

## STUDI KEMUNGKINAN DATA OVER VOICE MELALUI SISTEM PESAWAT TELEPON

**Hendra Setiawan**

Program Studi Teknik Elektro, Fakultas Teknologi Industri  
Universitas Islam Indonesia, Yogyakarta  
Jl. Kaliurang km 14,5 Yogyakarta  
[hendra.setiawan@uii.ac.id](mailto:hendra.setiawan@uii.ac.id)

### ABSTRAK

Makalah ini berisi uraian analisis kemungkinan pengiriman data di atas suara pada sistem pesawat telepon. Kajian dimulai dengan pengamatan sinyal suara percakapan di pesawat telepon pada kawasan frekuensi. Dari hasil analisis diperoleh adanya beberapa bagian di pita saluran suara yang sangat jarang terpakai oleh sinyal suara percakapan. Untuk mengetahui kemungkinan pemanfaatan bagian tersebut, sebuah sinyal sinusoidal dengan frekuensi tertentu dicampurkan/ditumpangkan pada sinyal suara. Sinyal hasil percampuran tersebut dimasukkan ke jalur suara sistem telepon dan dikirim ke bagian penerima sesuai dengan proses pengiriman suara secara normal. Di bagian penerima, sinyal suara dan sinyal sinusoidal dipisahkan dengan menggunakan rangkaian FIR filter. Untuk mengetahui tingkat kemiripan sinyal asli (sebelum dicampur) dan yang telah dipisahkan, digunakan perhitungan koefisien korelasi. Dari hasil simulasi dan perhitungan koefisien korelasi diketahui bahwa sinyal yang ditumpangkan mempunyai nilai koefisien korelasi rata-rata di atas 0,9 untuk frekuensi di bawah 250 Hz. Sedangkan nilai koefisien suara yang ditumpangkan bernilai minimal 0,9 ketika sinyal yang ditumpangkan mempunyai frekuensi di bawah 150Hz. Hasil ini menunjukkan bahwa pengiriman data lewat jalur suara pada sistem telepon selular sangat dimungkinkan dengan beberapa keterbatasan yang ada.

**Kata kunci:** Data over voice, koefisien korelasi, telepon, suara percakapan.

### ABSTRACT

*This paper presents a feasibility study of data transmission over voice in a telephone system. The study is started from a frequency domain observation of a speech signal in the telephony system. The observation results show some regions in the speech bandwidth that almost not occupied at all. In order to find a possibility to utilize those regions, a sine wave with a certain frequency is injected/superimposed to the speech. The mixed signal sent to the receiver side follows an exactly same process as the normal process of voice. In the receiver side, the mixed signal is separated into the sine wave and the speech using a FIR filter. To recognize the similarity of the original signal and the separated signal, a correlation coefficient is involved. From the simulations and the correlation coefficient calculation, the superimposed signal, i.e. sine wave, has 0.9 or higher correlation coefficient for the frequency below 250Hz. However, the speech signal has 0.9 or higher correlation coefficient when the superimposed signal is lower than 150Hz. Even though with some limitations, there is a big possibility to send a data in the telephone system simultaneously with the voice transmission.*

**Keywords :** Data over voice, correlation coefficient, telephone, speech

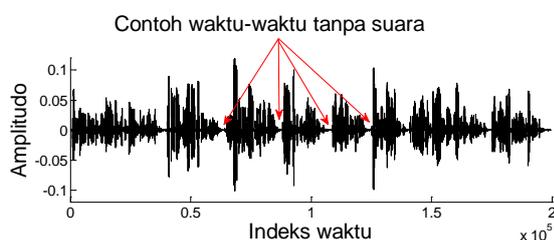
## PENDAHULUAN

Informasi merupakan salah satu kebutuhan manusia dalam rangka bersosialisasi dengan lingkungannya. Informasi dikirimkan oleh sumber ke penerima dengan cara dan bentuk bentuk yang beranekaragam. Salah satu cara manusia bertukar informasi adalah melalui suara.

