

Penyisipan Citra pada Audio dengan Kode PN Terdistribusi Gaussian

GELAR BUDIMAN¹, SUCI AULIA², I NYOMAN APRAZ RAMATRYANA³

^{1,3} Fakultas Teknik Elektro, Universitas Telkom

² Fakultas Ilmu Terapan, Universitas Telkom

Email : gelarbudiman@telkomuniversity.ac.id

Received 15 Februari 2019 | *Revised* 1 Maret 2019 | *Accepted* 24 Maret 2019

ABSTRAK

Pada makalah ini, perancangan audio watermarking memanfaatkan kode PN yang terdistribusi Gaussian atau Normal dengan menggunakan citra sebagai watermark yang disisipkan pada audio. Watermark yang berupa citra biner diubah ke dalam vektor 1 dimensi, kemudian dijumlahkan dengan kode PN terdistribusi normal yang disaring dengan filter psikoakustik. Setelah itu, sinyal dikalikan dengan faktor gain α sebelum dijumlahkan dengan host audio untuk mendapatkan watermarked audio. Hasil dari simulasi menunjukkan bahwa sistem memiliki kapasitas watermark yang tinggi pada 689.06 bps, imperseptibilitas yang baik pada SNR > 26 dB, dan tahan terhadap serangan LPF mulai frekuensi cut off 6 kHz keatas, serangan Additive Noise mulai 40 dB keatas, resampling pada rate 16 kHz, LSC 1% - 10%, dan kompresi MP3 untuk rate 192 kbps.

Kata kunci: Audio Watermarking, Kode PN, distribusi normal, filter psikoakustik

ABSTRACT

In this paper, the design of audio watermarking utilizes PN code that is Gaussian or Normal distributed by using the image as a watermark inserted in the audio. The watermark in the form of binary images is converted into a 1-dimensional vector, then summed up with a normally distributed PN code filtered by a psychoacoustic filter. After that, the signal is multiplied by α gain factor before adding it to the audio host to get the watermarked audio. The result of the simulation shows that the system has a high watermark capacity at 689.06 bps, good imperceptibility at SNR > 26 dB, and withstand LPF attacks starting from 6 kHz cut-off frequency and above, Additive Noise attacks from 40 dB up, resampling at 16 kHz, LSC 1% - 10%, and MP3 compression for 192 kbps rate.

Keywords: Audio Watermarking, PN code, normal distribution, psychoacoustic filter

1. PENDAHULUAN

Pertumbuhan internet yang makin pesat menyebabkan pertukaran multimedia terjadi lebih cepat. Pelanggaran hak cipta menjadi salah satu masalah yang timbul akibat pesatnya pertumbuhan internet. Salah satu bentuk multimedia yang paling sering dilanggar hak ciptanya adalah audio. *Digital watermarking* adalah suatu teknologi penyisipan *watermark* untuk melindungi hak cipta. Agar dapat melindungi hak cipta, *audio watermarking* harus memenuhi syarat *imperceptibility*, yaitu *watermark* tidak mempengaruhi kualitas dari audio aslinya setelah *watermark* disisipkan. *Robustness* atau ketahanan yaitu *watermark* yang telah disisipkan tidak bisa dihilangkan oleh pembajak menggunakan proses operasi dan serangan umum seperti pemotongan, pengompresan (Singh & Chadha, 2013). *Capacity* yaitu kapasitas jumlah bit yang dapat disisipkan pada sinyal audio pada waktu tertentu. *Robustness* dan *imperceptibility* adalah 2 aspek yang paling penting diantara keempat syarat *watermarking*, sehingga *watermark* mutlak harus memiliki 2 aspek tersebut. *Watermarking* digunakan untuk mengidentifikasi pemilik, pembayaran pada pemilik paten, dan autentikasi dengan menentukan apakah data sudah diubah dalam bentuk apapun dari bentuk aslinya. *Watermark* harus disisipkan pada data sedemikian sehingga *watermark* tidak terasa oleh pengguna. *Watermark* juga harus tidak terdengar ataupun terlihat untuk mencegah deteksi dan penghapusan oleh orang yang tidak berwenang. Selain itu *watermark* juga harus memiliki karakteristik yang mirip dengan sinyal aslinya dan harus tahan terhadap manipulasi atau operasi sinyal standar seperti, *filtering*, kompresi, *re-sampling*, dan lainnya (Dutta, Bhattacharyya, & Kim, 2009). *Watermarking* dapat dibedakan menjadi 2 dalam hal pengetahuan terhadap *host* di penerima, yaitu *blind watermarking* dan *non blind watermarking* (Khalil & Adib, 2015). *Blind watermarking* adalah proses *watermarking* dengan kondisi di penerima tidak perlu mengetahui *host audio* untuk mengekstraksi *watermark* yang telah disisipkan. Sementara *non blind watermarking* adalah proses *watermarking* dengan kondisi di penerima harus mengetahui *host audio* untuk mengekstraksi *watermark* yang telah disisipkan (Cox, Miller, Bloom, Fridrich, & Kalker, 2008).

Penelitian mengenai penyembunyian data pada audio atau *audio watermarking* telah banyak dilakukan dan telah banyak dipublikasikan. Sebagian besar peneliti yang mempublikasikan penelitian mengenai *audio watermarking* membahas bagaimana perancangan sistem *audio watermarking* yang dibuat dapat memenuhi kriteria *imperceptible*, tahan, dan kapasitas tinggi. Namun sampai saat ini permasalahan mengenai pemenuhan kriteria *audio watermarking* yang baik masih terus dikembangkan dan diperbaiki mengingat perkembangan serangan terhadap *watermarked audio* yang bertambah variasinya dari waktu ke waktu, kebutuhan manusia agar kualitas audio dapat lebih baik lagi, dan kebutuhan manusia agar informasi yang disembunyikan dapat jauh lebih besar kapasitasnya (Bajpai & Kaur, 2016). Pada penelitian ini dibatasi metode penelitian *audio watermarking* 1 kanal atau mono yang dilakukan dengan metode penggunaan kode PN yang hampir serupa dengan proses penyisipan menggunakan modulasi *spread spectrum* namun kode PN terdistribusi Gaussian serta penggunaan *audio watermarking* yang bersifat *non blind watermarking* atau *watermarking* dengan kondisi bahwa pada proses ekstraksi *host audio* asli harus diketahui. Tujuan dari penelitian yang dilakukan adalah merancang skema *audio watermarking* dengan kode PN yang terdistribusi Gaussian untuk mendapatkan ketahanan *watermarking* yang lebih kuat terhadap serangan yang merusak *watermark* karena sifat distribusi Gaussian yang memiliki pemusatan sinyal pada titik tengah dari amplituda sinyalnya. Titik tengah atau titik rata-rata pada Gaussian merupakan bagian sinyal yang tidak mudah diserang sehingga hipotesis hal ini memberikan peningkatan ketahanan yang lebih baik pada skema *watermarking* yang dirancang dengan tetap mempertahankan imperseptibilitas dan kapasitas *watermark* yang disisipkan.

