

# Novel Digital Audio Watermarking

Herianto

Jurusan Teknik Informatika ITB, Bandung 40132, email: if14077@students.if.itb.ac.id

**Abstract** – Makalah ini membahas bagaimana sebuah novel digital audio watermarking mampu melindungi konten digital audio dari penyalahgunaan pembajakan copyright. Watermarking ini sendiri dibentuk dengan memakai Audio Psychoacoustic Model dan DSSS(Direct Sequence Spread Spectrum) Method. Setelah watermarking disisipkan ke dalam sebuah konten digital audio, skema khusus digunakan untuk mengekstrak informasi copyright tanpa mengakses sinyal asli dengan menggunakan sebuah prosedur penjernihan dalam model Linear Predictive Coding. Beberapa test juga dilakukan ke beberapa sampel konten audio digital untuk mengevaluasi performansi dan kualitas watermarking tersebut.

**Kata Kunci** : watermarking, Audio Psychoacoustic Model Direct Sequence Spread Spectrum, Linear Predictive Coding.

## 1. PENDAHULUAN

Dalam beberapa dasawarsa terakhir, konten multimedia dalam bentuk digital sangat berkembang pesat baik dari segi ketersediaan maupun kualitas. Namun maraknya pembajakan yang dilakukan orang-orang yang tidak punya hak akan konten tertentu membuat kekhawatiran yang besar bagi penyedia dan pemilik sebuah konten audio digital. Berbeda dari tiruan tape analog, tiruan data digital sangat mirip dengan konten asli dan hamper tidak mengalami degradasi kualitas serta tidak mempunyai batasan terhadap kuantitas kopian yang bisa diciptakan. Hal ini didukung dengan begitu banyaknya peralatan yang mampu melakukan pembajakan dengan harga yang sangat terjangkau.

Salah satu pendekatan yang bisa dilakukan untuk pengamanan konten audio digital adalah dengan enkripsi dekripsi dalam kriptografi. Tapi hal ini dinilai terlalu rumit dan kurang solutif dalam penyelesaian masalahnya karena konten yang dienkripsi harus didekripsi dahulu sebelum bisa digunakan. Ketika konten tersebut sudah tidak dalam bentuk terenkripsi, penyedia dan pemilik konten tidak bisa lagi membuktikan copyright kepemilikan mereka.

Digital Audio Watermarking kini menjadi pilihan terbaik sebagai metode baru untuk melindungi konten tersebut dari pembajakan. Sebuah watermarking digital adalah sebuah sinyal yang disisipkan ke dalam konten audio asli yang nantinya bisa diekstrak ataupun dideteksi. Watermarking dimaksudkan akan dimasukkan secara permanen sehingga pengguna yang sah bisa dengan mudah mengaksesnya.

## 2. PEMBAHASAN

### 2.1. Analisis Kebutuhan

Berikut adalah kebutuhan-kebutuhan yang harus bisa dipenuhi dalam pemberian watermarking ke konten digital audio :

- a. Perceptual transparency  
Proses penyisipan watermark tidak boleh menunjukkan benda apapun yang bisa dikenali dengan jelas, atau dengan kata lain watermark tidak mempengaruhi kualitas sinyal asli.
- b. Robustness  
Robustness merupakan salah satu isu desain algoritma yang utama karena menentukan kecenderungan algoritma tersebut terhadap distorsi data yang muncul melalui pengolahan data. Watermark yang diberikan harus 'robust' dalam pengolahan sinyal yang lazim termasuk digital-to-analog dan analog-to-digital conversion, linear dan non-linear filtering, kompresi dan scaling.
- c. Security  
Watermark harus bisa resisten terhadap deteksi pembajakan dan juga diharapkan bisa menyulitkan dari usaha pemalsuan. Keamanan watermark memegang alamat kerahasiaan dari informasi yang disisipkan. Ketika kerahasiaan itu ingin dibuka, kunci harus digunakan dalam proses penyisipan dan ekstraksi.
- d. Recovery  
Proses ekstraksi watermark biasanya tidak perlu mengakses sinyal asli karena ukuran data yang cukup besar, hal ini disebut juga blind watermarking. Pada kenyataannya, properti ini bermanfaat untuk beberapa lingkungan seperti broadcasting dan portable players. Selain dalam beberapa kasus khusus, informasi yang disisipkan harus bisa di-recover tanpa sinyal asli.
- e. Data rate  
Skema watermarking harus mempunyai data rate yang cukup untuk mendukung informasi copyright. Informasi copyright terdiri dari informasi copy control, usage, dan serial number. Untuk memenuhi syarat informasi copyright minimal dibutuhkan beberapa puluh bit.