Suara merupakan salah satu bentuk perpindahan informasi melalui merambat energi di suatu medium. Untuk jarak yang jauh, pengiriman suara dilakukan dengan menggunakan alat bantu yang salah satunya dikenal dengan nama telepon. Sistem telepon merupakan salah satu peralatan komunikasi yang mengalami perkembangan sangat pesat sejak ditemukan di akhir abad 19 oleh Alexander Graham Bell (Zlatka, 2010). Melalui media ini informasi yang dikirimkan dapat berupa suara analog maupun data digital.

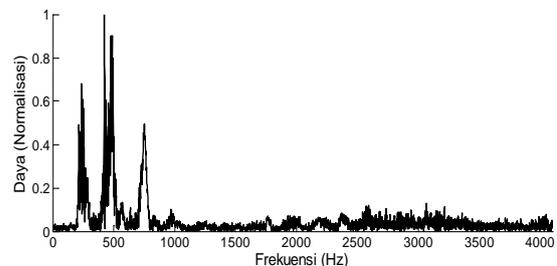
Menurut Zlatka, 2010, di banyak negara umumnya infrastruktur sambungan jalur telepon tersedia lebih baik daripada jaringan komunikasi lainnya seperti sistem selular dan internet. Selain itu jangkauannya yang luas menjadikannya andalan sistem komunikasi di daerah pedalaman. Untuk itu optimalisasi penggunaan jaringan pesawat telepon selain sebagai pertukaran informasi suara menjadi suatu yang wajib dikembangkan.

Percakapan yang dilakukan melalui pesawat telepon pada umumnya mempunyai karakteristik sebagai berikut. Pertama bahwa tidak setiap saat/waktu ada informasi suara di jalur telepon tersebut, atau ada saat-saat dimana individu yang melakukan komunikasi diam semua, walaupun mereka sedang berkomunikasi. Hal ini dapat dimaklumi karena perlunya intonasi dan jeda waktu antar kata atau kalimat ketika manusia berbicara. Gambar 1 menunjukkan contoh sinyal suara yang dikeluarkan. Dari gambar tersebut tampak adanya jeda-jeda waktu antar kata yang cukup banyak/sering.



Gambar 1. Contoh sinyal suara percakapan

Kedua, spektrum yang tersedia di jalur telepon ternyata tidak seluruhnya digunakan (Parker, 1988). Keadaan ini menjadikan efisiensi spektrum jalur telepon rendah (Kato, 2000 dan Gorman, 2000). Untuk memperjelas hal ini disajikan analisis sebagai berikut. Suara manusia mempunyai frekuensi dari 20Hz sampai dengan 20kHz. Pada sistem telepon, spektrum suara ini dimampatkan menjadi 300Hz sampai dengan 3300Hz. Dengan mempertimbangkan *bandguard*, maka untuk kanal transmisi suara di telepon umumnya adalah 4kHz. Dengan menggunakan kriteria Nyquist, frekuensi cuplik untuk pesawat telepon sekitar 8kHz. Untuk mengetahui penggunaan spektrum tersebut, sebuah suara percakapan ditransformasikan ke kawasan frekuensi dengan menggunakan *Fast Fourier Transform* (FFT) 8192 titik. Proses FFT lebih rinci dapat dilihat di Rabiner dan Schafer, 2011. Rata-rata hasil dari transformasi ini ditunjukkan di gambar 2 berikut ini.



Gambar 2. Contoh penggunaan spektrum di telepon

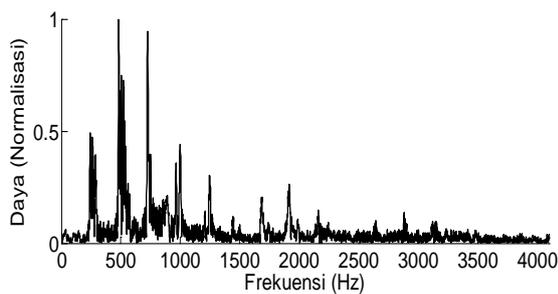
Pada Gambar 2 di atas terlihat bahwa suara percakapan sebagian besar terkonsentrasi di frekuensi 300Hz sampai dengan 800Hz. Sehingga ada bagian yang hampir tidak digunakan yaitu di bagian bawah 300Hz dan pita frekuensi di atas 800Hz.

