. Teknik LSB yang berkembang saat ini hanya mampu mengganti setengah dari keseluruhan bit media penampungnya untuk mempertahankan kualitas media tersebut (Johnson dan Jajodia 1998). Dalam teknik steganografi terdapat tiga aspek penting yang perlu diperhatikan, yaitu kapasitas yang berkaitan dengan banyaknya informasi yang dapat disembunyikan dalam media penampung, keamanan yang berkaitan dengan terjaganya kerahasiaan informasi dalam media penampung, dan ketahanan yang berkaitan dengan terjaganya keutuhan informasi dari banyaknya perubahan yang dilakukan terhadap media penampung yang telah berisi informasi rahasia (atau disebut *stego-image*) oleh pihak lawan (Provos dan Honeyman 2003).

Penelitian ini bertujuan mengimplementasikan teknik steganografi untuk meningkatkan kapasitas data rahasia yang bisa disisipkan pada media penampung. Teknik steganografi tersebut menggunakan metode penyisipan *variable-size* yang diusulkan oleh Lie dan Chang (1999). Metode ini dapat menggunakan bit LSB sedikitnya 4 untuk penyisipan data rahasia pada setiap *pixel* media penampung. Metode penyisipan *variable-size* ini mempunyai 3 komponen penentu utama, yaitu *capacity evaluation* (CE), *minimum error replacement* (MER), dan *improved grayscale* (IGSC). Komponen CE menentukan banyaknya LSB yang dapat disisipkan dari setiap *pixel* pada media penampung, *output*-nya dapat berjumlah 4 atau 5. Bagian kesalahan yang dapat ditimbulkan akibat nilai *output* CE diperbaiki oleh komponen MER dan IGSC. Komponen MER memperkecil tingkat kesalahan saat penyisipan dan IGSC untuk memisahkan kesalahan penyisipan jika terjadi kesalahan agar tidak berdekatan pada saat *pixel* bekerja.

METODE

Tahapan penelitian ini secara garis besar dibagi menjadi dua bagian, yakni implementasi dan pengujian hasil implementasi. Proses implementasi ada dua, yaitu penyisipan (Gambar 1) dan pendeteksian informasi (Gambar 2). Media penampung yang digunakan adalah berkas gambar digital jenis *grayscale* dan RGB berformat BMP, sedangkan data rahasia berupa gambar *grayscale* dengan format JPG. Hasil implementasi kemudian dianalisis menggunakan komponen CE, MER (*extended CE*), dan IGSC (*extended MER*). Pengamatan dilakukan pada kualitas media gambar setelah penyisipan menggunakan perhitungan PSNR, daya tampung media gambar, rata-rata jumlah bit per *pixel* dari media gambar, dan waktu penyisipan pada media gambar.

Teknik steganografi yang digunakan pada penelitian ini menggunakan metode LSB *variable-size* dengan mengganti k LSB pada media penampung dengan bit data atau informasi rahasia menggunakan persamaan:

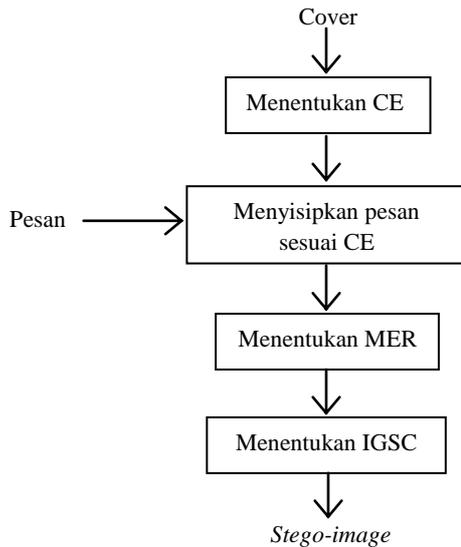
$$I_s(x,y) = I_c(x,y) - \text{mod}(I_c(x,y), 2^k) + B(x,y) \quad (1)$$

dengan:

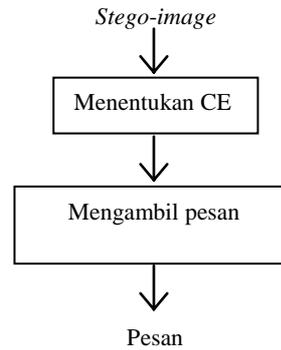
$I_s(x,y)$ = nilai *pixel* (x,y) dari *stego-image*,

$I_c(x,y)$ = nilai *pixel* (x,y) dari media penampung,

$B(x,y)$ = nilai desimal dari suatu blok data,



Gambar 1 Tahapan penyisipan pesan rahasia



Gambar 2 Tahapan deteksi kembali pesan rahasia

k = banyaknya bit data yang akan disisipkan.