### 2.2 Algoritma novel audio watermarking

Skema penyisipan watermarking yang dijelaskan mampu menerapkan perceptual transparency setelah proses penyisipan dengan mengeksploitasi efek

penyamaran pada sistem pendengaran manusia seperti skema yang dirujuk oleh Swanson, Ricardo Garcia dan L. Boney. Pendekatan ini dapat diaplikasikan terhadap sinyal lain seperti image dan video. Pada prosedur pendeteksi dengan implementasi whitening atau de-correlation, sebelum correlation, skema ini tidak membutuhkan sinyal audio asli untuk mengekstrak informasi audio.

Gambar 1 pada prinsipnya mengilustrasikan bahwa kita bisa memandang kebanyakan correlation berdasarkan proses penyisipan ketika menambahkan sinyal tambahan ke sinyal asli untuk mendapatkan sinyal watermark. Gambar 1 dapat direpresentasikan sebagai :

$$y(n) = x(n) + w(n) \quad (1)$$

dimana  $y(n)$  adalah sinyal audio yang sudah diberi watermark,  $x(n)$  sinyal asli dan  $w(n)$  adalah sinyal watermark itu sendiri. Untuk menampilkan pendeteksian watermark dengan correlation, hasil akhir nilai  $d$  akan menjadi :

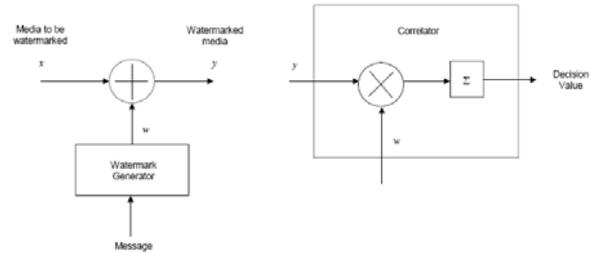
$$d = d_x + d_w \\ = \frac{1}{N} \sum_i x(i)w(i) + \frac{1}{N} \sum_i w^2(i). \quad (2)$$

Dari (2) jelas bahwa  $d_w$  adalah energy daripada watermark  $w(n)$  dan nilai  $d_x$  yang diharapkan adalah 0. Berdasarkan teorema central limit, jika  $N$  cukup besar dan jika kontribusi dari jumlahnya independen,  $d_x$  merupakan distribusi Gaussian. Jika ambang untuk deteksi watermark diset  $T = d_x/2$ , probabilitas dari false negative  $P_-$  (watermark ada, tapi detector menjawab tidak ada) akan sama dengan probabilitas false positive  $P_+$  (watermark tidak ada, tapi detector menjawab ada) :

$$P_+ = P_- = \frac{1}{2} \operatorname{erfc} \left( \frac{d_w}{\sqrt{8\sigma_x}} \right), \quad (3)$$

dimana  $\operatorname{erfc}()$  adalah fungsi complementary error. Dari (3) kita akan menggambar kondisi preliminary berikut untuk mengurangi probabilitas false.

1. Probabilitas false positive dan negative menurun ketika kekuatan watermark  $d_w$  yang akan
2. disisipkan meningkat. Kekuatan ini terbatas dan tergantung pada perceptual karakteristik dari sinyal audio.
3. Probabilitas false positive dan negative menurun ketika estándar deviasi  $\sigma_x$  meningkat. Ini dapat diambil dengan mengaplikasikan whitening atau de-correlation sebelum correlation di prosedur detection. Aplikasi prosedur whitening akan menghapus correlation dalam sinyal.



**Gambar 1 Penyisipan dan Pendeteksian Watermark**

### 2.3 Penyisipan Watermarking

Skema penyisipan watermark berdasarkan metode Direct Sequence Spread Spectrum (DSSS). Informasi khusus yang akan disisipkan dimodulasi dengan pseudo-noise (PN) sequence. Sinyal spread spectrum kemudian dibentuk dalam frekuensi domain dan dimasukkan ke dalam sinyal audio asli. Kekuatan penyisipan akan menentukan energy sinyal watermark dan dapat dipertimbangkan sebagai energy dari noise yang ditambahkan ke sinyal audio.

