Analisis dan Implementasi Penggunaan Biometrik Suara sebagai Pembangkit Kunci Kriptografi AES 128 bit

Dimas Priambodo[1], Maman Abdurohman[2], Iwan Iwut Tritoasmoro[3]

[1,2]Departemen Teknik Informatika Institut Teknologi Telkom, Bandung
[3]Departemen Teknik Elektro Institut Teknologi Telkom, Bandung
[1]dimaspriambodo@students.ittelkom.ac.id, [2]mma@ittelkom.ac.id, [3]iww@ittelkom.ac.id

Abstrak

Teknologi pengenalan suara merupakan salah satu teknologi biometrik yang tidak memerlukan biaya besar serta peralatan khusus. Suara merupakan salah satu dari bagian tubuh manusia yang unik dan dapat dibedakan dengan mudah. Sistem yang dibangun pada penelitian ini adalah sistem verifikasi suara yang dapat memverifikasi/membuktikan identitas yang diklaim berdasarkan suara yang diberikan. Kemudian, dari hasil ekstraksi ciri suara tersebut akan dibangkitkan kunci kriptografi yang akan digunakan untuk proses enkripsi/ dekripsi dengan menggunakan algoritma AES 128 bit.

Sistem ini dibangun menggunakan metode MFCC (*Mel Frequency Cepstrum Coefficients*) sebagai proses ekstraksi ciri dan metode VQ (*Vector Quantization*) sebagai proses pencocokan ciri. Proses MFCC akan mengkonversi sinyal suara menjadi beberapa vektor yang berguna untuk proses pengenalan. Vektor ciri hasil proses MFCC selanjutnya akan dibandingkan dengan vektor ciri yang tersimpan dalam basis data melalui proses VQ berdasarkan identitas yang diklaim oleh *user*. Baik buruknya akurasi sistem dalam mengenali karakteristik suara dipengaruhi oleh jumlah parameter *filterbank* pada MFCC dan jumlah parameter *centroid* pada VQ.

Kata Kunci: verifikasi pembicara, pembangkit kunci kriptografi, MFCC, VQ.

Abstract

Voice recognition technology is one of the biometric technology that does not requires a huge cost and special equipment. Voice is one of the unique parts of the human body and can be distinguished easily. The system built on this research is the speaker verification system to verify / prove the claimed identity based on the speech given. Then, based on the result of voice feature extraction, cryptographic keys will be generated to be used for encryption / decryption process using AES 128-bit algorithm.

The system is built using the MFCC (Mel Frequency cepstrum Coefficients) as the feature extraction process and VQ (Vector Quantization) as the process of feature matching. MFCC process will convert the voice signal into a useful vector for the recognition. Results of the MFCC feature vector will then be compared with the characteristic vectors stored in a database through the VQ based identity claimed by a user. Good or bad accuracy in recognizing the characteristic in the speaker verification system is affected by the number of parameters filterbank in the MFCC and the number of parameters centroid in the VQ.

Keywords: speaker verification, cryptographic key generator, MFCC, VQ.

1. Pendahuluan

Keamanan sebuah data adalah sebuah hal yang sangat penting. Oleh sebab itu metode untuk mengamankan sebuah file juga terus berkembang. Melakukan enkripsi terhadap sebuah file merupakan sebuah upaya untuk menjaga kerahasiaan data supaya data tersebut tidak dapat terbaca oleh pihak lain yang tidak berwenang. Namun metode enkripsi yang banyak dipakai saat ini hanya menggunakan sebuah kunci yang berupa PIN atau *password*, sehingga ketika nomor PIN atau *password* tersebut jatuh ke tangan orang lain, maka PIN atau *password* tersebut dapat dipergunakan untuk membuka sebuah file yang telah terenkripsi yang seharusnya tidak boleh dibaca oleh orang lain.

Enkripsi dengan menggunakan kunci yang dibangkitkan dari biometrik tubuh dapat menghindari kelemahan yang terdapat pada metode enkripsi konvensional. Salah satu biometrik tubuh yang dapat digunakan sebagai pembangkit kunci enkripsi adalah suara, karena suara manusia memiliki karakteristik yang unik pada setiap individu sehingga dimungkinkan untuk dijadikan sebagai pembangkit kunci enkripsi.