Berdasarkan analisis dua karakteristik percakapan yang terjadi di jalur telepon di atas, maka sangat dimungkinkan untuk menumpangkan/menyisipkan informasi lain selain suara di jalur telepon pada saat percakapan berlangsung. Skema yang dimungkinkan adalah: (1) penyisipan di kawasan waktu, (2) penyisipan data di kawasan frekuensi.

Penyisipan di kawasan waktu mempunyai kompleksitas yang lebih tinggi

karena memerlukan deteksi daya sinyal pada saat rentang waktu tertentu. Hasil deteksi ini akan menentukan apakah saat itu dimungkinkan disisipi data atau tidak. Perhitungan daya ini juga menyebabkan penambahan tunda waktu suara yang diterima sehingga berakibat pada penurunan tingkat kenyamanan komunikasi yang dilakukan. Sehingga penyisipan di kawasan frekuensi dirasa lebih menguntungkan daripada di kawasan waktu.

Terdapat dua alternatif penyisipan data di kawasan frekuensi, yaitu pada frekuensi di bawah frekuensi suara percakapan, atau penyisipan di atas frekuensi suara percakapan. Penyisipan data di atas frekuensi suara percakapan lebih sulit daripada penyisipan di frekuensi bawah karena batas frekuensi tertinggi percakapan tidak dapat ditentukan dengan pasti. Hal ini dikarenakan setiap orang mempunyai karakteristik suara yang berbeda-beda. Selain itu derau di frekuensi atas lebih besar daripada noise di frekuensi bawah. Gambar 3 menunjukkan penggunaan spektrum untuk objek yang berbeda. Pada gambar tersebut tampak bahwa frekuensi atas yang tidak digunakan berbeda dengan spektrum di Gambar 2. Daya noise di frekuensi atas juga relatif lebih besar dari frekuensi bawah.

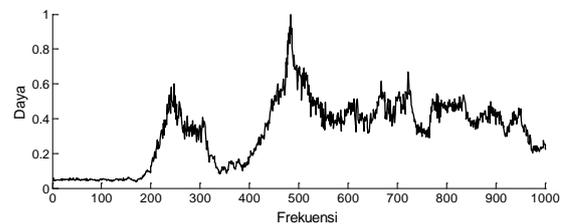


Gambar 3. Contoh lain penggunaan spektrum di telepon

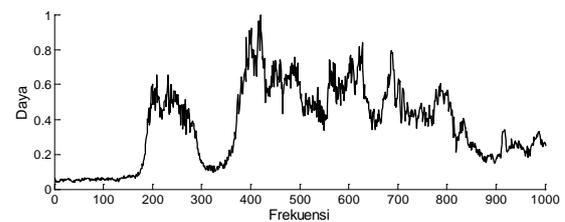
Berdasarkan analisis di atas, maka skema paling sederhana untuk menyisipkan informasi pada percakapan telepon adalah di frekuensi bawah. Untuk memastikan batas frekuensi bawah yang bebas dari interferensi suara percakapan, dilakukan beberapa pengambilan sample dan menganalisisnya di kawasan frekuensi. Hasil spektrum yang diperoleh ditunjukkan di Gambar 4.

