

SPEMBUATAN KODE BLOK, KODE KOMA DAN SHANON-FANO SERTA EFISIENSI PENGKODEANNYA

Fadilla Zennifa, Mutia Yanelda, Wera Heryani, Rennu Alvin, Anisa Putriana¹ Fitrilina²

1. Mahasiswa Teknik Elektro Universitas Andalas Program Studi Telekomunikasi

2. Dosen Teknik Elektro Universitas Andalas

ABSTRAK

Untuk menghasilkan suatu sistem komunikasi yang handal, dalam artian bebas dari error, perlu diterapkan suatu algoritma kode yang dapat mengkoreksi (error detection) sekaligus memperbaiki kesalahan bit (error correction). Algoritma pengkodean digunakan untuk mendeteksi kesalahan pada sumber informasi serta mengkompresi data untuk menghemat memori penyimpanan. Kompresi data tersebut antara lain menggunakan beberapa jenis teknik pengkodean diantaranya adalah block code, kode koma dan shanon-fano yang telah menjelaskan dan membantu kita dalam menangani file-file besar. dan contoh-contoh sederhana serta penerapannya. Diakhir paper kita dapat membandingkan setiap jenis kode dengan algoritmanya masing-masing sehingga dapat dikembangkan hingga menjadi lebih baik

Kata kunci: kode blok, kode koma, kode Shannon-Fano

1.1 Latar Belakang

Pengiriman informasi menuju ke penerima memerlukan suatu proses yang handal karena data yang dilewatkan melalui saluran komunikasi fisik dalam sebuah sistem komunikasi dapat mengalami berbagai jenis gangguan, distorsi dan interferensi yang dapat mengakibatkan output dari saluran komunikasi berbeda dengan apa yang sudah di input kan.

Data yang masuk pada sistem komunikasi dari sumber informasi (information

source) akan diproses pertama kali oleh suatu pengkodean sumber (source encoder) yang dirancang untuk mengkonversi informasi sumber menjadi bentuk kode. Biasanya, source encoder akan mengubah output sumber menjadi sebuah deretan informasi (information sequence) d. Output sumber dapat berupa sebuah gelombang kontinyu (continuous waveform) atau sebuah deretan dari symbol diskrit (sequence of discrete symbols).

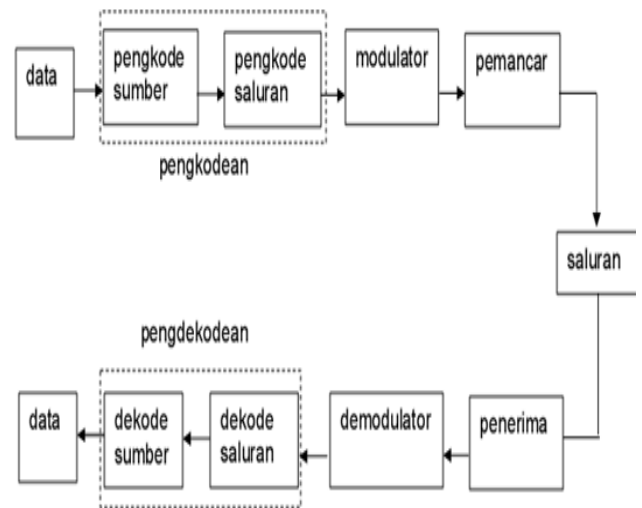
1.2 Tujuan

Tujuan dari *paper* ini adalah untuk mengetahui teknik pengkodean seperti kode blok, kode koma dan kode Shanon –Fano. Serta untuk mengetahui efisiensi dari masing masing jenis pengkodean tersebut

2. Teori Penunjang

Teori Source Encoding merupakan salah satu dari ketiga teorema dasar dari teori informasi yang diperkenalkan oleh Shannon (1948). Teori Source Encoding mencanangkan sebuah limit dasar dari sebuah ukuran dimana keluaran dari sebuah sumber informasi dapat dikompresi tanpa menyebabkan probabilitas error yang besar.