Penelitian *watermarking* dengan modulasi *spread spectrum* telah dimulai dengan *watermarking* pada citra (**Cox, Kilian, Leighton, & Shamoon, 1997**). Pada makalah tersebut Cox menjelaskan konsep *spread spectrum* dan penerapannya untuk menyisipkan data pada citra. Konsep *spread spectrum* yang dijelaskannya adalah bahwa bit *watermark* dimodulasi dengan *spread spectrum* kemudian hasil modulasinya dikalikan dengan faktor pengali α atau faktor penguat yang berpengaruh pada ketahanan dan imperseptibilitas *watermark* yang disisipkan. Kelemahan dari metode yang disampaikan Cox adalah besarnya interferensi pada *host audio* terhadap *watermark* yang disisipkan, sehingga berpengaruh pada ketahanan *watermark* yang tidak terlalu baik dalam menghadapi serangan *watermarking*. Darko dalam (**Kirovski & Malvar, 2003**) pertama kali mempublikasikan penelitian mengenai *audio watermarking* dengan konsep modulasi *spread spectrum* pada makalah bereputasi internasional. Darko menjelaskan konsep *spread spectrum* pada *audio watermarking* dengan komprehensif baik teori maupun penerapannya, namun tidak menjelaskan kinerjanya secara praktikal, baik imperseptibilitas dan ketahanan *watermarking*-nya. Malvar mempublikasikan makalah mengenai *audio watermarking* dengan *spread spectrum* yang lebih baik pada (**Malvar & Florêncio, 2003**). Malvar memperbaiki faktor penguat α yang tidak fleksibel pada makalah sebelumnya, dengan membuat formulasi sedemikian sehingga nilai α berubah-ubah sesuai fluktuasi daya sinyal rata-rata dari *host audio* dalam setiap segmennya. Kinerja yang ditunjukkan Malvar memberikan perbaikan dibandingkan (**Kirovski & Malvar, 2003**), namun tidak signifikan. Zhang pada (**Zhang, Xu, & Yang, 2012**) memperbaiki imperseptibilitas pada *audio watermarking* dengan *spread spectrum* dengan menambahkan filter psikoakustik pada *watermark* setelah dimodulasi *spread spectrum* dengan tetap mempertahankan ketahanan *watermarking* namun kapasitas *watermark* masih sangat kecil.

Penelitian mengenai *audio watermarking* dengan kondisi *non blind watermarking* dipublikasikan pada (**Foo, 2008**) dengan metode *compression expansion* yang memberikan hasil ketahanan yang tahan terhadap serangan, namun untuk serangan MP3 dijelaskan hanya terhadap *rate* 128 kbps, dan serangan LPF pada cut off yang cukup tinggi pada 10 kHz. Makalah lainnya yang menggunakan skema *non blind audio watermarking* adalah makalah (**Panda, Gera, & Bhattacharyya, 2014**) yang menggunakan metode *Empirical Mode Decomposition* (EMD) dengan hasil ketahanan yang kurang terhadap serangan terutama serangan kompresi MP3. Pada tahun yang sama, Jane pada makalah (**Jane, Elbaşı, & İlk, 2014**) mempublikasikan makalah *non blind watermarking* dengan metode DWT-SVD, namun Jane menyisipkan data pada citra, bukan audio.

Pada makalah ini dibahas bagaimana melakukan penyisipan citra pada audio dengan kode PN yang terdistribusi Gaussian menggunakan kunci pembangkit kode PN tertentu. Pertama, *watermark* berupa citra biner yang bentuknya 2 dimensi (2-D) dikonversi bentuknya menjadi 1 dimensi (1-D), kemudian dari vektor 1-D tersebut masing-masing bit ditebar sejumlah panjang yang sama dengan panjang kode PN yang dibangkitkan, setelah itu setiap bit yang telah ditebar itu dijumlahkan dengan kode PN yang telah difilter dengan filter psikoakustik yang kemudian hasilnya dikalikan dengan penguat koefisien α , lalu sinyal tersebut dijumlahkan dengan sinyal audio yang telah dibaca dan telah dikonversi ke mono, maka hasilnya adalah *watermarked audio* sehingga SNR, ODG dan kapasitas *watermark* dapat dihitung. Sementara pada proses ekstraksi, dilakukan penghilangan komponen *direct current* (DC) pada *watermarked audio* terlebih dahulu, kemudian normalisasi sehingga level daya sinyalnya sama dengan level daya sinyal *host* asli, setelah itu baru dilakukan pengurangan dengan *host audio* asli, dengan posisi pengurangannya adalah *host audio* asli. Setelah itu dilakukan integrasi dengan panjang integrasi sama dengan panjang kode PN yang ditebar pada saat proses penyisipan, dan akhirnya *watermark* didapat dengan melakukan *decision*

terhadap nilai tengah dari hasil integrasi, jika hasil integrasi lebih kecil dari nilai tengah atau rata-rata hasil integrasi, maka diperoleh bit *watermark* adalah 0, jika tidak maka bit *watermark* adalah 1, setelah itu dilakukan konversi 1-D ke 2-D, maka diperoleh citra *watermark* dan *bit error rate* (BER) dapat dihitung.