Pada pengerjaan tugas akhir ini akan dibuat metode enkripsi yang menggunakan biometrik suara manusia sebagai pembangkit kuncinya. Metode MFCC digunakan sebagai *feature extraction* untuk mengekstrak karakteristik suara, dan metode *Vector Quantization* digunakan sebagai *feature matching* untuk mencocokkan karakteristik suara. Metode enkripsinya sendiri menggunakan algoritma AES 128 bit.

2. Biometrik

Biometrik adalah karakteristik yang bersifat unik yang melekat pada setiap individu. Biometrik memiliki banyak kegunaan di bidang keamanan,

karena dapat digunakan untuk mengotentikasi identitas seseorang karena sifatnya yang unik untuk setiap individu. Biometrik yang digunakan sebagai alat pengenal identitas seseorang memiliki banyak kelebihan dibandingkan dengan metode-metode lain yang saat ini banyak digunakan, seperti *password* atau kartu akses, karena biometrik tidak dapat hilang, dipalsukan, terlupa, atau dicuri dengan mudah [10].

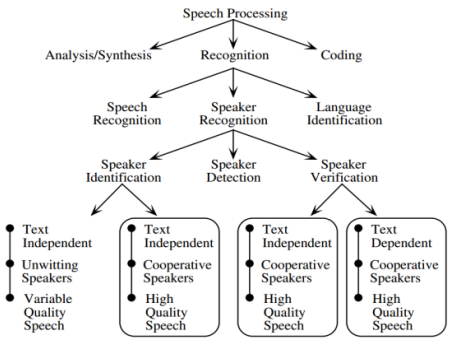
Dalam penelitian ini digunakan biometrik suara sebagai alat pengenal, karena biometrik suara memiliki beberapa keunggulan dibanding dengan penggunaan biometrik yang lain, diantaranya adalah:

- Mudah digunakan, karena bahasa (suara) adalah bentuk alami dalam berkomunikasi.
- Tidak memerlukan biaya yang besar dalam mengimplementasikannya, karena tidak memerlukan alat yang khusus.
- Dapat digunakan jarak jauh (*remote access*).

2.1. Pengenalan Suara

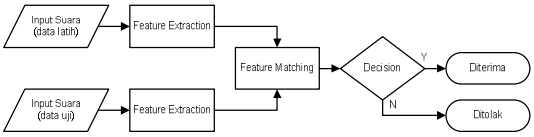
Pengenalan suara secara garis besar dapat dikategorikan menjadi 3 bagian, yaitu: *speech recognition*, *speaker recognition*, dan *language recognition* [1]. Dalam penelitian ini hanya khusus membahas mengenai *speaker recognition* dan lebih spesifik lagi mengenai *speaker verification*.

Berbeda dengan proses *speech recognition* yang berusaha mengenali kata/kalimat yang diucapkan oleh seseorang, *speaker recognition* adalah suatu proses yang bertujuan untuk mengenenali siapa yang sedang berbicara berdasarkan informasi yang terkandung dalam gelombang suara yang diberikan. Dalam setiap sistem *speaker recognition*, selalu terdapat 2 tahapan. Tahapan yang pertama adalah tahap pengenalan sistem. Dalam tahapan ini sistem dibekali kemampuan untuk mengenali setiap karakteristik dari suara yang diberikan. Kemudian tahapan yang selanjutnya adalah tahap pengujian sistem. Pada tahap ini sistem diminta untuk mengidentifikasi atau memverifikasi pembicara berdasarkan kemiripan antara karakteristik dari suara yang diberikan dan karakteristik dari suara yang telah dikenali oleh sistem sebelumnya [5]. *Speaker recognition* dapat dibagi menjadi 2 bagian, yaitu: *speaker identification* dan *speaker verification*.



Gambar 2.1. Kategori pengolahan sinyal suara

Prinsip utama dari setiap sistem *speaker verification* adalah *feature extraction* dan *feature matching* [5]. Fungsi dari *feature extraction* adalah untuk mengubah sinyal-sinyal suara menjadi beberapa parameter. Sedangkan *feature matching* berfungsi untuk membandingkan kemiripan parameter antara hasil ekstraksi data latih dan data uji. Diagram blok dibawah ini menggambarkan secara umum proses yang terjadi dalam *speaker verification.*



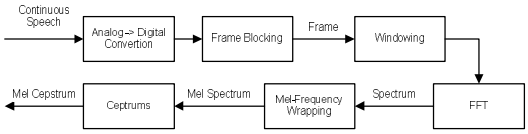
Gambar 2.2. Diagram proses speaker verification

