# Pengujian Beberapa Teknik Proteksi Watermark Terhadap Penyerangan

Dibi Khairurrazi Budiarsyah, 13509013
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
13509013@std.stei.itb.ac.id

***Abstraksi* — Perkembangan teknologi internet telah memudahkan manusia dalam banyak hal. Salah satunya adalah berbagi gambar, video, musik, dll. Kemudahan distribusi ini dapat menimbulkan permasalahan ketika file yang memiliki hak cipta disebarluaskan oleh orang-orang yang tidak bertanggung jawab. Salah satu cara untuk melindungi file dari penyebarluasan yang tidak bertanggung jawab adalah dengan menggunakan watermark. terdapat beberapa teknik watermarking berdasarkan domain yakni spatial domain (contohnya adalah metode LSB), transform domain, dan hybrid. Beberapa contoh dari teknik transform domain watermarking adalah discrete cosine transform (DCT), discrete fourier transform (DFT), discrete wavelet transform (DWT) maupun discrete laguerre transform (DLT). Dengan ditemukannya teknik watermarking, tentunya akan ada seseorang yang berusaha untuk merusak hal tersebut. Terdapat beberapa jenis serangan yang dapat dilakukan terhadap watermark, diantaranya adalah scrambling attack, synchronization attack, linear filtering dan noise removal, copy attack, ambiguity/deadlock attack, sensitivity analysis attack, gradient descent attack, dll. Beberapa pendekatan yang ingin diimplementasi terkait masalah ambiguity attack adalah Selective Detection, Multiple Watermark Embedding, dan Zero Knowledge Watermark Detection.***

***Index Terms*—Watermark, pengujian, ambiguity attack, Multiple Watermark Embedding.**

## I. PENDAHULUAN

Seiring dengan berkembangnya teknologi di dunia ini, arus penyebaran informasi menjadi lebih cepat dan mudah. Salah satu teknologi yang mendukung penyebaran informasi ini adalah internet. Dengan menggunakan internet, setiap orang dapat saling berbagi informasi dan data. Namun tidak semua informasi yang beredar di internet diketahui asal dan pemiliknya, dengan demikian suatu informasi dapat disalahgunakan atau digunakan secara tidak bertanggung jawab. Salah satu langkah penganganannya adalah dengan menggunakan hak cipta atau *copyright*. Salah satu cara meletakkan hal cipta pada suatu informasi adalah dengan menggunakan *watermark*. *Watermark* sendiri merupakan salah satu bentuk aplikasi dari *steganografi*.

*Steganografi* merupakan seni dalam menyembunyikan pesan dengan suatu cara tertentu sehingga hanya pengirim dan penerima saja yang mengerti. Kata *steganografi* berasal dari bahasa yunani yang mengandung kata *steganos* yang artinya rahasia dan *grafi* yang berarti sebuah tulisan atau sebuah gambar. Penggunaan *steganografi* ini telah dilakukan dari zaman dahulu misalnya dengan menggunakan tinta tak terlihat.

Perbedaan *steganografi* dengan *kriptografi* adalah *kriptografi* merupakan seni menyamarkan pesan sehingga tidak dapat dibaca orang lain sedangkan *steganografi* membuat pesan tersebut seolah-olah tidak ada. Hal ini merupakan kelebihan dari *steganografi* dimana *steganografi* tidak akan menimbulkan kecurigaan dari orang lain. Terkadang *steganografi* dan *kriptografi* digunakan secara bersamaan untuk menjamin kerahasiaan informasi.

Informasi rahasia dapat disembunyikan dalam sebuah gambar digital, musik, file teks, ataupun video. Dengan menggunakan algoritma yang tepat, seorang ahlipun akan sulit membedakan video yang telah disisipi pesan dengan video yang asli. Dengan demikian, penyebaran informasi dapat dilakukan dengan lebih aman.