Komponen CE pada teknik ini bertujuan menentukan kapasitas maksimum LSB dari setiap *pixel* media penampung. *Output* komponen CE dapat berjumlah 4 atau 5. Komponen CE menggunakan variasi *grayscale* dari ketetanggaan *pixel*, kemudian intensitasnya berguna untuk mengevaluasi kapasitas penyisipan. Langkah-langkah untuk menentukan CE ialah dengan mengasumsikan bahwa *grayscale* dari satu *pixel p* pada koordinat (x,y) dinotasikan sebagai $f(x,y)$, kemudian menentukan nilai $Kn(x,y)$. Nilai $f(x,y)$ pada *pixel p* akan diubah berdasarkan pada kapasitas penyisipan dan nilainya bergantung pada *grayscale*-nya dan variasi *grayscale* dari tetangga atas dan kiri (warna abu-abu pada Gambar 3).

Keuntungan menggunakan tetangga atas dan kiri untuk mengestimasi kapasitas penyisipan ialah pada saat atau sesudah *pixel* tertentu diproses, *grayscale* dari tetangga atas dan kiri tidak pernah berubah. Oleh karena itu, modul penyisipan dan ekstraksi akan sinkron pada saat mengestimasi kapasitas penyisipan dari setiap *pixel*. Selanjutnya, dilakukan perhitungan untuk menentukan $Kn(x,y)$.

$$\text{Max}(x,y) = \max\{f(x-1, y-1), f(x-1, y), f(x-1, y+1), f(x, y-1)\},$$

$$\text{Min}(x, y) = \min\{f(x-1, y-1), f(x-1, y), f(x-1, y+1), f(x, y-1)\},$$

$$D(x, y) = \text{max}(x, y) - \text{min}(x, y) \quad (2)$$

Kapasitas penyisipan $Kn(x,y)$ dari setiap *pixel* (x,y) didefinisikan sebagai:

$$Kn(x,y) = \log_2 D(x,y) \quad (3)$$

Kapasitas penyisipan seharusnya terbatas oleh *grayscale* dari *pixel* tertentu. Batas atas untuk kapasitas penyisipan pada *pixel* (x,y) didefinisikan sebagai berikut:

$(x-1,y-1)$	$(x-1,y)$	$(x-1,y+1)$
$(x,y-1)$	(x,y)	$(x,y+1)$
$(x+1,y-1)$	$(x+1,y)$	$(x+1,y+1)$

Gambar 3 Delapan ketetanggaan dari *pixel p*

$$U(x,y) = \begin{cases} 4, & f(x,y) \leq t \\ 5, & \text{selainnya} \end{cases} \quad (4)$$

Nilai t dapat diset dengan nilai tertentu, dengan syarat nilai $U(x,y)$ harus konsisten pada modul penyisipan dan ekstraksi. Batas bawah untuk kapasitas penyisipan diset 4 bit sehingga kapasitas penyisipan $K(x,y)$ dari setiap *pixel* bisa dihitung dengan persamaan (Lee dan Chen 2000):

$$K(x,y) = \min\{\max\{K_n(x,y), 4\}, U(x,y)\} \quad (5)$$

Komponen MER dipakai untuk memperkecil tingkat kesalahan saat penyisipan, dengan cara mencari nilai *grayscale* yang sedekat mungkin dengan nilai aslinya. Langkah-langkah untuk menghitung nilai MER dimulai dari $f(x,y)$ yang merupakan nilai *grayscale* asli, $g(x,y)$ yang merupakan nilai *grayscale* yang dihasilkan dari penyisipan k LSB, dan $g'(x,y)$ yang merupakan nilai *grayscale* yang dihasilkan dari pengubahan nilai LSB ke $k+1$ dari $g(x,y)$. *Minimum error grayscale* berada di antara $g(x,y)$ atau $g'(x,y)$. $e(x,y)$ adalah *error* yang terjadi antara $f(x,y)$ dan $g(x,y)$, dan $e'(x,y)$ adalah *error* yang terjadi antara $f(x,y)$ dan $g'(x,y)$. Jika $e(x,y) < e'(x,y)$, $g(x,y)$ akan digunakan untuk menggantikan $f(x,y)$, selainnya $g'(x,y)$ akan menggantikan $f(x,y)$.