Pengkodean sumber merupakan semua proses pengolahan sumber informasi agar dapat sesuai dengan proses berikutnya. Berikut merupakan bagan suatu sistem pentransmisian, dimana terdapat adanya source coding (pengkodean sumber), pengkodean kanal, multiple akses modulasi dan demodulasi. Source Encoding dapat dilihat pada gambar 1 berikut ini :



Gambar 1. Diagram Blok Source Encoding

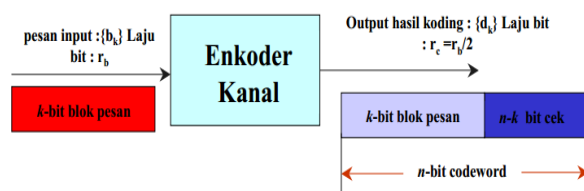
Source coding merupakan klasifikasi teknik kompresi, dimana pada pengkodean sumber dilakukan proses kompresi agar informasi yang dikirimkan tidak menduduki bandwidth yang lebar, baik saat masih berupa data analog maupun setelah menjadi data digital.

Kompresi berarti memampatkan atau mengecilkan ukuran file dimana merupakan proses pengubahan sekumpulan data menjadi suatu bentuk kode untuk menghemat kebutuhan tempat penyimpanan dan waktu untuk transmisi data.

2.1 Kode Block

Ciri khas dari kode block ini adalah semua kode word mempunyai panjang yang sama (panjang kode sama). Block codes menerima informasi dalam blok k-bit yang berurutan.

Setiap blok, menambah $n-k$ redundant bit yang berkaitan dengan aljabar bit pesan k . oleh karena itu, ini membuat secara keseluruhan blok encoded dari bit n , dengan $n > k$. Blok bit n disebut juga sebagai code word, dan n disebut sebagai panjang blok dari kode. Sederhananya, kita menyebut seperti block code (n,k) code block.



Pada kode blok, sebanyak r bit cek ditambahkan pada setiap k -bit pesan \rightarrow menjadi n -bit codeword.

Yang termasuk dalam kode blok adalah kode blok linier sistematis, kode siklik.

Dalam Bentuk Matrik dapat ditulis ;

$C = D G \rightarrow$ dimana G adalah matrik generator dari code

$G = [I_k : P]_{k \times n} \rightarrow$ contoh matrik generator untuk $(6,3)$

I_k : matrik identitas. $\rightarrow P =$ parity

Pada enkoder akan dibangkitkan kode, sesuai dengan matriks generatormya : G

$C \rightarrow$ vektor kode yang ditransmisikan pada kanal bernoise.

$R \rightarrow$ Noise Vektor yang diterima

$$R = C + E$$

$E \rightarrow$ Error Vektor yang diterima

Untuk deteksi/koreksi kesalahan, pada dekoder akan dihitung

syndrome : $S = CH^T \rightarrow$ untuk daerah penerima $\rightarrow S = RH^T$

H^T transpose matriks *parity check*, H berkaitan dengan matriks G , sbb :

$$G = [I_k : P] \iff H = [P^T : I_{n-k}]$$

Contoh : Kode Blok (7,4)

Dengan:

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

$I_4 \quad P$

$$H = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

$P^T \quad I_3$

$$H^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Untuk pesan : $D = (1 \ 0 \ 1 \ 1) \rightarrow$ codeword : $C = DG = (1 \ 0 \ 1 \ 1 \ 0 \ 0 \ 1)$

Bila pada penerima, codeword C mengalami error menjadi R , sbb:

$R = (1 \ 0 \ 0 \ 1 \ 0 \ 0 \ 1) = (1 \ 0 \ 1 \ 1 \ 0 \ 0 \ 1) + (0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0) = C + E$