Dari hasil keempat sample tersebut, terlihat bahwa frekuensi bawah yang bebas

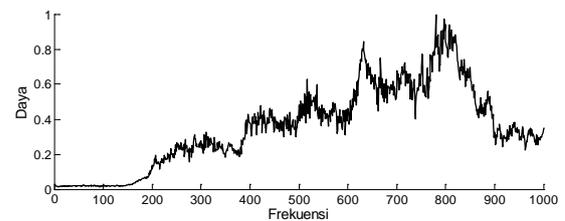
dari interferensi suara percakapan maksimum berkisar antara 100-200Hz. Sehingga frekuensi data yang ditumpangkan tidak boleh melebihi batas tersebut. Metode penumpangan dan pemisahan data pada saluran telepon ini selanjutnya disebut dengan “*data over voice*” yang lebih lanjut dibahas pada bab berikutnya.



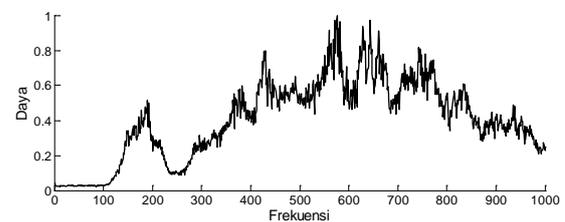
Sample (a)



Sample (b)



Sample (c)



Sample (d)

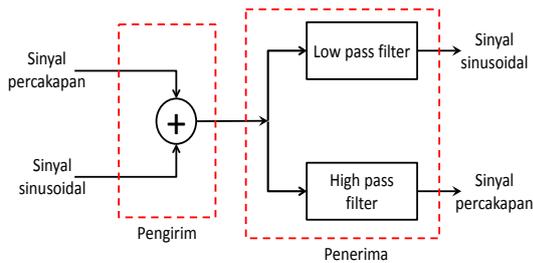
Gambar 4. Hasil analisis beberapa sample di kawasan frekuensi.

## METODE

Pada bagian ini dijelaskan mengenai model sistem *data over voice* pada saluran telepon. Dari analisis di bagian sebelumnya diketahui bahwa sinyal yang ditumpangkan mempunyai frekuensi maksimum antara 100-200Hz. Sedangkan suara yang ditumpangki

merupakan suara percakapan yang akan dimasukkan ke jalur telepon yang dicuplik dengan frekuensi cuplikan 8192Hz.

Gambar 5 memperlihatkan model sistem *data over voice* yang terdiri dari bagian pengirim dan penerima. Pada bagian pengirim terjadi proses pencampuran antara sinyal suara dengan data sedangkan di bagian penerima terjadi pemisahan antara keduanya. Pencampuran dilakukan dengan menggunakan rangkaian penjumlahan (*adder*) analog kemudian masuk ke ADC. Sedangkan pemisahan dilakukan dengan menggunakan filter FIR.



Gambar 5. Model sistem *data over voice*.

Data yang dikirimkan dalam pengujian ini merupakan sinyal sinusoidal yang amplitudonya tetap namun frekuensinya dibuat sebagai variabel peubah. Pengujian dimulai dari frekuensi 10Hz sampai dengan frekuensi 500Hz dengan kenaikan 10Hz. Amplitudo dibuat tetap yaitu '1' satuan.

Sinyal suara yang dijadikan sinyal percakapan merupakan hasil perekaman dengan menggunakan frekuensi sampling 8192Hz. Panjang perekaman adalah 25 detik. Dengan demikian panjang vektor suara maupun sinyal sinusoidal adalah 204800. Sinyal suara dinormalisasi sehingga maksimum simpangannya bernilai '1' satuan.

FIR filter terdiri dari filter lolos rendah (*low pass filter = LPF*) dan filter lolos tinggi (*high pass filter = HPF*). Filter dirancang dengan bantuan FDTTool yang ada di Matlab. Parameter digunakan dalam perancangan LPF dan HPF dinyatakan di Tabel 1 dan Tabel 2.

