

PENERAPAN AUDIO STEGANOGRAFI DALAM INTRASONICS

Risa Astari Dewi – NIM : 13506064

Program Studi Teknik Informatika, Institut Teknologi Bandung

Jl. Ganesha 10, Bandung

E-mail : if16064@students.if.itb.ac.id

Abstrak

Makalah ini membahas penerapan steganografi dalam *file* audio untuk menyimpan pesan berupa instruksi bagi mesin. Melalui *stego-object* tersebut, sebuah mesin dapat dikendalikan dari suara atau audio yang dimainkan. Dengan kondisi lingkungan tempat mesin nantinya, diperlukan *stego-object* yang *imperceptible* dan kokoh. Hal ini mendukung ketersediaan perintah yang diinginkan dan mesin mampu menerjemahkan perintah yang dimaksud.

Algoritma penyisipan dalam audio steganografi yang digunakan dengan metode LSB. Namun, metode tersebut belum cukup untuk meningkatkan kekokohan *stego-object*. Pada algoritma LSB standar jika ingin meningkatkan kekokohan *stego-object*, penyisipan dilakukan pada layer terdalam. Namun hal ini akan mengakibatkan semakin jelasnya perbedaan antara audio asli dengan *stego-object*. Resiko ini tidak mungkin diambil karena pendengaran manusia akan terganggu oleh suara bising yang diakibatkan proses penyisipan data dalam *file* audio. Pada makalah ini lebih lanjut akan membahas algoritma yang mampu meningkatkan hal itu, menghasilkan *stego-object* yang kokoh (*robust*) juga *imperceptible* bagi pendengaran manusia. Dengan memodifikasi metode LSB standar, algoritma akan melalui 2 langkah pendekatan.

Makalah ini juga membahas pengujian yang dilakukan terkait algoritma baru dan keterkaitannya dengan *intrasonic*. Disamping itu juga terdapat faktor pendukung yang akan lebih mengoptimalkan kinerja *intrasonic* itu sendiri.

Kata kunci : *intrasonic, audio steganografi, LSB, watermarking, imperceptible*

1. Pendahuluan

Perkembangan audio dalam teknologi semakin jauh dalam berimajinasi. Terlebih lagi dalam memanfaatkan audio menjadi media penyampaian data. Penggunaan audio secara umum dengan mendengarkan *file* dan data mampu ditangkap secara langsung. Kini ilmuwan mewujudkan imajinasi lainnya, menyimpan suatu informasi yang disembunyikan dalam audio yaitu audio steganografi. Pengimplementasiannya dengan menyisipkan informasi dalam audio dengan penggunaan algoritma tertentu. Dalam makalah ini membahas bentuk lain dari pengaplikasian audio steganografi yaitu penyimpanan instruksi mesin dalam audio. Makalah ini akan membahas penerapan dari metode penyisipan pesan. Penyisipan instruksi mesin ke dalam *file* audio. Sebuah mesin yang dikendalikan oleh intruksi yang disisipkan, akan merespon kejadian saat audio diperdengarkan. Metode ini merupakan

inovasi baru. Sebelumnya permasalahan serupa diatasi dengan *ultrasounds*[1].

Ultrasound adalah salah satu mekanisme pengiriman sinyal mesin yang luput dari telinga manusia. Sayangnya sinyal ini justru tertangkap oleh jenis hewan tertentu dan mengganggu binatang peliharaan bila diterapkan dalam rumah tangga. Menangani hal itu, timbul ide baru dengan menyisipkan instruksi mesin dalam audio yang kemudian oleh mesin penerima akan didekripsi untuk mengambil instruksi yang dimaksud.