Terdapat beberapa teknik yang dapat digunakan untuk melakukan *steganografi* pada media digital. Diantaranya adalah *spatial domain* dan *transform domain*. *Spatial domain* memodifikasi langsung nilai *byte* dari *cover object*, salah satu metodenya adalah metode modifikasi LSB. *Transform domain* memodifikasi hasil transformasi dalam ranah sinyal frekuensi, salah satu metodenya adalah dengan menggunakan metode spread spectrum.

## II. WATERMARK DAN JENIS-JENIS PENYERANGAN TERHADAP WATERMARK

### A. *Watermark*

*Watermarking* adalah teknik untuk menyisipkan informasi tertentu ke dalam sebuah data dengan suatu cara tertentu sehingga *watermark* tersebut sulit untuk dirusak atau dihapus. Teknik *watermarking* sudah ditemukan sejak 700 tahun yang lalu dimana pada akhir abad 13, pabrik kertas di Fabriano, Italia, membuat kertas yang diberi *watermark* dengan cara menekan bentuk cetakan

gambar atau tulisan pada kertas yang baru setengah jadi. Ketika kertas dikeringkan terbentuklah suatu kertas yang ber-*watermark*. Kertas ini biasanya digunakan oleh seniman atau sastrawan untuk menulis karya mereka, *watermark* tersebut digunakan untuk mengidentifikasi bahwa karya seni di atasnya adalah milik mereka.

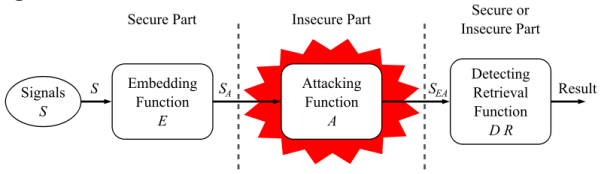
Terdapat dua macam *watermarking*, yaitu *visible watermarking* dan *invisible watermarking*. Pada *visible watermarking*, informasi yang ditambahkan akan terlihat pada gambar atau video. Biasanya informasi yang ditambahkan pada *visible watermarking* adalah text atau logo yang mengidentifikasi pemilik dari data. Pada *invisible watermarking*, *watermark* ditambahkan pada data, tetapi *watermark* tidak dapat dilihat dengan mata telanjang.

*Watermark digital* merupakan suatu tanda yang diletakkan pada *noise-tolerant signal* seperti data gambar dan audio. biasanya, *watermark* digunakan untuk mengidentifikasi siapa pemilik dari sinyal tersebut.

*Digital watermarking* bermulai dari Ide *watermarking* pada data digital yang dikembangkan di Jepang tahun 1990 dan di Swiss tahun 1993. *Digital watermarking* semakin berkembang seiring dengan semakin meluasnya penggunaan internet, objek digital seperti video, citra, dan suara yang dapat dengan mudah digandakan dan disebarluaskan. *Watermark* pada *Digital watermarking* dapat berupa teks, logo, data audio, hingga rangkaian bit yang tidak berarti.

Informasinya yang disembunyikan tidak harus ada hubungannya dengan sinyal file *carrier*-nya. dengan demikian, *watermark* digital dapat digunakan untuk mengautentifikasi file carrier atau menunjukkan identitas pemilik file tersebut. seperti halnya *watermark*, *watermark* digital hanya dapat dilihat setelah tercapainya suatu kondisi, misalkan dengan menggunakan suatu algoritma tertentu.

*Watermark* tradisional bisa diaplikasikan ke beberapa bentuk media seperti gambar dan video, dimana *watermark digital* dapat diletakkan pada audio, gambar, video, teks, atau model tiga dimensi. *watermark digital* tidak membuat ukuran dari file carrier berubah. *Watermark digital* dapat digunakan untuk keperluan perlindungan hak cipta, penelusuran sumber, dan pengawasan siaran.