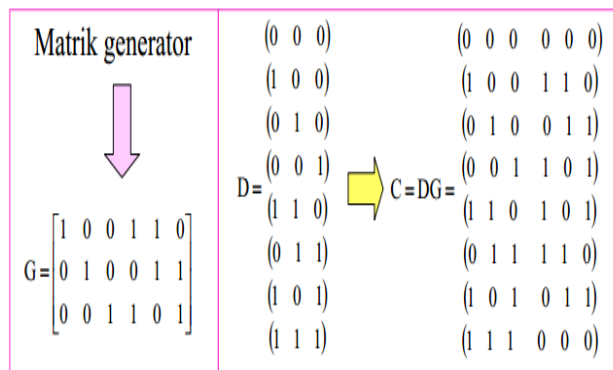
Pada dekoder akan dihitung *syndrome* : $S = RH^T = (101)$

Posisi salah di bit ke 3

Dekoding kode blok menggunakan *Table Lookup*

Misal kode blok (6,3) \Rightarrow terdapat $2^3=8$ codeword, karena error, codeword ini berubah menjadi salah satu dari 2^6-1 codeword. Dengan perhitungan *syndrome* dan menggunakan tabel, error dapat dikoreksi.

Contoh : untuk kode blok (6,3) dengan matrik generator:



Kemampuan deteksi dan koreksi kode blok linier

Kemampuan koreksi $\Rightarrow [(d_{\min}-1)/2]$

Kemampuan deteksi $\Rightarrow d_{\min}-1$

d_{\min} = jarak minimum = bobot minimum

codeword bukan nol \rightarrow komponen non zero

Contoh :

code word	bobot
0 0 0 0 0 0	0
0 0 1 1 1 0	③
0 1 0 1 0 1	③
0 1 1 0 1 1	4
1 0 0 0 1 1	③
1 0 1 1 0 1	4
1 1 0 1 1 0	4
1 1 1 0 0 0	③

\Rightarrow Bobot minimum = 3 $\Rightarrow d_{\min}=3$

- Kemampuan koreksi= $[(d_{\min}-1)/2] = 1$ bit
- Kemampuan deteksi = $d_{\min}-1 = 2$ bit

2.2 Kode Komma

Kode koma adalah suatu simbol kode spesifik yang berulang pada proses dekode.

Untuk setiap distribusi geometrik (distribusi eksponensial pada bilangan bulat), kode Golomb optimal. $1/n^2$ Dengan kode universal, distribusi Sekitar implisit kuasa hukum seperti (lebih tepatnya, distribusi Zipf). Untuk kode Fibonacci, distribusi secara implisit Sekitar, dengan $1/n^q$ dimana $q = 1/\log_2(\varphi) \simeq 1.44$, adalah rasio emas. Untuk koma kode terner (yaitu, pengkodean dalam basis 3, Diwakili dengan 2 bit per simbol), distribusi implisit adalah kuasa hukum dengan $q = 1 + \log_3(4/3) \simeq 1.26$. Dengan demikian distribusi ini telah dekat-optimal kode Dengan hukum mereka masing-masing kekuasaan.

2.3. Shanon- Fano

Teknik coding Shannon Fano merupakan salah satu algoritma pertama yang tujuannya adalah membuat code word dengan redundansi minimum. Ide dasar dari membuat code word dengan variable-code length, seperti Huffman codes, yang ditemukan beberapa tahun kemudian. Seperti yang disebutkan di atas, Shannon Fano coding didasarkan pada variable length-word, yang berarti beberapa simbol pada pesan (yang akan dikodekan) direpresentasikan dengan code word yang lebih pendek dari simbol yang ada di pesan. Semakin tinggi

probabilitasnya, maka code word semakin pendek. Dalam memperkirakan panjang setiap codeword maka dapat ditentukan dari probabilitas setiap simbol yang direpresentasikan oleh codeword tersebut. Shannon Fano coding menghasilkan codeword yang tidak sama panjang, sehingga kode tersebut bersifat unik dan dapat didekodekan. Cara efisien lainnya dalam variable-length coding adalah Shannon-Fano encoding. Prosedur dalam Shannon-Fano encoding adalah :

- Menyusun probabilitas simbol dari sumber dari yang paling tinggi ke yang paling rendah.
- Membagi menjadi 2 bagian yang sama besar, dan memberikan nilai 0 untuk bagian atas dan 1 untuk bagian bawah.
- Ulangi langkah ke 2, setiap pembagian dengan probabilitas yang sama sampai dengan tidak mungkin dibagi lagi
- Encode setiap simbol asli dari sumber menjadi urutan biner yang dibangkitkan oleh setiap proses pembagian tersebut.