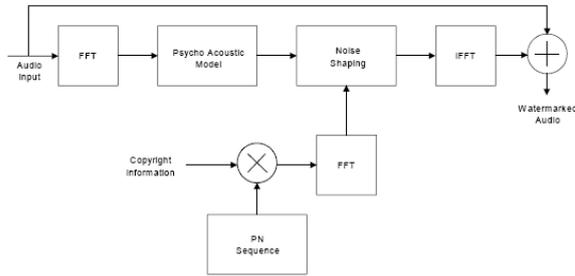
Goal dari schema ini adalah mendesain fungsi weighting sehingga energy watermark bisa dimaksimalkan terhadap kebutuhan akan acceptable distortion yang maksimal juga. Watermark diadaptasi sehingga energy maksimal dan perangkat pendengaran dijaga serendah mungkin. Masking model yang digunakan adalah model yang didefinisikan pada ISO-MPEG Audio Psychoacoustic Model. Fenomena frekuensi domain dimana sinyal low level dapat dibuat tidak terdengar merupakan masking terbesar yang dalam critical band dan masking ini efektif untuk sebuah lesser degree dalam neighboring bands. Kurva ambang masking final yang didapat dari menganalisis segmen audio diperkirakan dengan 12<sup>th</sup> order all-pole filter. Filtering ini kemudian diaplikasikan ke PN sequence menggunakan koefisien filter yang didapat dari all-pole modeling untuk membentuk sinyal watermark agar imperceptible dalam domain frekuensi. Ambang masking ini merupakan limit bagi pemaksimalan energy watermark sewaktu kualitas perceptual audio.

Dalam proses penyisipan, penyisipan operasi terhadap segmen pendek dari sinyal audio dilakukan secara berulang. Setiap satu segmen ini kemudian dinamakan sebagai frame. Diagram dari skema audio watermarking terlihat pada Gambar 2. Penggunaan dari Psychoacoustic Model dan metode DSSS, sistem kerja penyisipan watermark seperti berikut :

1. Mengkalkulasi ambang masking pada current analisis frame audio menggunakan Psychoacoustic model dengan analisis audio frame berukuran 1024 samples dan 1024 poin FFT
2. Membangkitkan PN sequence dengan panjang 1024
3. Melakukan FFT terhadap informasi copyright yang dimodulasi oleh PN sequence.
4. Menggunakan ambang masking, membentuk sinyal watermark agar imperceptible dalam domain frekuensi.
5. Mengkomputasi invers dari FFT dari sinyal

watermark yang terbentuk.

6. Membentuk sinyal final audio yang sudah diberi watermark dengan menambahkan sinyal watermark ke sinyal audio asli dalam domain waktu.



**Gambar 2 Diagram Prosedur Penyisipan Audio Watermarking**

#### 2.4 Pendeteksian Watermark

Ketika mendesain sistem pendeteksi watermark, performansi dan robustness yang diinginkan dari sistem harus dipertimbangkan. Watermark harus bisa dideteksi dibawah operasi proses sinyal biasa seperti konversi digital-to-analog dan analog-to-digital, linear dan nonlinear filtering, kompresi, dan scaling.

Prosedur deteksi watermark tidak perlu mengakses ke sinyal audio asli untuk mendeteksi sinyal watermark. Kebanyakan schema detector correlation berasumsi bahwa channel merupakan white Gaussian. Untuk non-white Gaussian channel, peningkatan performansi pendeteksian dengan pre-processing sebelum correlation mungkin terjadi. Aplikasi yang memakai prosedur whitening harus bisa menghapus correlation dalam sinyal untuk mencapai deteksi yang optimum. Untuk menghapus correlation dalam sunyal audio dapat digunakan sebuah model autoregresif bernama Linear Predictive Coding (LPC). Metode LPC dapat memperkirakan sinyal audio asli  $x(n)$  sebagai sebuah kombinasi linear dari  $p$  sampel audio yang lama seperti di bawah ini :