**Gambar 1. Fase life-cycle dari watermark, dengan fungsi embedding, attacking dan detection/retrieval**
**Sumber:http://en.wikipedia.org/wiki/Digital_watermarking**

Sebuah sistem watermarking biasanya dibagi menjadi tiga tahap, yaitu embedding, attack dan detection.

Pada embedding, sebuah algoritma menerima file carrier dan data watermark sehingga menghasilkan data yang telah diberi watermark. Data kemudian ditransmisikan atau disimpan.

Jika seseorang melakukan modifikasi terhadap file carrier tersebut, maka proses tersebut disebut attack karena ada modifikasi yang dilakukan yang bertujuan untuk merusak atau menghilangkan watermark yang terdapat pada data.

Detection, sering disebut juga extraction, adalah sebuah algoritma yang diaplikasikan kepada file carrier yang mungkin telah di attack oleh orang lain. Detection bertujuan untuk mendapatkan watermark dari data. Jika data tidak dimodifikasi ketika transmisi, maka watermark akan tetap ada dan dapat di ekstrak. Algoritma ekstraksi harus dapat menghasilkan watermark yang tepat, bahkan ketika modifikasi yang dilakukan cukup kuat. Jika watermarking yang dilakukan bertipe fragile, algoritma ekstraksi akan gagal jika terdapat perubahan pada data.

Digital watermarking dapat di klasifikasi dalam beberapa cara antara lain:

- Robustness
  Klasifikasi berdasarkan kekokohan dari watermark. Sebuah watermark dikatakan fragile jika gagal untuk di deteksi setelah data dimodifikasi sedikit, semi-fragile jika cukup kuat dan robust jika sangat kuat.

- Perceptibilty
  Klasifikasi berdasarkan penyembunyian watermark. Sebuah watermark dikatakan imperceptible jika data watermark tidak dapat dirasakan perbedaannya dengan data aslinya. Sebuah watermark dikatakan perceptible jika kehadiran watermark dapat dirasakan.

- Capacity
  Klasifikasi berdasarkan kapasitas ukuran informasi yang dapat disembunyikan kedalam data digital.

Selain itu, watermark dapat di klasifikasikan berdasarkan metode embed dan retrieve antara lain:

- Spread-spectrum
  *Watermark* di letakkan dengan menggunakan modifikasi *additive*. *Watermark Spread-spectrum* dikenal sebagai *watermark* yang cukup kokoh, tetapi hanya dapat menampung sedikit informasi karena inferensi dari host.

- Quantization

*Watermark* di letakkan dengan *quantization*. *Watermark Quantization* tidak kokoh, tetapi mempunyai kapasitas informasi yang besar karena inferensi *host* yang di *reject*

- Amplitude Modulation
  *Watermark* di letakkan dengan modifikasi *additive* tetapi dilakukan di *spatial domain*

## B. Teknik watermarking

Terdapat beberapa teknik watermarking berdasarkan *domain* yakni

- *spatial domain* : Metode LSB
- *transform domain*
  Terdapat beberapa metode transform domain, diantaranya adalah :
  - *discrete cosine transform* (DCT)
  - *discrete fourier transform* (DFT)
  - *discrete wavelet transform* (DWT)
  - *discrete laguerre transform* (DLT)
- *hybrid*.

## C. Penyerangan terhadap watermark

Terdapat empat kategori penyerangan terhadap watermark. Yakni :