Kode PN merupakan kode yang bersifat ortogonal dan dalam buku **(Shulze & Luders, 2005)** pada bab CDMA atau pada **(Yani, Rizal, & Prasetya, 2008)**, kode PN dihasilkan dari teknik atau proses *Linear Feedback Shift Register* (LFSR). Proses LFSR akan menghasilkan kode *random* bit 0 dan 1 secara acak namun ortogonal satu sama lain, setelah itu pada **(Xiang, Natgunanathan, Peng, Hua, & Liu, 2018)** kode acak 0 dan 1 tersebut dilakukan proses penguraian dengan prosedur Gram Schmidt menghasilkan kode yang terdiri dari bilangan desimal dan terdistribusi Gaussian dengan persamaan *probability density function* (pdf) terdistribusi Gaussian sesuai yang ditunjukkan pada Persamaan (1). Umumnya dari Persamaan (1), $\sigma = 1$ dan $\mu = 0$, sehingga Persamaan (1) dapat disederhanakan menjadi Persamaan (2). Ilustrasi nilai kode PN dan distribusinya dapat dilihat pada Gambar 1.

$$p(x \mid \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1)$$

dimana :

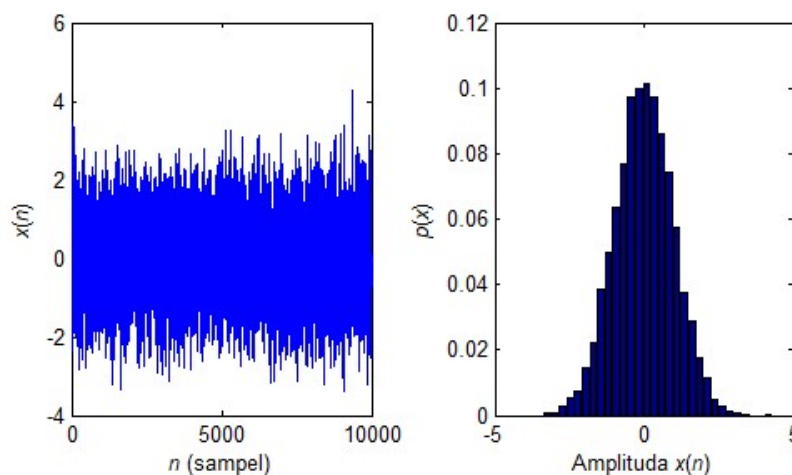
p = nilai pdf dari kode PN yang dibangkitkan

x = variabel *random* diskrit yang merupakan kode PN yang dibangkitkan

σ = standar deviasi dari x

μ = rata-rata dari x

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (2)$$



Gambar 1. Kode PN (kiri) dan Distribusi Gaussian pada Kode PN (kanan)

Filter psikoakustik yang digunakan pada penelitian ini berfungsi untuk meningkatkan imperseptibilitas *watermark* yang disisipkan atau meningkatkan kualitas *watermarked audio* secara perseptual pendengaran manusia yang didistribusikan ke *end user* yang akan menikmati audio tersebut. Filter psikoakustik hampir mirip fungsinya dengan *band stop filter* (BSF) yang akan menyaring suara yang dominan di frekuensi tengah atau frekuensi yang didominasi oleh sensitivitas pendengaran manusia, sehingga kenyaringan *watermark* pada frekuensi dimana pendengaran manusia sensitif akan ditekan yang berkonsekuensi interferensi *watermark* terhadap *host audio* akan berkurang atau imperseptibilitas *watermark*

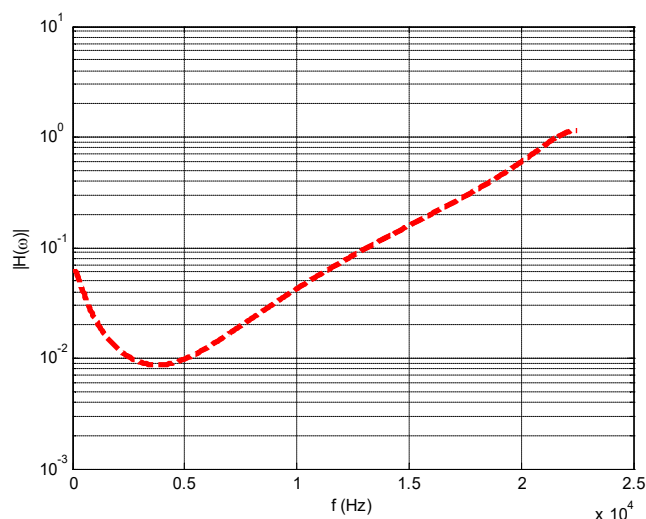
akan makin baik dan kualitas *watermarked audio* akan bertambah. Secara umum persamaan yang merepresentasikan filter psikoakustik adalah filter *infinite impulse respon* (IIR) seperti yang dijelaskan oleh **(Robinson, 1999)** dan **(Budiman, Suksmono, Danudirdjo, Usman, & Shin, 2016)** yang ditunjukkan pada Persamaan (3).

$$H(z) = \frac{0.07 - 0.147z^{-1} + 0.128z^{-2} - 0.053z^{-3} + 0.009z^{-4}}{1 - 0.15z^{-1} - 0.76z^{-2}} \quad (3)$$

$$z = e^{j\omega} \quad (4)$$

$$\omega = 2\pi \frac{f}{F_s} \quad (5)$$

dimana z merupakan variabel kompleks dalam domain frekuensi, f merupakan frekuensi dalam satuan Hz, ω merupakan frekuensi sudut dengan satuan radian/sampel dan F_s merupakan frekuensi sampling dari audio yang akan difilter dalam satuan sampel/detik. Respon magnituda filter dari Persamaan (3) ditunjukkan dengan grafik pada Gambar 2 **(Budiman et al., 2016)**.



Gambar 2. Respon Magnituda Filter Psikoakustik (Budiman et al., 2016)

Makalah ini disampaikan dengan struktur berikut, Bab 1 membahas latar belakang penelitian, permasalahan penelitian yang ditunjukkan dari review literature terdahulu yang terkait topik penelitian ini, usulan metode dan hipotesisnya serta konsep dasar yang mendasari penelitian. Bab 2 membahas metodologi penelitian yang menjelaskan langkah dan tahap penelitian yang dilakukan beserta proses penyisipan dan ekstraksi *watermark*. Bab membahas hasil penelitian dan analisis yang menjadi bahan diskusi. Dan bab 4 yang menunjukkan kesimpulan dari makalah ini.