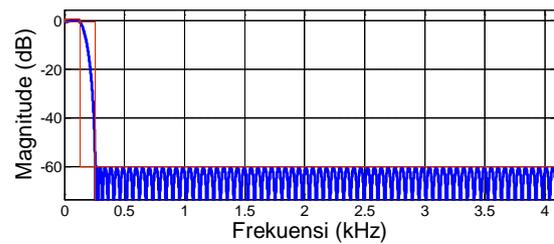
Tabel 1. Parameter perancangan FIR LPF

Frekuensi cuplik	8192Hz
Frekuensi lolos ( $f_{pass}$ )	128Hz
Frekuensi stop ( $f_{stop}$ )	256Hz
Jenis filter	Single rate FIR
Struktur filter	Direct form FIR
Orde mode	Minimum
Pelemahan di $f_{stop}$	-60dB

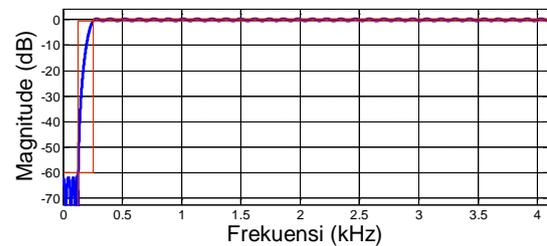
Tabel 2. Parameter perancangan FIR HPF

Frekuensi cuplik	8192Hz
Frekuensi lolos ( $f_{stop}$ )	128Hz
Frekuensi stop ( $f_{pass}$ )	256Hz
Jenis filter	Single rate FIR
Struktur filter	Direct form FIR
Orde mode	Minimum
Pelemahan di $f_{stop}$	-60dB

Berdasarkan koefisien FIR filter yang dibangkitkan, tanggapan impuls FIR yang dihasilkan ditunjukkan pada Gambar 6 untuk LPF dan Gambar 7 untuk HPF.



Gambar 6. Tanggapan impuls FIR-LPF



Gambar 7. Tanggapan impuls FIR-HPF

Untuk mengetahui tingkat kemiripan sinyal hasil pemisahan ( $\hat{s}$ ) dengan sinyal aslinya ( $s$ ), dilakukan perhitungan koefisien korelasi ( $\rho$ ). Menurut Anderson, 1942, koefisien korelasi dinyatakan dengan persamaan:

$$\rho = \frac{\frac{1}{N} \sum [(s - \bar{s}) \cdot (\hat{s} - \bar{\hat{s}})]}{\sqrt{\frac{1}{N} \sum (s - \bar{s})^2} \cdot \sqrt{\frac{1}{N} \sum (\hat{s} - \bar{\hat{s}})^2}} \quad (1)$$

dengan  $\bar{s}$  adalah nilai ekspektasi dari data sinyal  $s$ .

Nilai koefisien korelasi bernilai antar -1 sampai dengan 1. Nilai -1 berarti kedua Koefisien korelasi 1 berarti bahwa kedua sinyal tersebut sama persis. Nilai -1 berarti bahwa kedua sinyal tersebut sama namun berbeda polaritas. Sedangkan koefisien

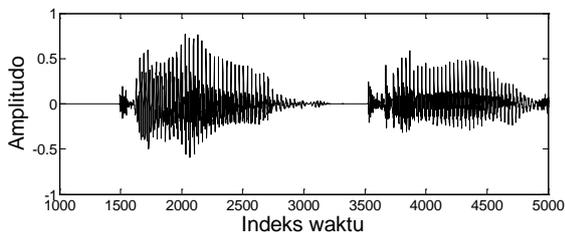
korelasi 0 berarti kedua sinyal tersebut tidak mempunyai kemiripan sama sekali.

## HASIL DAN PEMBAHASAN

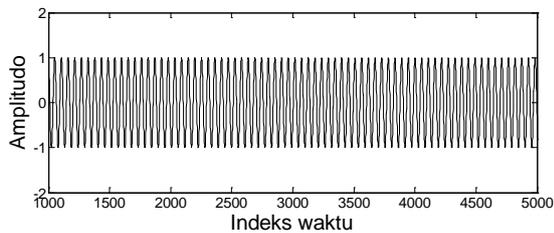
Pada bagian ini dipaparkan data hasil simulasi dan pengamatan, serta dilanjutkan dengan analisisnya.