- Removal Attack
  Removal attack merupakan percobaan untuk memisahkan dan menghilangkan watermark dari carriernya. Tujuannya adalah untuk memberikan distorsi pada gambar sehingga watermark yang ada menjadi tidak dapat ditemukan dan dibaca. Penyerangan dapat dikatakan sukses apabila watermark tidak dapat dibaca namun gambar tidak rusak dan dapat dikenali. beberapa operasi yang dapat dilakukan antara lain :
  - Lossy image compression (JPEG, JPEG 2000)
    Gambar yang beredar di internet biasanya sudah dikompresi. Agar watermark dapat bertahan dari proses kompresi ini. maka proses penempatan watermark harus berada pada domain yang sama dengan proses kompresi tersebut. Misalnya DCT lebih cocok digunakan pada file JPEG dibandingkan dengan spatial-domain watermarking. Dan DWT lebih cocok digunakan untuk JPEG2000.
  - Addition of Gaussian noise
    Dengan menggunakan *perceptially shaped noise*, penyerang dapat mengotak-atik batas dimana detektor watermark bekerja.
  - Denoising
    Memodelkan noise tambahan yang ditimbulkan dari adanya watermark relatif terhadap gambar aslinya. Beberapa metodenya antara lain local median, midpoint, trimmed mean filtering, dll.
  - Filtering
    Beberapa contoh penyerangannya antara lain highpass, lowpass, gaussian and sharpening filtering.
  - Statistical Averaging
    Dengan menganalisis beberapa dataset. Penyerang berusaha untuk mendapatkan file carrier dan file watermark, kemudian dilakukan unwatermark untuk menghilangkan watermark dari file carriernya.

- Geometrical Attacks
  *Geometrical attack* tidak bertujuan untuk menghilangkan *watermark*, tetapi bertujuan untuk menghancurkan *watermark* atau menonaktifkan deteksi. Caranya adalah dengan mencoba untuk merusak korelasi deteksi antara *watermark* yang terdeteksi dengan *watermark* aslinya dimana gambar yang diberi *watermark* di translasi, rotasi, *rescale*, atau di *crop*. Hal ini dapat dipenuhi dengan mengacak pixelnya. Penyerangan ini dapat dibagi menjadi *general affine transformation* dan *projective transformation*. Metode yang paling umum adalah *cropping*.

- Crypthographic Attacks
  Tujuan dari penyerangan ini adalah untuk merusak sistem keamanan dari *watermark*, dengan demikian *watermark* bisa dihilangkan atau ditumpuk dengan *watermark* yang lain. Salah satu metodenya adalah *brute force* untuk mencari informasi rahasia yang disembunyikan.

- Protocol Attacks
  Meletakkan *watermark* sendiri untuk membuat ambiguitas kepemilikian informasi. Konsep ini membuat *watermark* yang diletakkan pada sesuatu haruslah *non-invertible*. Contoh penerapannya adalah *deadlock attack* dan *copy attack*.

Selain dapat dikategorikan menjadi empat kategori yang telah disebutkan sebelumnya, penyerangan terhadap watermark dapat dikategorikan menjadi empat kategori yang berbeda pula, yakni :

- Unauthorized Embedding
  Penyerangan dilakukan terhadap fragile watermark. Meletakkan watermark ke file yang telah diberikan watermark untuk menimbulkan ambiguitas.

  Cara untuk menangkal serangan dari kategori ini adalah dengan menggunakan digital signature atau menggunakan kriptografi untuk mencegah peletakkan watermark baru oleh orang lain.

- Unauthorized Detection
  Melakukan deteksi dan ekstraksi *watermark*. Cara

untuk menangkal serangan dari kategori ini adalah dengan cara mengenkripsi watermark sebelum diletakkan pada data.

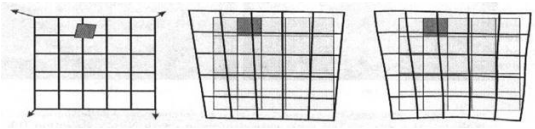
- Unauthorized Removal
  Melakukan penghilangan atau penyembunyian terhadap *watermark* sehingga *watermark* tidak dapat dideteksi.

- System Attack
  Melakukan penyerangan terhadap sistem *embed* dan *retrieve* dari detektor.