2.1.1. Mel-Frequency Cepstral Coefficients

MFCC (*Mel-Frequency Cepstral Coefficients*) merupakan salah satu metode berbasis pada transformasi *fourier* yang banyak digunakan dalam bidang *speech technology*, baik dalam *speech recognition* maupun *speaker recognition*. Metode ini digunakan dalam proses *feature extraction*, yaitu sebuah proses yang mengkonversi sinyal suara menjadi beberapa parameter untuk keperluan analisis dan pemrosesan selanjutnya. Beberapa keunggulan dari metode ini adalah [11]:

a. Mampu menangkap karakteristik suara yang sangat penting dalam proses pengenalan suara, atau dengan kata lain dapat menangkap informasi-informasi penting yang terkandung dalam sinyal suara.
b. Menghasilkan data seminimal mungkin tanpa menghilangkan informasi-informasi penting yang terkandung didalamnya.
c. Mereplikasi organ pendengaran manusia dalam melakukan persepsi terhadap sinyal suara.

MFCC sebenarnya merupakan adaptasi dari sistem pendengaran manusia, dimana sinyal suara akan difilter secara linear untuk frekuensi rendah (dibawah 1000Hz) dan secara logaritmik untuk frekuensi tinggi (diatas 1000Hz) [7, 8, 14, 16]. Skala ini disebut juga sebagai skala mel. Diagram blok dibawah menunjukkan proses yang terjadi di dalam MFCC.



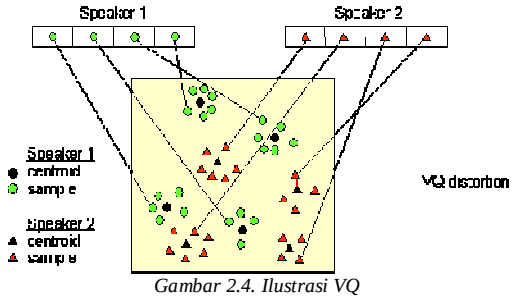
Gambar 2.3. Diagram Proses MFCC

2.1.2. Vector Quantization

Keputusan diterima atau tidaknya akses seseorang pada sistem *speaker verification* ditentukan oleh fungsi *feature matching*. *Feature matching* pada sistem *speaker verification* berfungsi untuk membandingkan karakteristik suara seseorang dengan karakteristik suara tertentu yang telah

dikenali oleh sistem. Ada banyak metode yang dapat digunakan pada fungsi *feature matching*, diantaranya adalah DTW (*Dynamic Time Warping*), HMM (*Hidden Markov Model*), dan VQ (*Vector Quantization*). Pada penelitian ini digunakan metode VQ sebagai fungsi *feature matching*, karena kemudahan dalam pengimplementasiannya dan tingkat keakuratannya yang tinggi [5].

Vector Quantization adalah sebuah proses untuk memetakan ruang vektor yang besar ke dalam sejumlah wilayah yang terbatas dalam ruang tersebut. Setiap wilayah disebut dengan *cluster*, dan dapat direpresentasikan oleh titik pusatnya yang disebut *codeword* atau *centroid*. Kumpulan dari semua *codeword* yang terdapat dalam ruang vektor tersebut disebut dengan *codebook*. *Vector quantization* adalah metode kompresi data yang bersifat *lossy* berdasarkan pada prinsip *block coding*. VQ dapat juga disebut sebagai *approximator* [12, 16].



Gambar 2.4. Ilustrasi VQ

2.1. Kriptografi

Secara umum, teknik kriptografi digunakan untuk menyandikan sebuah informasi atau pesan agar tidak dapat terbaca oleh orang lain yang tidak memiliki wewenang. Proses dalam menyandikan sebuah informasi (*plaintext*) menjadi sebuah pesan yang tersandi (*ciphertext*) disebut dengan proses enkripsi, sedangkan proses untuk merubah kembali sebuah *ciphertext* menjadi *plaintext* disebut proses dekripsi.

Pada teknik kriptografi modern proses enkripsi dan dekripsi menggunakan sebuah parameter atau kunci (*key*) sehingga algoritma kriptografi itu sendiri tidak perlu dirahasiakan, tetapi hanya kunci yang perlu dijaga kerahasiaannya.

Penggunaan teknik kriptografi bertujuan untuk menjaga kerahasiaan pesan (*confidentiality*), integritas data (*data integrity*), otentikasi (*authentication*), serta nirpenyangkalan (*non-repudiation*) [9]. Salah satu teknik kriptografi yang paling sering digunakan adalah AES.