### Hasil Pencampuran

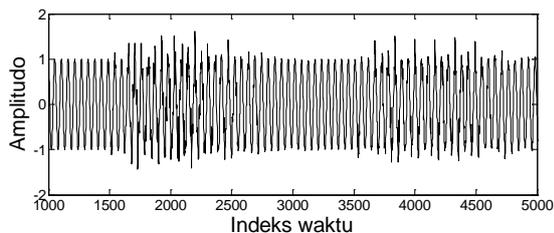
Pencampuran dilakukan antara sinyal suara dengan sinyal data berupa sinusoidal di sisi pengirim. Gambar 8 menunjukkan gambar sinusoidal ( $f = 150\text{Hz}$ ) dan suara percakapan sebelum dan sesudah dicampurkan. Dari gambar tersebut tampak bahwa sinyal percakapan seakan-akan lenyap tertelan sinyal sinusoidal. Perlu diketahui bahwa sinyal sinusoidal dan sinyal percakapan masing-masing dilakukan normalisasi sebelum dilakukan proses penjumlahan. Sehingga amplitudo maksimum sinyal sebelum digabungkan adalah '1', sedangkan setelah pencampuran adalah '2'.



(a) Sinyal percakapan



(b) Sinyal sinusoidal



(c) Hasil pencampuran (a) dan (b)

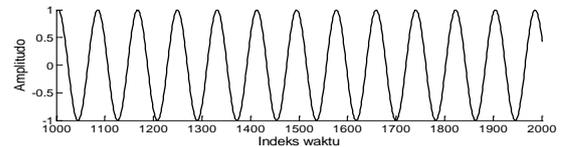
Gambar 8. Sinyal percakapan sebelum dan sesudah pencampuran

### Pemisahan Sinyal

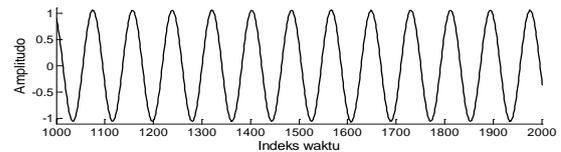
Sebagaimana dijelaskan sebelumnya, pemisahan sinyal percakapan dan sinyal

sinusoidal dilakukan dengan menggunakan filter FIR. Spesifikasi filter yang digunakan telah dijelaskan pada bagian sebelumnya.

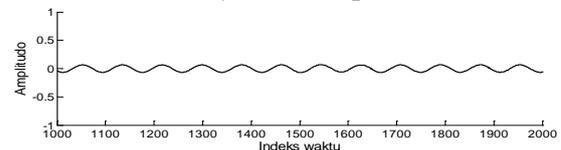
Pada percobaan ini dibuat frekuensi sinusoidal 80Hz. Gambar 9 memperlihatkan sinyal sinusoidal asli (a) dan hasil pemisahan (b). Dari gambar tersebut, terlihat kemiripan yang tinggi yang terbukti dengan nilai *error* (c) yang kecil.