Terdapat beberapa teknik penyerangan yang dapat dilakukan terhadap watermark. Diantaranya adalah :

- Scrambling attack
  File yang ada diacak atau dipecah menjadi file yang kecil sehingga detektor tidak dapat menemukan adanya *watermark*. Setelah proses deteksi, file tersebut di kembalikan ke bentuk awal. Langkah penanganannya adalah dengan

- Synchronize attack
  Serangan yang dilakukan mentransformasi bentuk pada file sehingga *watermark* yang terkandung didalamnya tidak dapat dilacak oleh detektor. Contoh serangan ini adalah *stirmark attack* yang diajukan oleh PetitColas. Salah satu contoh *stirmark attack* dapat dilihat pada gambar dibawah ini. Gambar paling kiri merupakan gambar asli dimana gambar kedua dan ketiga merupakan hasil *bending* dan *randomization*.



**Gambar 2. StirMark Attack**
**Sumber:www.ece.wisc.edu/~hu/ece738/notes/watermarkattack.ppt**

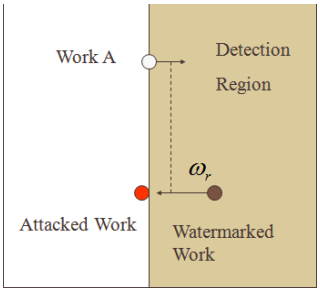

**Gambar 3. Hasil StirMark Attack**
**Sumber:www.ece.wisc.edu/~hu/ece738/notes/watermarkattack.ppt**

- Linear Filtering dan Noise Removal
  Penyerangan dilakukan dengan cara menghilangkan *noise* tambahan yang ditimbulkan dari adanya *watermark* pada file. Contoh aplikasinya adalah *Host Data Estimation*.

- Ambiguity attack
  Penyerangan dilakukan dengan meletakkan *watermark* lain sehingga terdapat dua *watermark* pada satu file. Hal ini dapat menimbulkan adanya ambiguitas kepemilikian terhadap file. Untuk menganganinya, maka *watermark* yang digunakan haruslah *non-invertible*.

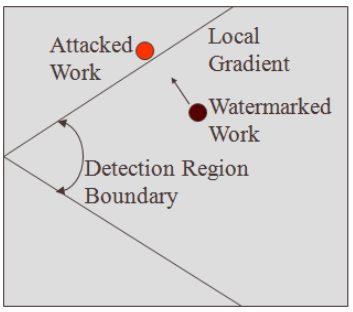
- Copy Attack
  Penyerangan dilakukan dengan cara mengetahui pola dari *watermark* yang diterapkan, kemudian menerapkan pola *watermark* tersebut ke file *carrier* yang lain. Salah satu contoh aplikasinya adalah *collage attack*.

- Sensitivity analysis attack
  Penyerangan dilakukan dengan cara mencari bagian dari file yang tidak ada watermarknya. Penyerangan ini ditemukan oleh Kalker.




**Gambar 4. Tahap-tahap Sensitivity Analysis Attack**
**Sumber:www.ece.wisc.edu/~hu/ece738/notes/watermarkattack.ppt**

- Gradient descendant attack
  Penyerangan jenis ini juga ditemukan oleh Kalker. Penyerangan dilakukan dengan cara memisahkan gradien. Dengan mengetahui gradien tersebut, bagian yang di deteksi oleh detektor akan diketahui sehingga dapat diserang.




**Gambar 5. Tahap-tahap Gradient Descent Attack**
**Sumber:www.ece.wisc.edu/~hu/ece738/notes/watermarkattack.ppt**

## III. BEBERAPA ALGORITMA PENCEGAH AMBIGUITY ATTACK YANG DIGUNAKAN PADA PERCOBAAN INI

### A. Multiple Watermark Embedding oleh Sencar, Qiming, dan Nasir

Ide utama dari algoritma ini adalah dengan menerapkan beberapa watermark dalam satu file. Cara pendeteksiannya adalah dengan mendeteksi secara acak subset dari watermark yang telah di embed tersebut.

$$M = E(C,W) = C + \alpha W \quad (1)$$

Dimana M, C Є $R^n$, W Є $\{-1, 1\}^n$, α Є R. C merupakan carrier, W merupakan watermark, dan M merupakan file yang akan di publish. E merupakan fungsi embed yang digunakan.