$$x(n) \approx a_1x(n-1) + a_2x(n-2) + \dots + a_px(n-p), \quad (4)$$

dimana koefisien  $a_1, a_2, \dots, a_p$  diasumsikan konstan terhadap analisi segmen audio. (4) dikonvert kedalam sebuah persamaan dengan memasukkan sebuah nilai error  $u(n)$ , sehingga :

$$x(n) = \sum_{i=1}^p a_i x(n-i) + u(n), \quad (5)$$

dimana  $u(n)$  merupakan eksitasi atau sinyal residu dari  $x(n)$ . Penggunaan (5), sinyal audio yang sudah diwatermark direpresentasikan seperti :

$$y(n) = \sum_{i=1}^p a_i x(n-i) + u(n) + w(n). \quad (6)$$

Berdasarkan dari analisis diatas, sinyal audio yang sudah diwatermark dihitung menjadi :

$$y(n) = \sum_{k=1}^p a_k y(n-k) + m(n), \quad (7)$$

dimana  $m(n)$  merupakan sinyal residu dari sinyal audio yang diwatermark. Asumsikan bahwa dengan karakteristik analisis Linear Predictive,  $m(n)$  mempunyai karakteristik  $u(n)$  maupun  $w(n)$ . Kombinasi linear dari sampel audio lama dihitung sebagai estimasi  $\tilde{y}(n)$  didefinisikan seperti :

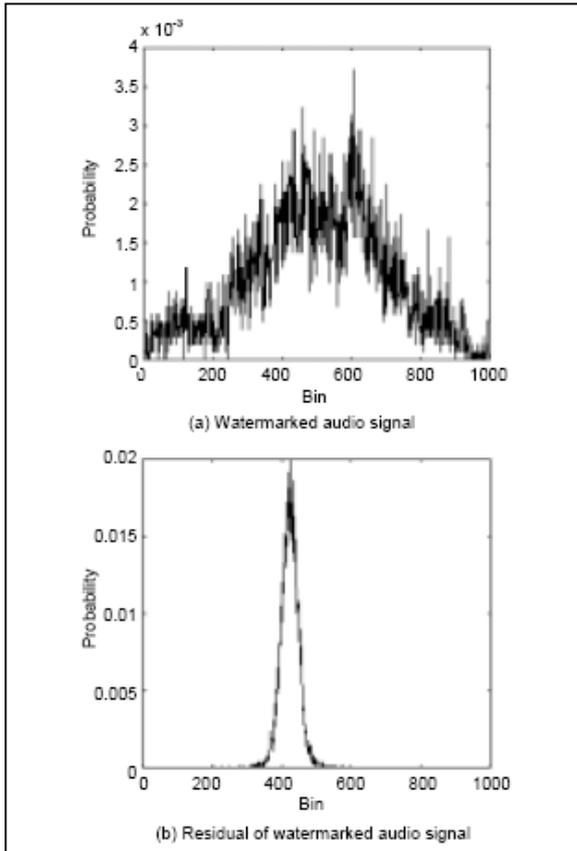
$$\tilde{y}(n) = \sum_{k=1}^p a_k y(n-k). \quad (8)$$

Sekarang error prediksinya  $e(n)$  terdefinisi sebagai :

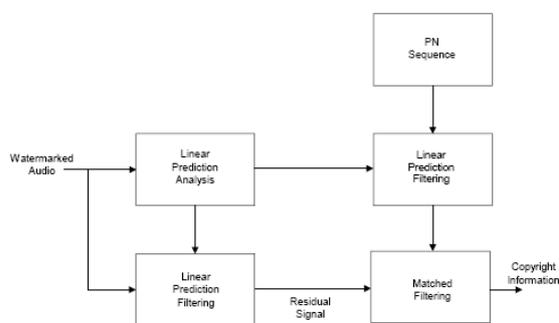
$$\begin{aligned} e(n) &= y(n) - \tilde{y}(n) \\ &= y(n) - \sum_{k=1}^p a_k y(n-k) \\ &= m(n). \end{aligned} \quad (9)$$

Persamaan (9) membuktikan bahwa sinyal  $m(n)$  bisa diestimasi dengan spektrum audio yang telah dihapus, yang mempunyai karakteristik baik sinyal eksitasi dari audio asli  $u(n)$  maupun sinyal watermark  $w(n)$  dari sinyal audio yang diwatermark dengan menggunakan analisis linear predictive.