2. METODOLOGI PENELITIAN

Makalah ini merupakan hasil penelitian audio watermarking yang memanfaatkan distribusi Gaussian pada kode PN yang disematkan pada audio asli. Untuk mengambil kembali watermarknya, diperlukan *host audio* asli pada proses ekstraksi yang akan dibandingkan dengan *watermarked audio* dan dilakukan proses *decision* memanfaatkan kelebihan dari

distribusi Gaussian sebagai distribusi dari bilangan random yang terdapat pada kode PN yang dibangkitkan. Perancangan dan tahapan metode penelitian dijelaskan berikut.

2.1. Identifikasi Masalah

Permasalahan utama yang mendasari penelitian ini adalah terkait tiga hal, antara lain ketahanan *watermark* terhadap serangan yang masih rendah, imperseptibilitas yang rendah, dan kapasitas *watermark* yang rendah. Pada uraian di Bab 1, telah dijelaskan beberapa literature dari penelitian terdahulu yang masih menghasilkan masalah tersebut. Untuk menyelesaikan masalah tersebut maka diperlukan pendalaman materi terkait sebelum melakukan penelitian. Materi terkait yang telah dialami adalah:

- a. Konsep *audio watermarking* untuk kasus *non blind watermarking*
- b. Kode PN dan karakteristiknya
- c. Distribusi Gaussian dan karakteristiknya
- d. Teknik *Spread spectrum* pada *Audio watermarking*

2.2. Persiapan Data

Data yang digunakan pada penelitian ini terdiri atas data gambar *watermark* dan data audio *host*. Gambar *watermark* merupakan gambar biner TelkomUniversity.png yang beresolusi 75x225 seperti ditunjukkan pada gambar 3. Sementara audio *host* yang digunakan berjumlah 5 file audio antara lain rock.wav, piano.wav, gitar.wav, drums.wav dan bass.wav.



Gambar 3. Gambar Watermark Biner

2.3. Perancangan Sistem dan Simulasi

Sistem yang dirancang terdiri atas subsistem penyisipan data *watermark*, subsistem penyerang sinyal, dan subsistem pengambilan data *watermark*. Gambar 4 menunjukkan proses atau subsistem penyisipan, sementara Gambar 5 menunjukkan proses atau subsistem pengambilan data *watermark*. Pada subsistem penyisipan data *watermark* langkah-langkah yang dilakukan adalah sebagai berikut:

- Langkah 1: Baca *host audio* dan konversi audio stereo ke mono sehingga diperoleh *host audio* yang siap untuk disisipkan atau $x(n)$.
- Langkah 2: Baca *watermark* $w(m, n)$ kemudian diubah kedalam ukuran 1 dimensi menjadi $w(n)$. Jika resolusi gambar *watermark* adalah $P \times Q$, maka ukurannya setelah diubah menjadi 1 dimensi adalah $1 \times PQ$.
- Langkah 3: Tebar setiap bit pada $w(n)$ sepanjang N menjadi $w_s(n)$ sehingga ada PQ segmen dan setiap segmennya terdiri dari N sampel. N merupakan panjang kode PN yang dibangkitkan. Prinsip penebaran bit mengambil prinsip *pulse shaping* untuk sinyal diskrit. Sehingga ukurannya menjadi $1 \times PQN$.
- Langkah 4: Bangkitkan kode PN terdistribusi Gaussian dengan ukuran kode PN adalah $1 \times N$ sehingga didapatkan $c(n)$.
- Langkah 5: Lakukan pemfilteran kode PN dengan filter psikoakustik sehingga diperoleh $c_p(n)$.

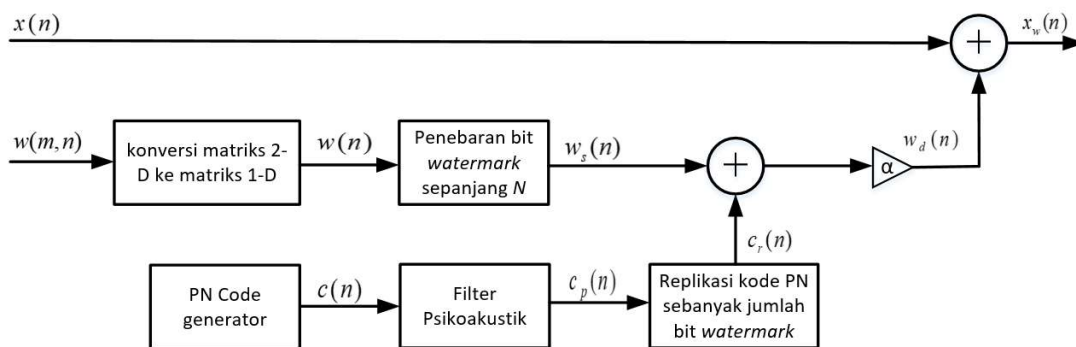
Langkah 6: Lakukan replikasi kode PN sejumlah bit $watermark$ dari $w(n)$ atau sejumlah PQ. Sehingga diperoleh $c_r(n)$ dengan ukuran $1 \times PQN$.

Langkah 7: Jumlahkan $w_s(n)$ dari langkah 3 dengan $c_r(n)$ dari langkah 6 dan kalikan dengan faktor pengali α . Maka diperoleh Persamaan (6).

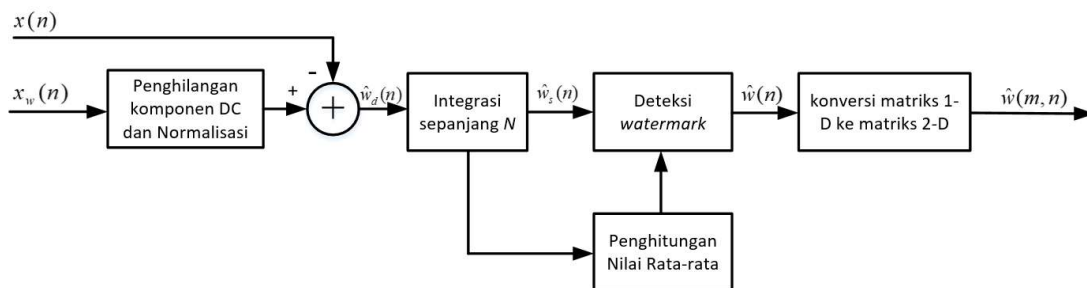
$$w_d(n) = \alpha(w_s(n) + c_r(n)) \quad (6)$$

Langkah 8: Lakukan penyisipan dengan menjumlahkan $x(n)$ dan $w_d(n)$ sehingga diperoleh $watermarked\ audio\ x_w(n)$ sesuai Persamaan (7).

$$x_w(n) = x(n) + w_d(n) = x(n) + \alpha(w_s(n) + c_r(n)) \quad (7)$$



Gambar 4. Subsistem Penyisipan *Watermark*



Gambar 5. Subsistem Ekstraksi *Watermark*

Sementara pada subsistem ekstraksi data *watermark* langkah-langkah yang dilakukan adalah sebagai berikut:

Langkah 1: Baca *host audio* atau $x(n)$ sebagai syarat yang harus dalam *non-blind watermarking*.