Pemahaman steganografi sendiri secara umum adalah menjadikan gambar, audio atau video sebagai medium penyisipan bagi pesan. Keberhasilan steganografi bergantung pada ketidakmampuan visual dan auditori manusia. *Stego-object* memerlukan alat bantu untuk mengekstrak pesan yang disisipkan. Pada audio steganografi berbasis computer, pesan rahasia

disisipkan dalam suara digital. Penyisipan dilakukan dengan cara mengubah sekuensial binary dari *file* suara. Aplikasi yang telah ada mampu menyisipkan pesan dalam WAV, AU, bahkan *file* MP3. Audio steganografi mengambil keuntungan dari *psychoacoustical* manusia atau kemampuan psikus terhadap akustik. Salah satu kelemahan yang dimanfaatkan adalah *imperceptible*, hasil *stego-object* tidak dapat dipersepsikan secara visual/auditori. Audio *stego-object* akan terdengar sama persis dengan *file* audio aslinya.

2. Audio steganografi

Sangat mungkin mengubah kode modulasi pada *file* suara sedemikian sehingga perubahan tidak dapat dideteksi oleh pendengaran manusia. Perubahan tersebut berisi data dari sumber. Misalkan pada wav *file*, *file* ini memiliki standar representasi 2 bytes untuk satu bagian wav. Manipulasi bisa dilakukan pada urutan terendah, jika ukurannya cukup, perubahan bisa tidak dirasakan.

Kriteria yang menilai sebuah stenograf dikategorikan bagus, *imperceptible* (keberadaan pesan tidak dapat dipersepsikan), fidelity (mutu cover-object tidak jauh berubah dari akibat penyisipan) dan recovery (data yang disembunyikan bisa diambil kembali)[4]. Pada pembahasan makalah ini, untuk memungkinkan pengaplikasian berhasil diberikan dua kebutuhan. Pertama, cover object dan *stego-object* tidak bisa dibedakan secara persepsi. Kedua, data dengan high data rate bisa disisipkan. Disemua aplikasi steganografi, selain memerlukan *high bit rate* pada *embeded data*, juga membutuhkan algoritma yang mampu mendeteksi dan mengembalikan bit tersembunyi tanpa mengakses *file* yang asli.

Metode LSB (*Least Significant bit*) adalah salah satu teknik yang dipelajari dalam penyembunyian data dan watermarking digital audio. Keuntungan LSB adalah mudah dalam mengimplementasikan dan proses encoding yang cepat. Disisi lain kelemahannya adalah tidak tahan terhadap perubahan terhadap cover object.

2.1 Metode LSB Standar

Penyisipan data pada least significant bits (LSB) dari contoh audio merupakan salah satu algoritma termudah dengan kemampuan

penambahan informasi pada *high data rate*. LSB encoder biasanya mengambil subset dari contoh audio utama yang dipilih dengan kunci rahasia. Operasi substitusi pada LSB dilakukan pada subset ini, nilai bit asli diganti dengan bit yang ingin disembunyikan. Proses ekstraksi dengan membaca nilai bits dari audio stego object. Maka dari itu, decoder memerlukan semua contoh dari stego audio yang digunakan saat penyisipan. Namun, penyeleksian acak tersebut akan menimbulkan AWGN (*additive white Gaussian noise*) dalam skala kecil. Diketahui bahwa system auditori manusia sangat sensitif dengan AWGN. Fakta ini membatasi jumlah LSB yang bisa *imperceptible* saat dilakukan penyisipan.

Seiring dengan meningkatnya penggunaan LSB saat pemrograman, modifikasi terhadap kedalaman layer LSB semakin besar, kemungkinan pesan tersembunyi mampu dideteksi juga meningkat dan persepsi yang transparan terhadap stego object menurun. Sehingga muncul batasan untuk penggunaan tingkat kedalaman layer LSB pada contoh audio yang bisa digunakan untuk menyembunyikan data.

Maksimum kedalaman layer LSB secara rata-rata yang bisa digunakan tanpa menyebabkan distorsi yang bisa diketahui layer keempat LSB dengan 16 bits per contoh *audio sequences* digunakan. Di sisi lain, tingkat kekokohan dari *stego-object* akan bertambah seiring meningkatnya kedalaman layer LSB yang digunakan. Oleh karena itu diperlukan modifikasi algoritma standar LSB untuk meningkatkan kekokohan (*robustness*) dan juga memberikan hasil yang *imperceptible*.