Metode ini mentransformasi sinyal non-white audio yang diwatermark ke sebuah sinyal yang sudah whitening dengan menghapus spektrum audio. Gambar 3 dan Gambar 4 memperlihatkan fungsi densitas probabilitas pada sebuah segmen kecil dari sinyal audio ter-watermark sebelum filtering dengan LPC dan residu atau sinyal error setelah filtering dengan LPC. Fungsi densitas ini tentunya tidak halus dan mempunyai banyak variansi. Di samping itu fungsi densitas dari sinyal residu setelah filtering dengan LPC mempunyai distribusi yang lebih halus dan variansi yang lebih sedikit. Distribusi tersebut dapat dengan mudah dimodelkan menggunakan fungsi densitas probabilitas Gaussian. Dengan menggunakan sinyal residu dan PN sequence sebagai template, dapat dilakukan filtering yang cocok untuk mengekstraksi informasi yang disisipkan. Gambar 4 menampilkan diagram schema pendeteksian watermark.



Gambar 3 Fungsi densitas Probabilitas sebelum dan sesudah filtering



Gambar 4 Diagram Prosedur Pendeteksian Audio Watermark

#### 4. PENGUJIAN

Pengujian dilakukan terhadap enam file audio yang masing-masing berdurasi 15 detik. Sinyal audio yang digunakan disampel pada 44.1 kHz dengan 16 bits/sampel menggunakan sound card YAMAHA DS241. Informasi yang disisipkan ke audio pada rate 128 bits per 15 detik. Gambar 5 menampilkan sebuah frame dari sinyal audio dan sinyal watermarking yang akan disisipkan.

##### 4.1 Pengujian Robustness

Pengujian merujuk ke SDMI (Secured Digital Music Initiative) dengan langkah-langkah :

1. D/A, A/D conversion

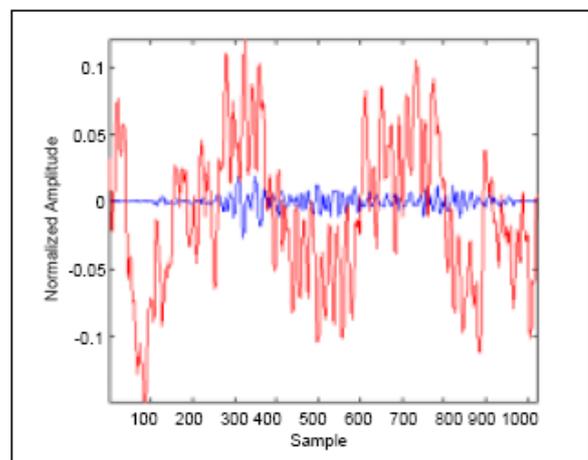
Menggunakan sound card 16-bit full duplex dan

SONY TCD-D8 DAT. Pertama sinyal audio yang ter-watermark di dalam DAT dikirim ke PC. Proses konversi berlangsung 2 kali.

2. Mencampur dan mengkompresi amplitude  
Sinyal audio dengan two-channel stereo dicampur menjadi channel mono signal lalu amplitude sinyal tersebut dikompres dengan non-linear factor.
3. Band-pass filtering  
Menggunakan 2<sup>nd</sup> order Butterworth filter yang mempunyai frekuensi pemotongan 100 Hz dan 6 kHz.
4. Echo addition  
Sebuah sinyal echo dengan delay 100 ms dan decay 50% ditambahkan.
5. MPEG compression  
Memakai MPEG-1 Audio Layer 3 dengan bit rate 64 kbps/stereo dan MPEG-2 AAC Audio Coding dengan bit rate 128 kbps/stereo.
6. Noise addition  
Menambahkan white noise dengan sebuah level konstan 38 dB ke audio ter-watermark dibawah level power dari sinyal audio
7. Equalization  
Menggunakan 10-band equalizer dengan karakteristik :  
Frequency [Hz]: 31, 62, 125, 250, 500, 1000, 2000, 4000, 8000, 16000  
Gain [dB]: -6, +6, -6, +6, -6, +6, -6, +6, -6, +6