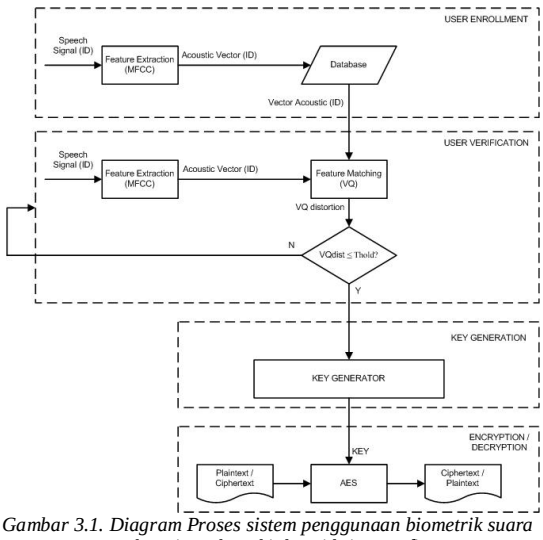
3. Perancangan Sistem

Sistem penggunaan biometrik suara sebagai pembangkit kunci kriptografi AES 128 bit merupakan sistem yang dapat memverifikasi identitas seseorang berdasarkan input suara yang diberikan, kemudian membangkitkan kunci untuk proses enkripsi atau dekripsi berdasarkan karakteristik biometrik suaranya. Sistem ini dibangun tanpa menggunakan *background noise filter*, sehingga seluruh pengujian hanya dapat dilakukan pada ruang yang kedap suara dengan tingkat kebisingan sekitar 30 dB.

Sistem ini dirancang menggunakan metode MFCC untuk mengekstrak karakteristik dari suara seseorang. MFCC merupakan sebuah metode yang diadaptasi dari sistem pendengaran manusia, dimana sinyal suara akan difilter secara linear untuk frekuensi rendah di bawah 1000 Hz, dan secara logaritmik untuk frekuensi tinggi diatas 1000 Hz. Sedangkan metode VQ digunakan untuk membandingkan kedua karakteristik suara pada proses *user verification*. Hasil dari pembandingan kedua suara menggunakan metode VQ akan menghasilkan sebuah nilai distorsi yang menunjukkan kemiripan karakteristik dari kedua suara yang dibandingkan. Semakin kecil nilai distorsi yang dihasilkan menunjukkan bahwa karakteristik kedua suara yang dibandingkan semakin mirip.

Pembangkitan kunci kriptografi akan secara otomatis dilakukan oleh sistem ketika *user* berhasil melakukan proses verifikasi. Sinyal suara yang diberikan *user* ketika melakukan proses verifikasi akan digunakan kembali untuk proses pembangkitan kunci. Hasil dari proses pembangkitan kunci akan menghasilkan kunci sepanjang 128 bit yang dapat digunakan untuk melakukan proses enkripsi maupun dekripsi dengan menggunakan algoritma AES 128 bit.

Proses yang terjadi dalam sistem penggunaan biometrik suara sebagai pembangkit kunci kriptografi AES 128 bit secara garis besar dapat dilihat pada diagram blok berikut:



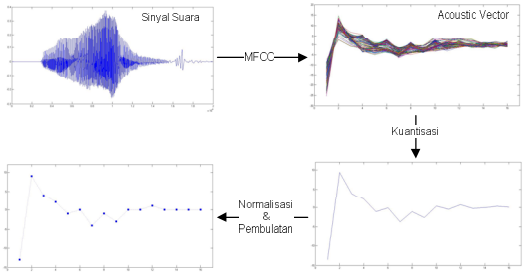
Gambar 3.1. Diagram Proses sistem penggunaan biometrik suara sebagai pembangkit kunci kriptografi

3.1. Pembangkitan Kunci

Proses pembangkitan kunci akan secara otomatis dilakukan oleh sistem ketika *user* berhasil

melakukan proses verifikasi. Input suara yang diberikan pada saat proses verifikasi akan diambil kembali untuk diekstrak karakteristik suaranya dan kemudian akan diubah menjadi parameter-parameter yang akan digunakan pada proses pembangkitan kunci.

Pada proses pembangkitan kunci, sinyal suara akan diekstrak kembali karakteristiknya menggunakan metode MFCC. Hasil dari proses ekstraksi ciri menggunakan MFCC ini akan menghasilkan sekelompok *acoustic vector*. Kemudian dari sekelompok *acoustic vector* ini akan dikuantisasi sehingga hanya akan menjadi 1 buah *acoustic vector*. Dari 1 buah *acoustic vector* ini kemudian akan dibagi ke dalam 16 bagian, dan pada masing-masing bagian akan dilakukan proses normalisasi dan pembulatan nilai sehingga akan menghasilkan nilai sepanjang 16 byte atau setara dengan 128 bit.