Fungsi deteksi menghasilkan nilai boolean

$$D(M,W) = \begin{cases} true, if\ \tau < \sum_1^n M[i]\ x\ W[i] \\ false, otherwise \end{cases} \quad (2)$$

Dimana $\tau$ merupakan nilai batas.

### B. Zero Knowledge Watermark Detection oleh Qiming dan Chang

Qiming dan Chang membuat sebuah detektor yang memberikan *constrain* pada *watermark* yang dapat mencegah kecurangan berupa *ambiguity attack* pada *watermark*.

## IV. PENGUJIAN

Karena implementasi dari dua algoritma tersebut belum dapat diselesaikan dengan baik, maka pengujian tidak dapat dilakukan

## V. KESIMPULAN

Makalah ini menceritakan tentang jenis-jenis penyerangan pada watermark kemudian fokus kepada beberapa pendekatan yang diajukan oleh beberapa peneliti untuk menangkal serangan ambiguitas pada watermark. Namun pengujian belum dapat diselesaikan dengan baik dikarenakan implementasi dari dua algoritma yang ingin digunakan belum dapat diselesaikan.

## REFERENSI


Li, Q., & Chang, E. C. (2006, September). Zero-knowledge watermark detection resistant to ambiguity attacks. In *Proceedings of the 8th workshop on Multimedia and security* (pp. 158-163). ACM..

Fabien A. P. Petitcolas, Ross J. Anderson, Markus G. Kuhn. Attacks on copyright marking systems, in David Aucsmith (Ed), Information Hiding, Second International Workshop, IH'98, Portland, Oregon, U.S.A., April 15-17, 1998, Proceedings, LNCS 1525, Springer-Verlag, ISBN 3-540-65386-4, pp. 219-239.H. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1985, ch. 4.

Samcovic, A., & Turan, J. (2008). Attacks on Digital Wavelet Image Watermarks. *JOURNAL OF ELECTRICAL ENGINEERING-BRATISLAVA-,59*(3), 131.

Fabien A. P. Petitcolas. Watermarking schemes evaluation. I.E.E.E. Signal Processing, vol. 17, no. 5, pp. 58–64, September 2000.E. H. Miller, "A note on reflector arrays (Periodical style—Accepted for publication)," *IEEE Trans. Antennas Propagat.*, to be published.

Bhattacharya, S., & Cortesi, A. (2010, October). Zero-knowledge Software Watermarking for C Programs. In *Proceedings of the 2010 International Conference on Advances in Communication, Network, and Computing* (pp. 282-286). IEEE Computer Society.

Sencar, H. T., Li, Q., & Memon, N. (2007, September). A new approach to countering ambiguity attacks. In *Proceedings of the 9th workshop on Multimedia & security* (pp. 205-214). ACM.

Taha Senear, H., & Memon, N. (2007). Combatting ambiguity attacks via selective detection of embedded watermarks. *Information Forensics and Security, IEEE Transactions on, 2*(4), 664-682.

http://www.ece.wisc.edu/~hu/ece738/notes/watermarkattack.ppt. waktu akses : 20 Maret 2013.

[1] informatika.stei.itb.ac.id/~rinaldi.munir/Kriptografi/2009-2010/Makalah1/Makalah1_IF3058_2010_038.pdf

[2] http://elreg-03.blogspot.com/2010/01/implementasi-teknik-watermarking.html. waktu akses : 3 maret 2013

[3] http://digilib.ittelkom.ac.id/index.php?option=com_content&view=article&id=865:watermark&catid=21:itp-informatika-teori-dan-pemograman&Itemid=14. Waktu akses : 3 maret 2013.

http://www.petitcolas.net/fabien/watermarking/stirmark/. waktu akses 23 maret 2013


## PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 26 Maret 2013

ttd

Dibi Khairurrazi Budiarsyah
13509013