2.2 Pengembangan metode LSB

Pengembangan dengan metode yang berbeda dengan sebelumnya memungkinkan perubahan batasan penyisipan data, dari layer keempat LSB hingga layer keenam LSB. Menggunakan 2 langkah pendekatan. Langkah pertama, bit *watermark* disisipkan ke layer ke-*i* LSB dari host audio. Langkah kedua, suara bising yang diakibatkan oleh penyisipan *watermark* ditataulang[2].

Metode LSB standar hanya dengan menukarkan bit audio asli pada layer ke- i ($i=1, \dots, 16$) dengan bit dari *watermark* bit stream. Jika bit asli dan *watermark* berbeda dan layer ke-*i* LSB yang digunakan rusak akibat watermarking maka

kuantis langkah adalah 2^{i-1} (kisaran amplitude -32768,32767). Penyisipan pasti mengalami kerusakan jika bit asli adalah 0 dan bit *watermark* adalah 1 dan begitu pula sebaliknya.

Ide utama dari algoritma yang diajukan adalah penyisipan bit *watermark* yang hanya menimbulkan distorsi minimal pada host audio. Hal ini terlihat jelas, jika hanya satu dari 16 bits telas pasti dan sebanding dengan bit *watermark*, bit lainnya bisa diputar balik untuk meminimalisir kerusakan penyisipan. Contohnya, jika nilai asli adalah $0\dots01000_2=8_{10}$, dan bit *watermark* yang disisipkan pada layer ke-4 LSB adalah 0 daripada $0\dots00000_2=0_{10}$ (sesuai standar LSB), algoritma ini akan menghasilkan nilai contoh $0\dots00111_2=7_{10}$, jelas lebih mendekati nilai asli. Akan tetapi algoritma untuk mengekstrak masih sama, mengembalikan bit *watermark* dengan cara membaca nilai bit dari layer LSB yang telah terdefinisi pada contoh audio.