Komponen IGSC digunakan untuk memisahkan kesalahan penempelan agar tidak berdekatan pada tempat *pixel* bekerja. Dalam komponen IGSC, *error* penyisipan biasanya disebarkan pada tetangga kanan dan bawah *pixel* (warna putih pada Gambar 3). $e(x,y)$ dinotasikan sebagai *error* penyisipan dari *pixel* p pada koordinat (x,y) , nilai *grayscale* empat tetangga bawah dan kanan kemudian dimodifikasikan dengan persamaan berikut (Lee dan Chen 2000):

$$\begin{aligned} f(x, y + 1) &= f(x, y + 1) + 1/4 e(x, y), \\ f(x + 1, y - 1) &= f(x + 1, y - 1) + 1/4 e(x, y), \\ f(x + 1, y) &= f(x + 1, y) + 1/4 e(x, y), \\ f(x + 1, y + 1) &= f(x + 1, y + 1) + 1/4 e(x, y) \end{aligned} \quad (6)$$

Pengujian hasil implementasi dengan PSNR digunakan untuk mengukur distorsi yang terjadi antara gambar yang telah mengalami manipulasi dengan gambar aslinya. Semakin besar nilai PSNR, semakin baik gambar tersebut karena gambar tersebut lebih sedikit mengalami distorsi dan sebaliknya. Satuan dari nilai PSNR ialah desibel (dB). Nilai PSNR dapat dihitung dengan persamaan berikut:

$$\text{PSNR} = 10 \log_{10} \left(\frac{\text{MAX}_I^2}{\text{MSE}} \right) = 20 \log_{10} \left(\frac{\text{MAX}_I}{\sqrt{\text{MSE}}} \right) \quad (7)$$

Nilai *mean squared error* (MSE) dihitung dengan persamaan berikut:

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i,j) - K(i,j)\|^2 \quad (8)$$

m = jumlah baris/tinggi gambar,

n = jumlah kolom/lebar gambar,

$I(i,j)$ = nilai piksel dari gambar asli,

$K(i,j)$ = nilai piksel gambar yang dimodifikasi,

MAX_I = nilai piksel maksimum dari gambar (untuk gambar yang direpresentasikan menggunakan 8 bit per sampel, maka MAX_I adalah 255).

HASIL DAN PEMBAHASAN

Dari proses penyisipan data menggunakan komponen CE, MER, dan IGSC dengan $t = 135$ untuk level 1, $t = 143$ untuk level 2, $t = 159$ untuk level 3, dan $t = 191$ untuk level 4 diperoleh hasil sebagai berikut.

Nilai PSNR dari Media Penampung *Grayscale* dan RGB pada Level 1, 2, 3, dan 4 Menggunakan Komponen CE, MER, dan IGSC

Nilai PSNR menurun dari level 1 ke level 4 baik untuk media penampung *grayscale* dan media penampung RGB pada komponen CE, MER, IGSC, dan fx LSB (Tabel 1). Hal ini berarti pada level 4 distorsi yang terjadi lebih besar dibandingkan dengan level lainnya pada proses penyisipan data baik menggunakan komponen CE, MER, IGSC ataupun fx LSB. Hal ini terjadi karena pada level 4, bit data yang disisipkan pada media penampung lebih banyak dibandingkan dengan level lainnya. Dengan demikian, semakin banyak bit data yang disisipkan, kemungkinan bit-bit pada media penampung yang berubah akan semakin banyak sehingga distorsi yang terjadi semakin besar.

Dari Tabel 1 terlihat juga bahwa penyisipan data menggunakan komponen MER untuk media penampung *grayscale* dan RGB pada setiap level memiliki nilai PSNR yang lebih besar dibandingkan dengan penyisipan menggunakan komponen CE ataupun IGSC. MER digunakan untuk mencari nilai *grayscale* sedekat mungkin dengan nilai aslinya sehingga distorsi yang terjadi lebih kecil dibandingkan dengan komponen CE.