Langkah 2: Baca *watermarked audio* atau $x_w(n)$ dengan ukuran yang sama dengan $x(n)$.

Langkah 3: Kurangkan $x_w(n)$ oleh $x(n)$, sehingga menghasilkan $\hat{w}_d(n)$, ditunjukkan pada Persamaan (8).

$$\hat{w}_d(n) = x_w(n) - x(n) \quad (8)$$

Langkah 4: Lakukan integrasi diskrit atau penjumlahan diskrit sepanjang N pada setiap segmen dari $w_d(n)$ sehingga menghasilkan $\hat{w}_s(n)$.

Langkah 5: Lakukan penghitungan rata-rata dari $\hat{w}_s(n)$ pada setiap segmennya dan simpan sebagai $\mathcal{T}(n)$.

Langkah 6: Lakukan deteksi *watermark* dengan membandingkan setiap segmen dari $\hat{w}_s(n)$ terhadap $\mathcal{T}(n)$, jika $\hat{w}_s(n) < \mathcal{T}(n)$, maka diperoleh $\hat{w}(n) = 0$, dan jika $\hat{w}_s(n) \geq \mathcal{T}(n)$, maka diperoleh $\hat{w}(n) = 1$ sesuai Persamaan (9).

$$\hat{w}(n) = \begin{cases} 0, & \text{jika } \hat{w}_s(n) < T(n) \\ 1, & \text{jika } \hat{w}_s(n) \geq T(n) \end{cases} \quad (9)$$

Langkah 7: Lakukan konversi matriks 1-D ke 2-D dari $\hat{w}(n)$ sehingga menghasilkan gambar *watermark* hasil ekstraksi $\hat{w}(m, n)$.

2.4. Pengujian

Pengujian simulasi dilakukan dengan Matlab versi 2014A dengan spesifikasi PC adalah menggunakan prosesor i7 dan RAM 16 GB. Parameter yang diubah-ubah untuk pengaturan *trade off* antara imperseptibilitas dan ketahanan adalah α (faktor penguat *watermark*) dan N (panjang kode PN untuk setiap segmen). Serangan yang diambil untuk pengujian adalah LPF dengan frekuensi cut off 3k, 6k, dan 9k, *resampling* dengan frekuensi 22k, 11k, and 16k, *Linear Speed Change* (LSC) dengan perubahan persentase 1%, 5%, 10%, and kompresi MP3 dengan *rate* 32k, 64k, 128k, and 256k. Parameter performansi yang diukur pada penelitian ini adalah parameter imperseptibilitas yang terdiri dari parameter objektif (SNR dan ODG) dan subjektif (MOS), parameter ketahanan (BER), dan kapasitas (C). Formulasi bit error *rate* atau BER dapat dilihat pada Persamaan (10).

$$BER = \frac{J_e}{J_b} \quad (10)$$

dimana J_e adalah jumlah bit *watermark* yang terdeteksi salah, dan J_b adalah jumlah seluruh bit *watermark*. Persamaan untuk parameter objektif *Signal to Noise Ratio* (SNR) dapat dilihat pada Persamaan (11).

$$SNR = 10 \log_{10} \frac{\sum_{n=0}^{L-1} x^2(n)}{\sum_{n=0}^{L-1} (x(n) - x_w(n))^2} \quad (11)$$

dimana $x(n)$ adalah sinyal *host audio* asli, L adalah panjang sinyal *watermarked audio*, $x_w(n)$ adalah sinyal *watermarked audio*, dan n adalah sampel dalam domain waktu diskrit. Sementara persamaan untuk kapasitas dapat dilihat pada Persamaan (12).

$$C = \frac{L_w F_s}{L} \quad (12)$$

dimana L_w adalah panjang bit *watermark* yang disisipkan dan F_s adalah frekuensi sampling dari *host audio*.

Tabel 1. Standar Kualitas Audio Berdasarkan ITU-R BS.1387 (ITU-R, 1998)

Nilai ODG	Nilai SDG/MOS	Keterangan
0	5	<i>Watermark</i> tidak terasa
-1	4	<i>Watermark</i> terasa namun tidak mengganggu audio
-2	3	<i>Watermark</i> terasa namun sedikit mengganggu audio
-3	2	<i>Watermark</i> terasa dan audio terganggu
-4	1	<i>Watermark</i> terasa dan audio sangat terganggu

Objective Different Grade (ODG) merupakan proses penghitungan yang kompleks untuk menghitung kualitas audio secara perseptual yang mengacu pada standar ITU-R BS.1387 pada **(ITU-R, 1998)** dengan skala nilai dari -4 sampai 0 yang dapat dilihat pada Tabel 1. *Subjective Different Grade* (SDG) atau *Mean Opinion Score* (MOS) merupakan parameter subjektif hasil penilaian secara perseptual dari survey terhadap beberapa responden yang memiliki skala nilai 1 sampai 5 yang dapat dilihat pada Tabel 1.