Pada algoritma penyisipan, layer ke-(i+1) LSB (bit a_i) mengalami modifikasi pertama dengan memasukan bit pesan. Kemudian algoritma dibawah dijalankan.

```

if host sample a>0
    if bit 0 is to be embedded
        if  $a_{i-1}=0$  then  $a_{i-1}a_i-$ 
 $2\dots a_0=11\dots 1$ 
        if  $a_{i-1}=1$  then  $a_{i-1}a_i-$ 
 $2\dots a_0=00\dots 0$  and
            if  $a_{i+1}=0$  then  $a_{i+1}=1$ 
            else if  $a_{i+2}=0$  then
 $a_{i+2}=1$ 
            ...
            else if  $a_{15}=0$  then  $a_{15}=1$ 
        else if bit 1 is to be
embedded
            if  $a_{i-1}=1$  then  $a_{i-1}a_i-$ 
 $2\dots a_0=00\dots 0$ 
            if  $a_{i-1}=0$  then  $a_{i-1}a_i-$ 
 $2\dots a_0=11\dots 1$  and

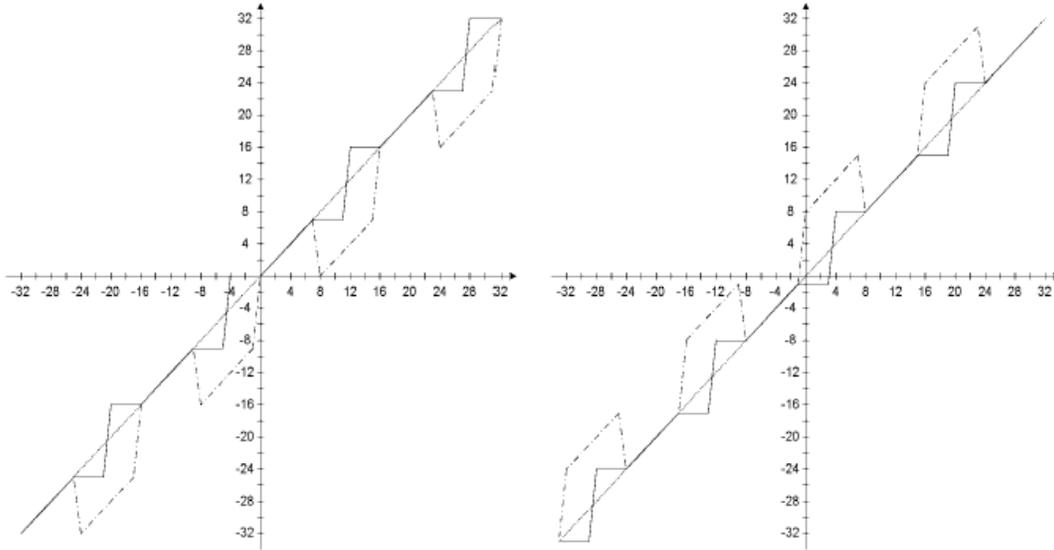
```

```

        if  $a_{i+1}=1$  then  $a_{i+1}=0$ 
        else if  $a_{i+2}=1$ 
then  $a_{i+2}=0$ 
            ...
            else if  $a_{15}=1$ 
then  $a_{15}=0$ 
if host sample a<0
    if bit 0 is to be embedded
        if  $a_{i-1}=0$  then  $a_{i-1}a_i-$ 
 $2\dots a_0=11\dots 1$ 
        if  $a_{i-1}=1$  then  $a_{i-1}a_i-$ 
 $2\dots a_0=00\dots 0$  and
            if  $a_{i+1}=1$  then  $a_{i+1}=0$ 
            else if  $a_{i+2}=1$  then
 $a_{i+2}=0$ 
            ...
            else if  $a_{15}=1$  then  $a_{15}=0$ 
        else if bit 1 is to be
embedded
            if  $a_{i-1}=1$  then  $a_{i-1}a_i-$ 
 $2\dots a_0=00\dots 0$ 
            if  $a_{i-1}=0$  then  $a_{i-1}a_i-$ 
 $2\dots a_0=11\dots 1$  and
                if  $a_{i+1}=1$  then  $a_{i+1}=0$ 
                else if  $a_{i+2}=1$  then
 $a_{i+2}=0$ 
                ...
                else if  $a_{15}=1$  then  $a_{15}=0$ 

```

Ilustrasi hasil algoritma yang diajukan dengan algoritma standar LSB dapat dilihat pada gambar berikut. Gambar kiri untuk kasus bit *watermark* sama dengan 0 dan gambar kanan untuk kasus bit *watermark* sama dengan 1. Gambar menunjukkan contoh karakteristik penyisipan pada layer keempat LSB, nilai hasil algoritma yang diajukan direpresentasikan oleh garis putus-putus. Dapat dilihat bahwa metode tersebut menghasilkan kerusakan yang lebih kecil dari metode standar. Jika layer keempat LSB digunakan, jangkauan kerusakan berkisar dari 1 sampai 4 QS, dimana dengan metode standar nilai error hingga 8 QS[3].



Gambar 1 Kiri : karakteristik penyisipan metode standar vs algoritma LSB yang diajukan (bit yang disisipkan 0). Kanan: perbandingan dengan bit yang disisipkan 1.