Nilai PSNR penyisipan data menggunakan komponen IGSC untuk media penampung *grayscale* dan RGB pada level 1, 2, 3, dan 4 lebih kecil dibandingkan dengan komponen MER, tetapi lebih besar dibandingkan dengan komponen CE. Penyisipan data menggunakan komponen IGSC untuk media penampung *grayscale* dan RGB pada level 1 dan 2 memiliki nilai PSNR lebih kecil dibandingkan dengan penyisipan fx LSB, akan tetapi memiliki nilai PSNR yang lebih besar daripada level 3 dan 4. Hal ini menunjukkan bahwa penyisipan data menggunakan metode *variable-size* lebih baik dibandingkan dengan penyisipan fx LSB untuk level besar. Selain itu, secara visual, gambar yang dihasilkan dengan metode *variable-size* juga lebih baik dibandingkan dengan penyisipan fx LSB.

Daya Tampung Media Penampung *Grayscale* dan RGB pada Level 1, 2, 3, dan 4 Menggunakan Komponen CE, MER, dan IGSC.

Tabel 2 menunjukkan daya tampung komponen CE, MER, IGSC untuk media penampung *Grayscale* dan RGB. Kenaikan daya tampung terjadi dari level 1 ke level 4 untuk media penampung *grayscale* dan RGB dengan dimensi yang sama pada komponen CE, MER, dan IGSC. Hal ini terjadi karena semakin tinggi level yang digunakan maka bit data yang dapat disisipkan akan semakin banyak. Media penampung RGB juga terlihat memiliki daya tampung yang lebih besar dibandingkan dengan *grayscale*. Hal ini terjadi karena media penampung RGB mempunyai tiga *layer* untuk menyisipkan data sehingga daya tampungnya pun akan tiga kali lipat dari media penampung *grayscale* dengan dimensi yang sama.

Tabel 2 juga menunjukkan bahwa penyisipan data menggunakan komponen CE, MER, dan IGSC memiliki daya tampung yang hampir sama, namun lebih besar jika dibandingkan dengan penyisipan fx LSB.

Rata-rata Jumlah Bit per *Pixel* dari Media Penampung *Grayscale* dan RGB pada Level 1, 2, 3, dan 4 Menggunakan Komponen CE, MER, dan IGSC

Rata-rata jumlah bit per *pixel* pada media penampung RGB dapat dihitung dengan persamaan berikut:

$$\text{bit_per_pixel} = \frac{\text{daya_tampung}}{\text{jumlah_pixel}} \quad (9)$$

Tabel 1 Nilai PSNR komponen *capacity evaluation* (CE), *minimum error replacement* (MER), dan *improved grayscale compensation* (IGSC) untuk media penampung *grayscale* dan RGB

<i>Grayscale</i>				
Level	CE	MER	IGSC	Fix LSB
1	48.6457	49.1489	49.1489	50.1837
2	42.4872	43.5465	43.3694	43.5281
3	36.4114	37.7224	37.2810	36.9676
4	31.2577	32.3538	31.6172	31.2451
RGB				
Level	CE	MER	IGSC	Fix LSB
1	49.8631	50.4572	50.4572	50.4647
2	42.1415	43.1926	42.9416	43.5563
3	35.2592	36.7377	36.5503	36.4299
4	30.9989	32.1296	31.3328	31.2245

Tabel 2 Daya tampung komponen CE, MER, dan IGSC untuk media penampung *grayscale* dan RGB

<i>Grayscale</i>			
Level	CE	MER	IGSC
1	1.1909	1.1891	1.1891
2	2.1221	2.1184	2.1185
3	3.0525	3.0498	3.0525
4	4.0104	4.0107	4.0100
RGB			
Level	CE	MER	IGSC
1	1.3321	1.3278	1.3278
2	2.2219	2.2168	2.2211
3	3.1174	3.1136	3.1210
4	4.0296	4.0272	4.0286

Rata-rata jumlah bit per *pixel* pada media penampung *grayscale* dengandimensi $m \times n$ dihitung dengan persamaan berikut:

$$\text{per_pixel} = \frac{\text{dayatampung}}{(m-2)(n-2)} \quad (10)$$

Rata-rata jumlah bit per *pixel* yang diperoleh untuk masing-masing komponen disajikan pada Tabel 3.