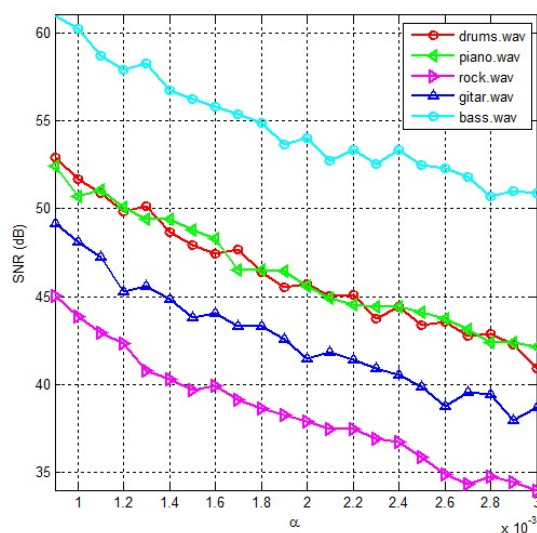
3. HASIL DAN DISKUSI

Pada bab ini dipaparkan mengenai hasil simulasi yang terdiri atas pengaruh parameter pengujian N dan α terhadap imperseptibilitas, ketahanan dan kapasitas, ketahanan *watermark* terhadap berbagai serangan, degradasi kualitas *watermark*, dan kualitas *watermarked audio* pada parameter yang telah disesuaikan.

3.1 Pengaruh Parameter terhadap Kinerja *Watermarking*

Pada subbab ini parameter α diubah-ubah untuk dilihat bagaimana pengaruhnya terhadap imperseptibilitas yang diwakili oleh SNR dan ODG, serta ketahanan yang diwakili BER. Pengujian perubahan α ini dilakukan dari batas nilai α : 0.0009 sampai 0.003 untuk jumlah sampel kode PN acak atau N adalah 128 sampel dalam setiap frame. Batas ini ditentukan sedemikian karena masih dalam batas dimana secara subjektif suara *watermark* masih *imperceptible*. Gambar 6 menunjukkan pengaruh α terhadap SNR. Hasilnya dapat dilihat bahwa kenaikan nilai α membuat adanya penurunan nilai SNR atau penurunan kualitas *watermarked audio*, dimana penurunan tersebut memiliki perbedaan nilai untuk setiap *host audio* yang berbeda-beda. Artinya nilai α yang sama tidak berarti akan menghasilkan SNR yang sama untuk *host audio* yang berbeda-beda. Imperseptibilitas untuk setiap *host audio* yang berbeda akan berbeda-beda pula.

Melihat hasil dari Gambar 6, maka pengujian berikutnya adalah bagaimana mencari nilai α yang masih berada pada level imperseptibilitas yang baik secara subjektif dimana secara teoritis nilai α pada metode yang mirip, namun dalam domain frekuensi dapat diformulasikan pada **(Harahap, Budiman, & Novamizanti, 2016)**. Pengujian dilakukan dengan menggunakan PC dan *headphone* dengan speaker pasif yang volumenya diatur pada 40% sebagaimana volume yang digunakan oleh manusia jika mendengarkan via *headphone*. Pengujian dilakukan pada nilai $N=128$. Hasilnya ditunjukkan pada Tabel 2. Pada tabel tersebut kita lihat bahwa nilai α maksimal yang masih menghasilkan suara *watermarked audio* yang *imperceptible* atau memiliki nilai SDG 5 berbeda-beda untuk setiap *host audio*-nya dengan nilai SNR dan ODG yang juga berbeda-beda.

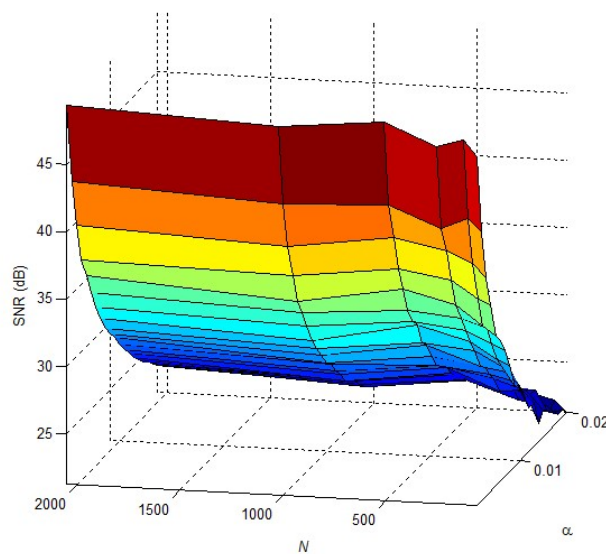


Gambar 6. Pengaruh α terhadap SNR

Tabel 2. Pengaturan Nilai α Maksimal pada Imperseptibilitas Baik (SDG>4)

<i>Host audio</i>	Alfa	SNR (dB)	ODG
drums.wav	0.002	46.52	-3.9
piano.wav	0.01	32.51	-2.84
rock.wav	0.003	33.63	-2.94
gitar.wav	0.01	27.67	-0.43
bass.wav	0.03	30.57	-2.9

Pengujian diatas merupakan pengujian pada nilai $N=128$ yang menghasilkan kapasitas *watermark* pada 344.53 bps, berikutnya dilakukan pengujian dengan mengubah-ubah nilai N dan α serta melihat pengaruhnya terhadap SNR untuk *host* gitar (.wav). Pengujian ini dilakukan dengan batas nilai α : 0.001 sampai 0.02, dan N : 64 sampai 2048. Gambar 7 dan Tabel 3 memperlihatkan hasil pengujiannya. Terlihat bahwa perubahan N tidak terlalu berpengaruh terhadap SNR, parameter yang sangat berpengaruh terhadap perubahan SNR hanyalah α . Namun dari Tabel 3 pada $\alpha =0.01$, nilai N yang makin besar akan berpengaruh terhadap berkurangnya kapasitas *watermark* maksimal yang dapat disisipkan. Jika diharapkan *watermark* memiliki kapasitas minimal 100 bps, maka N yang diperlukan untuk perancangan sistem ini harus $N \leq 256$.