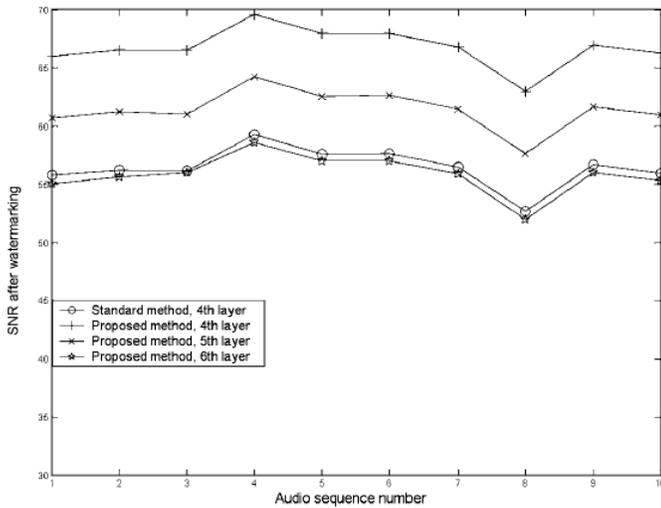
Diharapkan dengan menggunakan algoritma dua langkah ini, mampu meningkatkan kedalaman dari penyisipan *watermark* yang lebih daripada layer keempat dan juga meningkatkan kekokohan *stego-object*.

2.3 Pengujian Terhadap Algoritma

Uji coba untuk algoritma baru sudah pernah diadakan[3]. Algoritma diujikan pada 10 contoh audio dengan genre musik yang berbeda (pop, rock, jazz, techno). Contoh ini dipilih karena

dianggap merepresentasikan berbagai genre musik yang berbeda karakteristik dan dinamikanya. Semua musik kemudian diberi *watermark* menggunakan algoritma baru dan algoritma LSB standar. Contoh adalah mono audio *file* 44.1 kHz yang direpresentasikan dengan 16 bits per *sample*. Durasi *sample* antara 10 sampai 15 detik.

Gambar berikut menampilkan perbandingan kedua metode menurut SNR (signal to noise ratio). Bisa dilihat perbedaan hasil penyisipan keduanya.



Gambar 2 Nilai SNR dari 10 uji coba dengan LSB standar dan algoritma baru

Dua garis paling bawah menggambarkan LSB standar (garis dengan bulatan) dan algoritma baru (garis berbintang). LSB standar diterapkan pada layer keempat contoh audio (layer keempat LSB merupakan batas maksimal untuk menerapkan metode ini) sedangkan algoritma baru diterapkan pada layer keenam LSB. Menurut SNR, keduanya menghasilkan nilai yang tidak jauh berbeda. Hanya saja algoritma baru menghasilkan nilai yang lebih kecil dari LSB standar.

Uji coba selanjutnya dilakukan secara subjektif dengan memperdengarkan audio kepada 10 orang relawan. 3 diantaranya memiliki tingkat musik dasar, lanjut, atau musisi aktif. Pada tes pertama, relawan akan mendengarkan *file* audio asli dan yang telah diberi *watermark* secara bergantian. Tabel atas merepresentasikan hasil dari tes pertama. Pada tes kedua, relawan kembali mendengarkan *file* audio tapi secara acak. Hasilnya direpresentasikan pada tabel bawah.

	Country	Violin	Pop
Discrimination values (%)			
Standard method (4 th LSB)	52	49	51
Standard method (5 th LSB)	59	40	57
New method (6 th LSB)	51	50	51
New method (7 th LSB)	55	45	55
Mean opinion score (MOS)			
Standard method (4 th LSB)	5.0	4.9	5.0
Standard method (5 th LSB)	4.6	4.5	4.7
New method (6 th LSB)	5.0	5.0	5.0
New method (7 th LSB)	4.6	4.6	4.6

Gambar 3 MOS (*mean opinion scores*) dan *discrimination values*

Dari gambar dapat dilihat nilai diskriminasi dan MOS antara penyisipan pada layer ke-6 LSB dengan algoritma baru dan layer ke-4 dengan LSB standar hampir sama. Hal ini menunjukkan algoritma baru berhasil menyisipkan data pada layer ke-4 hingga ke-6 LSB tanpa mengakibatkan tranparansi persepsi terhadap sinyal audio. Sehingga, peningkatan terhadap kekokohan *stego-object* bisa diwujudkan.