Table 1 List item musik yang diuji

| Music Items                           | Characteristics    |
|---------------------------------------|--------------------|
| Maiden Voyage – Leny Andrade          | Base, Female vocal |
| Played Twice – The Fred Hersch Trio   | Piano, Drum, Base  |
| I Love Paris – Johnny Frigo           | Violin             |
| Sweet Georgia Brown – Monty Alexander | Drum, Base         |
| Grandma's Hand – Livingston Taylor    | Male a cappella    |
| Flute Concerto in D – Vivaldi         | Flute Concerto     |



Gambar 5 Frame sinyal audio dan hasil watermark

**Table 2 Hasil deteksi pada pengujian robustness**

| Type of Attack        | Bit Error | BER (%) |
|-----------------------|-----------|---------|
| D/A, A/D              | 10/768    | 1.30    |
| Mix Down              | 0/768     | 0.00    |
| Amplitude Compression | 0/768     | 0.00    |
| Band-Pass Filtering   | 18/768    | 2.34    |
| MPEG-1 Layer3         | 23/768    | 2.99    |
| MPEG-2 AAC            | 7/768     | 0.91    |
| Echo Addition         | 0/768     | 0.00    |
| Noise Addition        | 21/768    | 2.73    |
| Equalization          | 0/768     | 0.00    |

**Table 3 List musik pada pengujian**

| Music Items         | Diffgrade | # of Transparent Items |
|---------------------|-----------|------------------------|
| Maiden Voyage       | 0.02      | 8                      |
| Played Twice        | -0.01     | 6                      |
| I Love Paris        | 0.05      | 8                      |
| Sweet Georgia Brown | 0.13      | 7                      |
| Grandma's Hand      | -0.02     | 5                      |
| Flute Concerto      | 0.01      | 7                      |

#### 4.2 Pengujian Subjective Listening

Dilakukan berdasarkan double blind triple stimulus dengan reference tersembunyi. Pengujian ini dirangkum dalam Tabel 3 di bawah Diffgrade dan # dari Transparent Items. Diffgrade setara dengan rating subjective yang diberikan ke item yang diuji dikurangi rating yang diberikan kepada reference tersembunyi. Diffgrade yang mendekati 0.00 mengindikasikan kualitas tinggi. Diffgrade dapat bernilai positif yang mengindikasikan ketidakbenaran identifikasi dari item yang diwatermark misalnya kualitas item tersebut transparent. Skalanya dibagi lima : imperceptible(>0.00), not annoying(0.00 to -1.00), slightly annoying(-1.00 to -2.00), annoying(-2.00 to -3.00) dan very annoying(-3.00 to -4.00).

## 5. KESIMPULAN

Makalah ini menggambarkan algoritma baru untuk digital audio watermarking. Tujuan yang ingin dicapai dalam schema penyisipan watermark adalah prteceptual transparency setelah proeses penyisipan dengan mengeksploit efek masking dari sistem pendengaran manusia. Schema penyisipan ini mengadapatasikan watermark sehingga energy dari watermark itu maksimal dibawah constraint dari penjagaan perangkat pendengaran yang serendah mungkin. Prosedur pendeteksi mengekstrak informasi copyright tanpa mengakses sinyal aslinya dengan menjalankan whitening atau de-correlation sebelum correlation. Prosedur whitening mampu secara efektif menghapus correlation dalam sinyal. Hasil pengujian menunjukkan bahwa chema watermarking adalah robust terhadap proses penyerangan sinyal biasa seperti mencampur, kompresi amplitude dan kompresi data.

## DAFTAR REFERENSI

- [1] M. Swanson, M. Kobayashi, and A. Tewfik, "Multimedia Data Embedding and Watermarking Technologies," *Proc. of IEEE*, vol.86, no. 6, June 1998, pp. 1064-1087.
- [2] M. Swanson, B. Zhu, A. Tewfik, and L. Boney, "Robust Audio Watermarking Using Perceptual Masking," *Signal Processing*, vol.66, no. 3, May 1998, pp. 337-355.
- [3] Recardo A. Garcia, "Digital Watermarking of Audio Signals Using a Psychoacoustic Auditory Model and Spread Spectrum Theory," *Proc. of 107th AES Convention*, preprint 5073, Sept. 1999.
- [4] Munir, Rinaldi, Diktat Kuliah IF5054 Kriptografi, Institut Teknologi Bandung, 2007.
- [5] <http://www.library.stttelkom.ac.id/istela>