(a) Sinyal asli sebelum penggabungan



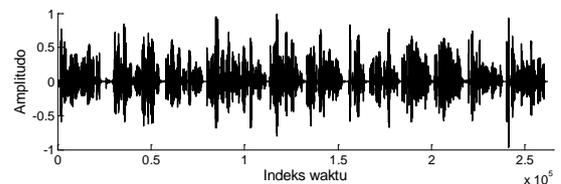
(b) Sinyal setelah pemisahan



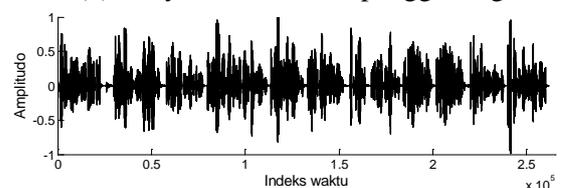
(c) Selisih/error

Gambar 9. Potongan sinyal sinusoidal sebelum dan sesudah proses pemisahan

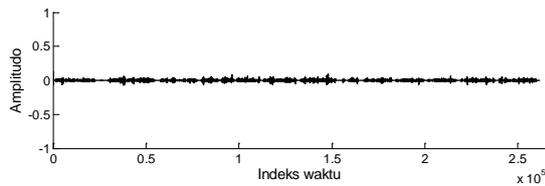
Hasil pemisahan sinyal percakapan diperlihatkan di Gambar 10. Gambar 10 (a) adalah sinyal asli sebelum dicampurkan, gambar 10 (b) adalah sinyal percakapan setelah pemisahan, dan selisih antara keduanya ditampilkan di gambar 10 (c).



(a) Sinyal asli sebelum penggabungan



(b) Sinyal setelah pemisahan



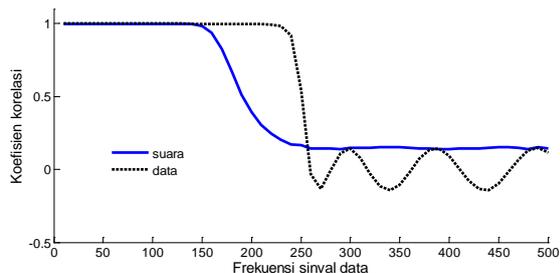
(c) Selisih/error

Gambar 10. Potongan sinyal percakapan sebelum dan sesudah pemisahan

Berdasarkan gambar 9 dan Gambar 10 terlihat adanya nilai kesalahan yang kecil. Nilai kesalahan ini lebih didominasi karena perbedaan amplitudo antara sinyal sebelum dan sesudah filter, namun secara substansi masih memuat informasi yang sama. Hal ini terbukti dengan hasil perhitungan koefisien korelasi yang bernilai di atas 0,99.

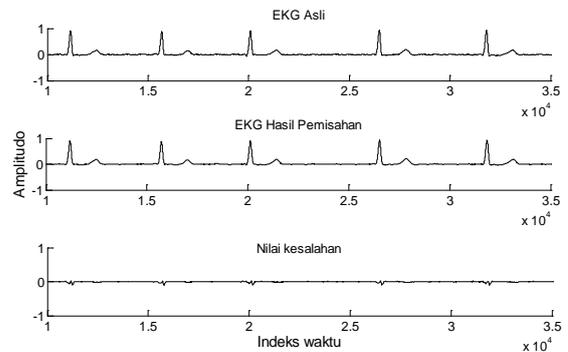
Untuk mengetahui pengaruh frekuensi data yang ditumpangkan dengan hasil pemisahan, dilakukan perhitungan koefisien korelasi untuk nilai frekuensi sinusoidal yang berbeda-beda. Nilai frekuensi yang diujikan bernilai 10Hz sampai dengan 500Hz dengan rentang kenaikan 10Hz.

Pengujian yang dilakukan memperoleh hasil sebagaimana diperlihatkan pada Gambar 11. Dari gambar tersebut dapat dimaknai bahwa sinyal data hasil pemisahan masih memiliki kemiripan yang tinggi dengan sinyal data aslinya apabila sinyal sata tersebut mempunyai frekuensi di bawah 250Hz. Sedangkan pada batas tersebut sinyal suara percakapan telah mengalami distorsi yang sangat besar sehingga koefisien korelasinya mendekati nol. Suara percakapan hasil pemisahan mempunyai koefisien korelasi di atas 0,9 apabila data yang ditumpangkan kepadanya mempunyai frekuensi tidak lebih dari 150Hz.