Gambar 3.2. Ilustrasi pembangkitan kunci

Daftar Pustaka:

[1] Campbell, J. P. (2001). *Speaker Recognition.* Fort Meade: Department of Defense.

[2] Chin, S., Lau, K., & Leu, L. (2002). *A Speaker Verification System.* UK: University of Victoria.

[3] Cornaz, C., & Hunkeler, U. (2003). *An Automatic Speaker Recognition System.* Ecole Polytechnique Federale De Lausanne.

[4] Hasan, R., Jamil, M., Rabbani, G., & Rahman, S. (2004). *Speaker Identification Using Mel Frequency Cepstral Coefficients.* Bangladesh: International Conference on Electrical & Computer Engineering.

[5] Kumar, C. S. (2011). *Design of An Automatic Speaker Recognition System Using MFCC, Vector Quantization and LBG Algorithm.* Guntur: Nalanda Institute of Technology.

[6] Menezes, A. J., Oorschot, P. C., & Vanstone, S. A. (1996). *Handbook of Applied Cryptography*. CRC Press.

[7] Molgaard, L. L., & Jorgensen, K. W. (2005). *Speaker Recognition: Special Course.* IMM DTU.

[8] Muda, L., Begam, M., & Elamvazuthi, I. (2010). *Voice Recognition Algorithms Using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques.* Malaysia: Journal of Computing.

[9] Munir, R. (2006). *Kriptografi*. Bandung: Informatika.

[10] Myers, L. (2004). *An Exploration of Voice Biometrics.* SANS Institute.

[11] Putra, D., & Resmawan, A. (2011). *Verifikasi Biometrik Suara Menggunakan Metode MFCC dan DTW.* Bali: Universitas Udayana.

[12] Srinivasan, A. (2012). *Speaker Identification and Verification Using Vector Quantization and Mel Frequency Cepstral Coefficients.* India: Research Journal of Applied Science, Engineering and Technology.

[13] Stallings, W. (2011). *Cryptography and Network Security.* US: Pearson.

[14] Tiwari, V. (2010). *MFCC and its Applications in Speaker Recognition.* India: International Journal on Emerging Technologies.

[15] Zilvan, V., & Muttaqien, F. H. (2011). *Identifikasi Pembicara Menggunakan Algoritme VF15 dengan MFCC sebagai Pengekstraksi Ciri*. INKOM.

[16] ___. (2010). *Speaker Recognition - Project Report.*

[17] ___. (n.d.). Retrieved September 17, 2012, from ittelkom.ac.id: http://digilib.ittelkom.ac.id/index.php?option=com_content&view=article&id=841:mel-frequency-cepstral-coefficient&catid=15:pemrosesan-sinyal&Itemid=14

[18] ___. (n.d.). Retrieved July 4, 2012, from edipermadi.wordpress.com: http://edipermadi.wordpress.com/tag/inverse-sbox/

[19] ___. (n.d.). Retrieved August 23, 2012, from emeraldinsight: http://www.emeraldinsight.com/journals.htm?articleid=1747892&show=html

[20] ___. (n.d.). Retrieved September 19, 2012, from faculty.nps.edu: http://faculty.nps.edu/dl/mobile/eo3404/B-Discrete-Fourier-Transform/b2_spectralEstimation_2.html

[21] ___. (n.d.). Retrieved August 27, 2012, from neural.cs.nthu.edu.tw: http://neural.cs.nthu.edu.tw/jang/books/audiosignalprocessing/speechfeaturemfcc.asp

[22] ___. (n.d.). Retrieved July 13, 2012, from www.rpi.edu: http://www.rpi.edu/dept/phys/ScIT/InformationTransfer/sigtransfer/signalcharacteristics.html

[23] ___. (n.d.). Retrieved August 23, 2012, from www.sciencedirect.com: http://www.sciencedirect.com/science/article/pii/S0010482509001267

[24] ___. (n.d.). Retrieved July 13, 2012, from Wikipedia: http://wikipedia.org/wiki/advanced_encryption_standard

[25] ___. (n.d.). Retrieved July 13, 2012, from Wikipedia: http://wikipedia.org/wiki/nato_phonetic_alphabet