**Gambar 7. Pengaruh α dan N terhadap SNR**

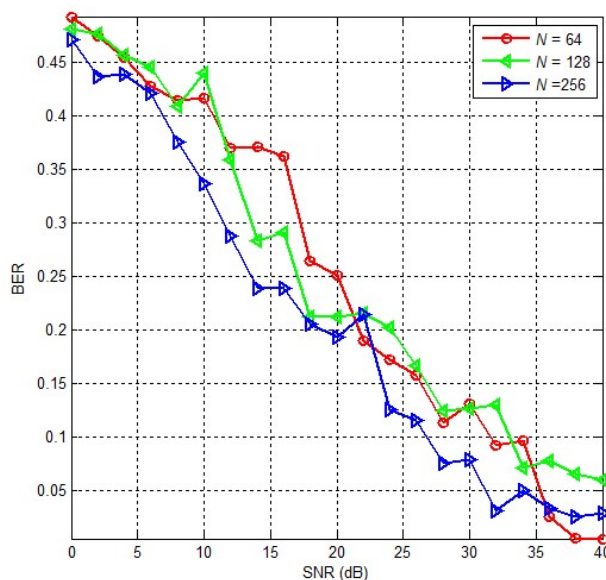
3.2 Kinerja Ketahanan Sistem

Pada subbab ini dilakukan beberapa pengujian dengan parameter α sesuai pada Tabel 2 dan $N=64$ dan 256, agar kapasitas *watermark* yang diperoleh masih lebih besar dari 100 bps. Serangan yang digunakan untuk pengujian pertama adalah serangan penambahan noise pada *watermarked audio* untuk semua *host* dengan SNR dari 0 dB sampai 30 dB pada *host audio* gitar.wav. Hasilnya adalah rata-rata BER dari semua *host* ditampilkan pada Gambar 8.

Tabel 3. Pengaruh Nilai N pada Imperseptibilitas dan Kapasitas

N	ODG	SNR (dB)	C (bps)
64	-0.80962	26.91649	689.0625
128	-0.47539	27.83535	344.5313
256	-0.2532	28.1906	172.2656
512	-0.08079	29.53925	86.13281
1024	-0.2937	28.4301	43.06641
2048	-0.15363	29.20878	21.5332

Gambar 8 memperlihatkan kondisi bahwa semakin besar daya *watermarked audio* terhadap daya noise yang direpresentasikan dengan SNR pada sumbu x maka semakin baik ketahanan *watermark* yang ditunjukkan dengan semakin rendahnya nilai BER pada sumbu y. Sementara pengaruh N disini adalah bahwa nilai N yang makin besar membuat nilai BER lebih baik walaupun tidak terlalu signifikan perbedaannya.



Gambar 8. Pengaruh Serangan *Additive Noise* Terhadap BER









Pengujian berikutnya adalah pengujian untuk semua *host audio* dengan parameter α sesuai pada Tabel 2 dan $N=64$ dan 256 , agar kapasitas *watermark* yang diperoleh masih lebih besar dari 100 bps. Serangan yang dilakukan untuk menguji sistem *audio watermarking* yang telah dirancang sesuai dengan penjelasan pada subbab 2.4. Hasil pengujian diperlihatkan pada Tabel 4. Terlihat bahwa rata-rata BER untuk $N=64$ (0.1) lebih besar daripada rata-rata BER untuk $N=256$ (0.07). Hal ini menunjukkan bahwa semakin besar N , maka semakin tahan *watermark* terhadap serangan secara keseluruhan. Hal ini menjadi *trade off* karena semakin besar N , maka dalam hal kapasitas *watermark* semakin berkurang seperti telah dijelaskan sesuai Tabel 3. Kenaikan kapasitas mengakibatkan penurunan ketahanan *watermark* terhadap serangan. Secara keseluruhan berdasarkan hasil dari Tabel 3 metode pada makalah ini tahan terhadap serangan LPF mulai frekuensi cut off 6 kHz keatas, serangan Additive Noise mulai 40 dB keatas, resampling pada *rate* 16 kHz, LSC 1% - 10%, dan kompresi MP3 untuk *rate* 192 kbps.

Tabel 4. Ketahanan Sistem dengan $N=64$ dan 256

Jenis Serangan	Parameter Serangan	BER									
		$N=64$ sampel/frame					$N=256$ sampel/frame				
		Drums	Piano	Rock	Gitar	Bass	Drums	Piano	Rock	Gitar	Bass
LPF	3k	0.30	0.00	0.28	0.20	0.00	0.10	0.00	0.44	0.12	0.00
	6k	0.14	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00
	9k	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Noise	0 dB	0.47	0.39	0.50	0.56	0.53	0.61	0.33	0.36	0.08	0.42
	20 dB	0.37	0.03	0.25	0.00	0.14	0.44	0.03	0.00	0.00	0.00
	40 dB	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Resampling	11k	0.18	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.42
	16k	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	22k	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.39
	24k	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20
Linear Speed Change	1%	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	5%	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04
	10%	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05
Kompresi MP3	32k	0.59	0.40	0.19	0.57	0.47	0.58	0.00	0.17	0.04	0.46
	64k	0.57	0.11	0.19	0.07	0.29	0.14	0.00	0.03	0.00	0.22
	128k	0.37	0.00	0.06	0.00	0.24	0.00	0.00	0.00	0.00	0.07
	192k	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Rata-rata		0.19	0.05	0.09	0.08	0.10	0.11	0.02	0.06	0.01	0.13

Nilai BER hasil pengujian ketahanan *watermark* menunjukkan kualitas gambar *watermark* yang diekstraksi dari *watermarked audio* yang diserang. Makin tinggi nilai BERnya maka makin buruk kualitas gambar *watermark* hasil ekstraksi. Namun ada batas nilai BER maksimal pada gambar *watermark* yang masih dapat diterima manusia karena secara visualisasi manusia masih dapat memahami isi dari gambar *watermark* tersebut walaupun terkena kerusakan. Batas maksimal nilai BER dari gambar *watermark* sangat tergantung pada resolusi gambar *watermark* yang disisipkan pada audio. Pada *audio watermarking* dengan metode diatas, disisipkan citra logo UniversitasTelkom.png dengan resolusi 75x225 dimana tampilan aslinya dapat dilihat pada Tabel 5 pada BER=0. Saat gambar *watermark* mengalami kerusakan dan diperoleh BER=0.01 sampai 0.22 maka terlihat bahwa gambar tersebut masih dapat dipahami oleh manusia, namun ketika gambar *watermark* memiliki tingkat kerusakan dengan BER=0.33 sudah mulai manusia sedikit memahami isi gambar tersebut. Hal ini menunjukkan bahwa jika ketahanan *watermark* memiliki BER di bawah 0.3 pada resolusi gambar 75x225 sesuai gambar UniversitasTelkom.png, maka gambar *watermark* hasil ekstraksinya masih dapat diterima, atau dengan kata lain *watermark* masih tahan terhadap serangan jika kualitas *watermark* tersebut masih berada pada kisaran BER < 30%.