Gambar 11. Hasil perhitungan koefisien korelasi terhadap perubahan frekuensi sinyal data

Berdasarkan hasil yang telah dipaparkan di atas, sistem data over voice pada saluran telepon dapat berjalan dengan baik dengan syarat informasi/data yang ditumpangkan mempunyai frekuensi tidak lebih dari 150Hz. Contoh aplikasinya adalah untuk sistem monitoring yang laju update-nya yang tidak besar seperti suhu, tinggi permukaan air, dan kecepatan rata-rata angin. Selain itu dapat dimanfaatkan untuk analisis penyakit secara jarak jauh (Moran, 2006), dan *interactive voice response* (Joyanta, 2012).

Gambar 12. Pegiriman sinyal ECG dengan sistem *data over voice* melalui jalur telepon

Salah satu aplikasi yang dapat menggunakan sistem ini adalah menggunakan sistem ini adalah monitoring detak jantung. Sinyal *electrocardiograph* (ECG) mempunyai fekuensi di bawah 100Hz, sehingga sangat dimungkinkan untuk diaplikasikan pada sistem *data over voice*. Gambar 12 menunjukkan sinyal ECG yang dikirimkan dengan sistem *data over voice*. Dari hasil simulasi tersebut tampak kemiripan sinyal sebelum dikirim dan setelah dipisahkan dengan nilai koefisien korelasi 0.99.

## SIMPULAN DAN SARAN

Pada makalah ini telah dibahas mengenai beberapa kemungkinan penumpangan data ke sinyal percakapan melalui jalur telepon atau *data over voice*. Dari beberapa kemungkinan yang ada, penyisipan data/informasi di frekuensi bawah merupakan alternatif terbaik. Sistem yang dibangun terdiri dari pecampur sinyal di bagian pengirim dan pemisah sinyal di bagian penerima. Pemisahan sinyal dilakukan dengan menggunakan filter FIR-LPF dan HPF dengan batas frekuensi di 128Hz dan 256Hz. Hasil simulasi yang dilakukan menunjukkan unjuk kerja yang baik dengan nilai koefisien korelasi diatas 0,9 untuk

sinyal data yang mempunyai frekuensi dibawah 150Hz.

Saran perbaikan yang mungkin dilakukan adalah menggabungkan LPF dan HPF untuk meminimalkan kompleksitas. Kemungkinan pengembangan yang lain adalah melakukan filtering pada sinyal percakapan sebelum ditumpangi sinyal data. Hal ini untuk menyakinkan bahwa keduanya tidak saling berinterferensi.

#### UCAPAN TERIMAKASIH

Penelitian ini merupakan bagian dari hibah penelitian unggulan perguruan tinggi direktorat pendidikan tinggi (DIKTI) tahun 2015

#### DAFTAR PUSTAKA

Anderson, R. L. 1942. Distribution of the Serial Correlation Coefficient. *The Annals of Mathematical Statistics*. 13(1): 1-13

Gorman, M. G. (2000). *U.S. Patent No. 6,141,356*. Washington, DC: U.S. Patent and Trademark Office.

Joyanta B. etc. 2012. Telephony Speech Recognition System: Challenges. Makalah disajikan dalam *National Conference on Communication Technologies & its impact on Next Generation Computing CTNGC 2012*. hlm. 30-36.

Kato, H. (2000). *U.S. Patent No. 6,044,266*. Washington, DC: U.S. Patent and Trademark Office.

Moran, R. J., dan Reilly, R. B. 2006. Telephony-Based Voice Pathology Assessment Using Automated Speech Analysis. *IEEE Trans. On Biomedical Engineering*. 53(3):468-477.

Parker, S. S., Scheffler, T. H., dan Schirmer, P. W. (1988). *U.S. Patent No. 4,718,082*. Washington, DC: U.S. Patent and Trademark Office.

Rabiner L. B., dan Schafer R. W. 2011. *Teory and Applications of Digital Speech Processing*. 5<sup>th</sup> edition. New Jersey. Pearson.

Zlatka M. 2010. An Architecture of a Telephone-based System for Speech Data Collection. *Research Project Report*. Karlsruhe Institute of Technology.