3. Penerapan Dalam *Intrasonic*

Sebagaimana yang telah dibahas pada awal makalah, penerapan audio steganografi dalam *intrasonic*. Salah satu kekurangan metode LSB

standar adalah hasil penyisipan yang tidak kokoh, sedangkan pada *intrasonic* nantinya faktor ini sangat penting untuk menjaga ketersediaan perintah kendali pada mesin. Dengan diujukannya algoritma yang baru, diharapkan tingkat kekokohan *stego-object* meningkat.

Penerapan *intrasonic* sendiri memang diarahkan ke dunia hiburan dan multimedia menurut imajinasi Scientific Generics, pencetus ide *intrasonic*. Ide ini muncul melihat kebosanan anak-anak saat menonton televisi. *Intrasonic* diterapkan pada mainan. Mainan ini akan memberi respon jika mendengarkan audio berisi sinyal kendali dari televisi atau multimedia lainnya. Dari segi *hardware* pendukung, mesin penerima hanya memerlukan mikropon, baterai dan *single-chip decoder*. Decoder yang dirancang tidaklah rumit, karena pada LSB sendiri proses pengambilan data yang disisipkan dapat dengan mudah dilakukan, hanya perlu mengembalikan bit *watermark* dengan cara membaca nilai bit dari layer LSB yang telah terdefinisi pada contoh audio. Tentu saja untuk penerapan pada mainan, sinyal kendali yang disisipkan sangat sederhana.

Faktor pendukung lainnya adalah akustik dari lingkungan penerima. Dalam contoh kasus mainan ini, yang harus diperhitungkan adalah jarak antara mainan dengan sumber suara dan tingkat kebisingan lingkungan sekitarnya. Semakin jauh jarak penerima dari sumber, suara yang ditangkap akan mengalami distorsi sehingga susah untuk dibaca. Walaupun penerapan algoritma baru telah membuat kokoh hasil penyisipan, tetapi hasil tangkapan suara yang terdistorsi akan mempengaruhi pembacaan bit dari audio, tingkat kepastiannya pun menurun. Begitu juga jika lingkungan sekitar penerima dipenuhi oleh suara lainnya. Untuk penerapan dalam mainan sederhana hal ini sulit dilakukan. Mesin perlu ditambahkan filter untuk menyaring suara yang dimaksudkan. Jika tidak, suara lain bisa merusak susunan bit dari sumber suara aslinya.

4. Kesimpulan

Penerapan *intrasonic* akan menjadi optimal dengan menggunakan pendekatan dua langkah pada algoritma baru. Hal ini ditunjukkan dengan hasil uji coba yang membuktikan tidak adanya perbedaan antara metode LSB standar dengan algoritma baru jika dinilai dari pendengaran

subjektif manusia. Namun pada algoritma baru tingkat kekokohan stego-object lebih besar dari LSB standar karena penyisipan dilakukan pada layer yang lebih dalam (layer 4 hingga 6 LSB). Algoritma baru ini ditujukan agar didapat *stego-object* yang kokoh setelah disisipkan sinyal kendali mesin dan audio *stego-object* yang *imperceptible*.

Selain itu ada faktor tambahan yang bisa lebih mengoptimalkan pengaplikasian *intrasonic*. Dari segi akustik lingkungan penerima, tingkat kebisingan lingkungan dan hardware yang ditanamkan pada mesin penerima.

Daftar pustaka

[1]Fox, Barry, NewScientist.com, "TV sounds can hide machine instructions" URL: <http://www.futuretalk.org/01/quarter3/09021348.html>. Tanggal akses 20 Maret 2009.

[2] Cvejic, Nedelko. Algorithms For Audio Watermarking and Steganography.2004. University of Oulu, Finland.

[3]Cvejic, Nedeljko & Seppanen, Tapio. Increasing Robustness of LSB Audio Steganography by Reduced Distortion LSB Coding. University of Oulu, Finland.

[4]Munir, Rinaldi. Steganografi.2006. Departemen Teknik Informatika, Institut Teknologi Bandung.