Teorema

- Untuk suatu prefix binary code dengan panjang rata-rata codeword maka

$$L_{avg} \geq H(S)$$

- Terdapat prefix binary code dimana

$$L_{avg} < H(S) + 1$$

Shannon's Fundamental Theorem of Discrete Noiseless Coding

Untuk sumber S dengan entropy $H(S)$, dimungkinkan mengalokasikan codeword deretan k simbol dengan kondisi prefix dipenuhi, dan panjang rata-rata codeword L_k

$$H(S) \leq \frac{L_k}{k} < H(S) + \frac{1}{k}$$

Efisiensi macam-macam code diukur dengan:

$$effisiensi = \frac{H(S)}{L_{avg}} \cdot 100\%$$

Dimana :

$H(s)$ adalah entropi Sumber

L_{avg} adalah panjang kode rata-rata

3. Analisa Cara Pembuatan Kode Blok, Kode Koma Dan Shanon-Fano Serta Efisiensi Pengkodeannya

Berikut ini adalah contoh-contoh dari jenis teknik pengkodean:

1. Kode Blok
2. Kode Koma

Contoh:

Sumber Simbol	P	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆
A	0.6	00	0	0	0	0	0
B	0.25	01	10	10	01	10	10
C	0.1	10	110	110	011	11	11
D	0.05	11	1110	111	111	01	0

☀ Manakah yang merupakan kode blok

Rata-rata panjang kode adalah

$$C_1 : 2 \times 0.6 + 2 \times 0.25 + 2 \times 0.1 + 2 \times 0.05 = 2$$

$$C_2 : 1 \times 0.6 + 2 \times 0.25 + 3 \times 0.1 + 4 \times 0.05 = 1.60$$

$$C_3 : 1 \times 0.6 + 2 \times 0.25 + 3 \times 0.1 + 3 \times 0.05 = 1.55$$

$$C_4 : 1 \times 0.6 + 2 \times 0.25 + 3 \times 0.1 + 3 \times 0.05 = 1.55$$

$$C_5 : 1 \times 0.6 + 2 \times 0.25 + 2 \times 0.1 + 2 \times 0.05 = 1.40$$

$$C_6 : 1 \times 0.6 + 2 \times 0.25 + 2 \times 0.1 + 1 \times 0.05 = 1.35$$

Maka yang merupakan kode blok adalah

C₁ karena memiliki panjang yang sama

☀ Berdasarkan contoh diatas yang termasuk kode koma adalah C₂ dengan 0 berperan sebagai koma

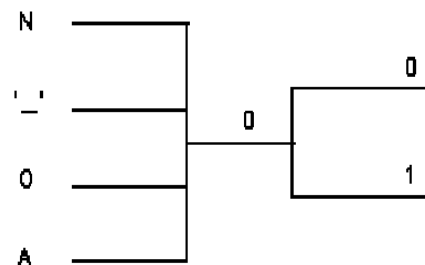
Contoh Apabila kita mempunyai kalimat:

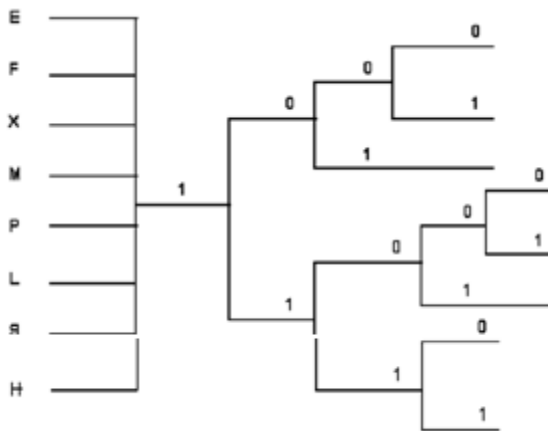
"EXAMPLE OF SHANNON FANO"

Pertama, kita kalkulasi probabilitas dari setiap simbol:

Symbol	Codewords
E	2/23
X	1/23
A	3/23
M	1/23
P	1/23
L	1/23
O	3/23
F	2/23
S	1/23
H	1/23
N	4/23
space	3/23

Coding tree dari contoh tersebut adalah :





Contoh untuk efisiensi shanon fano

- Contoh

$$S = \{A, B, C, D, E\}$$

$$P = \{0.35, 0.17, 0.17, 0.16, 0.15\}$$

- Pengkodean Shannon-Fano:

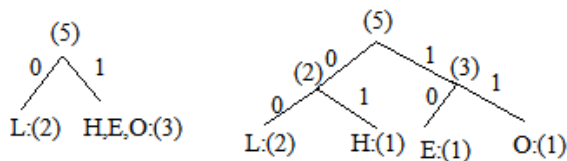
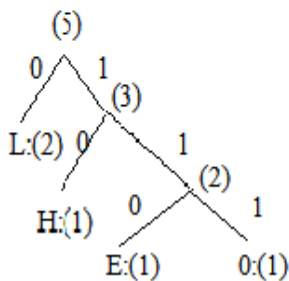
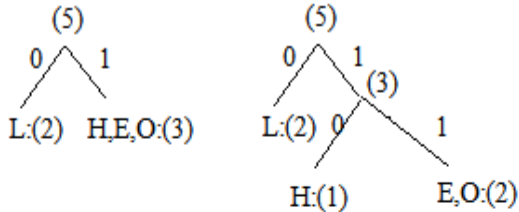
- Bagi S kedalam s_1 dan s_2 (pilih yang memberikan perbedaan $p(s_1)$ dan $p(s_2)$ terkecil
- $s_1 = (A,B) \rightarrow p(s_1) = p(A) + p(B) = 0,52$
- $s_2 = (C,D,E) \rightarrow p(s_2) = p(C) + p(D) + p(E) = 0,48$
- Panggil ShannonFano()

Contoh lain H E L L O

Simbol : H E L O

Jumlah 1 2 1

Binary tree nya adalah:



x_i	p_i	codeword
A	.35	00
B	.17	01
C	.17	10
D	.16	110
E	.15	111

- Panjang code rata-rata:

$$L_{sh} = 0,35*2 + 0,17*2 + 0,17*2 + 0,16*3 + 0,15*3 = 2,31$$

- Efisiensi = $(2,23284/2,31)*100 = 96,66\%$

[C%20code%20comma&f=false](#) diakses pada 31 Maret 2013 pukul 15.45

4. KESIMPULAN

Diantara 3 jenis kode di atas yang paling efisien adalah : teknik coding shannon fano karena beberapa simbol pada pesan yang akan dikodekan direpresentasikan dengan code word yang lebih pendek dari simbol yang ada di pesan. Semakin tinggi probabilitasnya, maka code word semakin pendek

DAFTAR PUSTAKA

http://it.becs.ac.in/content/tuhina_samanta/infcod/Lec2Source_Coding.pdf diakses pada 31 Maret 2013 pukul 16.00

http://books.google.co.id/books?id=0CI8bd0upS4C&pg=PA189&lpg=PA189&dq=source+coding,+code+comma&source=bl&ots=IIDKlHpj7V&sig=Ln5AH1z_cg8fgde3khMzwIT2UmE&hl=en&sa=X&ei=6PZXUcegBtHprQeNkoGwCw&redir_esc=y#v=onepage&q=source%20coding%20