Dapat dilihat bahwa rata-rata jumlah bit per *pixel* untuk media penampung *grayscale* dan RGB dengan dimensi yang sama bertambah seiring dengan meningkatnya level. Penyisipan data menggunakan komponen CE, MER, dan IGSC memiliki rata-rata jumlah bit per *pixel* yang hampir sama, namun lebih besar dibandingkan dengan penyisipan *fx* LSB. Hal ini disebabkan oleh rata-rata jumlah bit per *pixel*, yang sebanding dengan daya tampung.

Waktu Penyisipan dengan Menggunakan Komponen CE, MER, dan IGSC dari Media Penampung *Grayscale* dan RGB pada Level 1, 2, 3, dan 4

Tabel 4 menunjukkan bahwa penyisipan data menggunakan komponen MER memiliki waktu yang lebih lama dibandingkandengan penyisipan menggunakan komponen CE. Hal ini terjadi karena pada penyisipan data menggunakan komponen MER, algoritme yang digunakan adalah algoritme untuk menentukan CE ditambah algoritme untuk menentukan MER.

Waktu eksekusi penyisipan data menggunakan komponen IGSC paling besar karena algoritme yang digunakan adalah algoritme untuk menentukan CE, MER, ditambah algoritme menentukan IGSC.

Tabel 3 Rata-rata jumlah bit per *pixel* komponen CE, MER, dan IGSC untuk media penampung *grayscale* dan RGB

<i>Grayscale</i>			
Level	CE	MER	IGSC
1	054030	053950	053950
2	096277	096111	096115
3	138490	138365	138489
4	181946	181961	181931
RGB			
Level	CE	MER	IGSC
1	181307	180719	180719
2	302414	301724	302308
3	424298	423785	424786
4	548461	548131	548327

Tabel 4 Waktu penyisipan menggunakan komponen CE, MER, dan IGSC untuk media penampung *grayscale* dan RGB (detik)

<i>Grayscale</i>			
Level	CE	MER	IGSC
1	8.1250	15.954	16.000
2	9.1720	16.984	17.062
3	9.4530	17.078	17.391
4	11.265	17.578	17.594
RGB			
Level	CE	MER	IGSC
1	8.1250	15.954	16.000
2	9.1720	16.984	17.062
3	9.4530	17.078	17.391
4	11.265	17.578	17.594

SIMPULAN

Teknik steganografi berbasis LSB pada gambar dengan metode penyisipan *variable-size* mampu menyembunyikan informasi rahasia dengan baik. Selain itu, penggunaannya dapat disesuaikan dengan ukuran media gambar yang digunakan dan informasi yang akan disisipkan.

Penelitian ini dapat dikembangkan lebih lanjut ke arah peningkatan keamanan, dengan menerapkan kriptografi untuk mengenkripsi data sebelum disisipkan ke dalam media gambar. Cara lain untuk meningkatkan aspek keamanan adalah dengan menyisipkan data ke posisi tertentu saja dari ke seluruhnya k bit LSB. Dalam hal ini, sebuah kunci diperlukan untuk membangkitkan suatu bilangan dan menentukan posisi penyisipan.

DAFTAR PUSTAKA

- Cox IJ, Miller ML, Bloom JA, Fridrich J, Kalker T. 2008. *Digital Watermarking and Steganography*, Ed ke-2. Burlington (US): Elsevier.
- Johnson NF, Jajodia S. 1998. Exploring steganography: seeing the unseen. *IEEE Comp Soc.* 31(2):26-34. 10.1109/MC.1998.4655281.
- Lee YK, Chen LH. 2000. A high capacity image stegano-graphic model. *IEEE Proc Vis Image Signal Proc.* 147(3):288-294. 10.1049/ip-vis:20000331.
- Lie WN, Chang LC. 1999. Data hiding in images with adaptive number of least significant bits based on human visual systems. *Proc ICIP.* 99(1):286-290. 10.1109/ICIP.1999.821615.
- Provos N, Honeyman P. 2003. Hide and seek: an introduction to steganography. *IEEE Sec Priv.* 1(3):32-44. 10.1109/MSECP.2003.1203220.