Tabel 5. Degradasi Kualitas Watermark

BER	Gambar <i>Watermark</i>	BER	Gambar <i>Watermark</i>
0		0.2	
0.05		0.22	
0.1		0.33	
0.15		0.4	

4. KESIMPULAN

Makalah ini memaparkan metode *non blind audio watermarking* dengan memanfaatkan kode acak PN yang disematkan sesuai nilai *watermark* yang disisipkan. Kode PN yang disematkan terdistribusi Gaussian sehingga memberikan ketahanan yang lebih baik karena memiliki karakteristik pemusatan sinyal pada nilai rata-ratanya. Pengujian yang dilakukan memberikan hasil berupa kapasitas *watermark* yang tinggi pada 689.06 bps, imperseptibilitas yang baik pada SNR > 26 dB, dan tahan terhadap serangan LPF mulai frekuensi cut off 6 kHz keatas, serangan Additive Noise mulai 40 dB keatas, resampling pada *rate* 16 kHz, LSC 1% - 10%, dan kompresi MP3 untuk *rate* 192 kbps.

UCAPAN TERIMA KASIH

Kami sampaikan terimakasih kepada PPM Universitas Telkom yang telah mendanai penelitian dan publikasi makalah ini.

DAFTAR RUJUKAN

- Bajpai, J., & Kaur, A. (2016). A literature survey - Various *audio watermarking* techniques and their challenges. In *Proceedings of the 2016 6th International Conference - Cloud System and Big Data Engineering, Confluence 2016* (pp. 451–457). <https://doi.org/10.1109/CONFLUENCE.2016.7508162>
- Budiman, G., Suksmono, A. B., Danudirdjo, D., Usman, K., & Shin, D. H. (2016). A modified multicarrier modulation binary data embedding in audio file. *International Journal on*

- Electrical Engineering and Informatics*, 8(4), 762–773.
<https://doi.org/10.15676/ijeei.2016.8.4.5>
- Cox, I. J., Kilian, J., Leighton, F. T., & Shamoon, T. (1997). Secure spread spectrum watermarking for multimedia. *IEEE Transactions on Image Processing*, 6(12), 1673–1687. <https://doi.org/10.1109/83.650120>
- Cox, I. J., Miller, M. L., Bloom, J. A., Fridrich, J., & Kalker, T. (2008). *Digital Watermarking and Steganography* (2008th ed.). Burlington: Morgan Kaufmann, Elsevier.
- Dutta, P., Bhattacharyya, D., & Kim, T. (2009). Data Hiding in Audio Signal: A Review. *International Journal of Database Theory and Application*, 2(2), 1–8.
- Foo, S. W. (2008). Non-blind audio-watermarking using compression-expansion of signals. In *IEEE Asia-Pacific Conference on Circuits and Systems, Proceedings, APCCAS* (pp. 1288–1291). <https://doi.org/10.1109/APCCAS.2008.4746263>
- Harahap, H., Budiman, G., & Novamizanti, L. (2016). Implementasi Teknik *Watermarking* menggunakan FFT dan Spread Spectrum *Watermark* pada Data Audio Digital. *ELKOMIKA*, 4(1), 98–109.
- ITU-R, R. RECOMMENDATION ITU-R BS.1387-1 Method for objective measurements of perceived audio quality, BS.1387-1 International Telecommunications Union-Recommendation, Geneva (1998). Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Method+for+objective+measurements+of+perceived+audio+quality+1#0%5Cnhttp://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:1387,+Method+for+objective+measurements+of+perceived+audio+qualit>
- Jane, O., Elbaşı, E., & İlk, H. G. (2014). Hybrid non-blind watermarking based on DWT and SVD. *Journal of Applied Research and Technology*, 12(4), 750–761. [https://doi.org/10.1016/S1665-6423\(14\)70091-4](https://doi.org/10.1016/S1665-6423(14)70091-4)
- Khalil, M., & Adib, A. (2015). From blind to non-blind audio watermarking: A new approach based on blind source separation. In *1st International Conference on Electrical and Information Technologies ICEIT'2015* (pp. 237–240).
- Kirovski, D., & Malvar, H. S. (2003). Spread-Spectrum Watermarking of Audio Signals. *IEEE Transactions on Signal Processing*, 51(4), 1020–1033.
- Malvar, H. S., & Florêncio, D. A. F. (2003). Improved spread spectrum: A new modulation technique for robust watermarking. *IEEE Transactions on Signal Processing*, 51(4), 898–905. <https://doi.org/10.1109/TSP.2003.809385>
- Panda, J., Gera, K. R., & Bhattacharyya, A. (2014). Non-Blind Audio watermarking Scheme

- Based on Empirical Mode Decomposition. *International Journal of Advanced Science, Engineering and Technology*, 3(3), 38–43.
- Robinson, D. J. M. (1999). the Human Auditory System. *107th Convention of the Audio Engineering Society*, 1–13.
- Shulze, H., & Luders, C. (2005). *Theory and Applications of OFDM and CDMA*. A John Wiley&Sons Ltd. <https://doi.org/10.1002/0470017406>
- Singh, P., & Chadha, R. (2013). A Survey of Digital *Watermarking* Techniques, Applications and Attacks. *International Journal of Engineering and Innovative ...*, 2(9), 165–175. <https://doi.org/10.1109/INDIN.2005.1560462>
- Xiang, Y., Natgunanathan, I., Peng, D., Hua, G., & Liu, B. (2018). Spread Spectrum *Audio watermarking* Using Multiple Orthogonal PN Sequences and Variable Embedding Strengths and Polarities. *IEEE/ACM Transactions on Audio Speech and Language Processing*, 26(3), 529–539. <https://doi.org/10.1109/TASLP.2017.2782487>
- Yani, K., Rizal, A., & Prasetya, B. (2008). Analisis Kinerja Algoritma Short Time Fourier Transform (STFT) Untuk Deteksi Sinyal Carrier Frequency Hopping Spread Spectrum (FHSS) CDMA. In *Seminar Sistem Informasi Indonesia (SESINDO2008)*.
- Zhang, P., Xu, S. Z., & Yang, H. Z. (2012). Robust *audio watermarking* based on extended improved spread spectrum with perceptual masking. *International Journal of Fuzzy Systems*, 14(2), 289–295.