Ефикасност Хафмановог кода при проширењу извора
Huffman Code Efficiencies for Extension of Sourses

Бранко Павловић, Тијана Накићеновић Богојевић
SDC International, Београд
bane@sdc.co.yu, tijana@sdc.co.yu

Садржај – *Познато је да се ефикасност кода може повећати проширењем извора. Што је веће проширење, ефикасност је ближа јединици. У овом раду показано је да се ефикасност не повећава увек следећим проширењем. У неким случајевима следеће проширење има мању ефикасност од претходног, мада увек остаје у теоријским границама. Феномен је показан на променама у бинарном Хафмановом коду.*

Abstract – *It is well known that the efficiency of noisess coding may be improved by coding extensions of the source; for large extension the efficiency is arbitrarily close to unity. This paper shows that the efficiency is not always improved just by coding the next extension. In some cases the code of a larger extension is markedly less efficient than its predecessor, altough always within the theoretical limits of efficiency. We show how the phenomen arises from changes to the Huffman coding tree as the source probabilities change.*

1. УВОД

Посматрамо дискретан извор информација без меморије S са вероватноћама појављивања симбола $\{P_1, P_2, P_3,..\}$. Симболи S_i из скупа S се кодирају речима различитих дужина L_i и преносе кроз канал без сметњи.

Ентропија извора симбола H(S) је:

$$H(S) = \sum_i P_i ld \frac{1}{P_i},$$

(коришћена је уобичајена ознака за логаритам са основом 2).

Средња дужина кодне речи L је:

$$L = \sum_i P_i L_i .$$

Ефикасност кода је:

$$\eta = \frac{H(S)}{L}$$

Из теорије информација је познато да је:

$H(S) \le L$, односно $\eta = \dfrac{H(S)}{L} \le 1$.

Такође је веома добро позната и Прва Шенонова (Shannon) теорема ([1], str. 72), која каже да се ефикасност кода може повећати проширивањем извора, односно груписањем симбола извора у групе од 2,3 или више и кодирањем тих група симбола.
За n-то проширење извора симбола важи:

$\lim\limits_{n\to\infty} \dfrac{L}{n} = H(S)$, (где је са L означена средња дужина кодне речи n-тог проширења), односно важи следеће ограничење за ефикасност кода:

$$1 \ge \eta > 1 - \frac{1}{n}.$$

Можемо закључити да се код довољно великих проширења извора ефикасност кода приближава јединици, односно добија се максимално ефикасан код.

Посматрамо Хафманов (Huffman) код [4]. Квалитет кода се може представити ефикасношћу кода (η) или његовом редундансом. Редунданса кода R је:

$$R = L - H(S).$$

Галагер (Gallager) [3] је показао да је горња граница редундансе Хафмановог кода P + 0,0861, где је P вероватноћа најфреквентнијег симбола. Џонсен (Johnsen) [5] и Капочели (Capocelli) [2] су извели прецизније границе редундансе, проверавајући различите опсеге P.

Наведени аутори су проучавали само основни Хафманов код, јер су сматрали да се проширења извора могу посматрати као основни код са одговарајућим бројем симбола. На основу ограничења за ефикасност кода: $1 \ge \eta > 1 - \dfrac{1}{n}$, очекује се да се ефикасност увек повећава са проширивањем извора, као што тврди Абрамсон ([1], str. 87). Ову тврдњу ћемо проверити на два примера Хафмановог кода.

Пример Хафмановог кода са вероватноћама симбола: P_1=0,85 и P_2=0,15

Прво проширење

Симбол	Вероватноћа	Код
A	0,85	*1*
B	0,15	*0*

Из кода се види да је средња дужина кодне речи 1, али и формула то даје:

$$L=\sum_{i=1}^{2}P_iL_i=0{,}85*1+0{,}15*1=1$$

Ентропија за овај случај је:

$$H(S)=\sum_{i=1}^{2}P_i ld\frac{1}{P_i}=0{,}85ld\frac{1}{0{,}85}+0{,}15ld\frac{1}{0{,}15}=$$
$$=0{,}6098$$

Ефикасност је:

$$\eta=\frac{H(S)}{L}=0{,}6098$$

Друго проширење

Симбол	Веров.	Код	Кодирање			
AA	0,7225	*1*	0,72	*1*	0,72	*1*
AB	0,1275	*00*	0,15	*01*	0,28	*0*
BA	0,1275	*011*	0,13	*00*		
BB	0,0225	*010*				

Сада је средња дужина кодне речи:

$$L=\sum_{i=1}^{4}P_iL_i=0{,}72*1+0{,}13*2+0{,}13*3+0{,}02*3$$
$$=1.4225$$

Ентропија:

$$H(S)=\sum_{i=1}^{4}P_i ld\frac{1}{P_i}=0{,}72ld\frac{1}{0{,}72}+2*0{,}13ld\frac{1}{0{,}13}+$$
$$+0{,}02ld\frac{1}{0{,}02}=1{,}213$$

И ефикасност:

$$\eta=\frac{H(S)}{L}=0{,}8544$$

Аналогним поступком добијају се ефикасности за 3., 4. и 5. проширење. Добијене вредности су приказане у табели 1.

Проширење	Ефикасност
1	0,6098
2	0,8544
3	0,9663
4	0,9837
5	0,9926

Табела 1. Ефикасност извора {0,85, 0,15}

На основу резултата из табеле 1, види се да се ефикасност брзо приближава максимуму тј. јединици.

2. НЕРЕГУЛАРНИ СЛУЧАЈЕВИ

Посматрамо сада бинарни Хафманов код за неке друге вероватноће, очекујући исти закључак.

Пример Хафмановог кода са вероватноћама симбола: P_1=0,8 и P_2=0,2

Аналогним поступком као у претходном примеру добијају се ефикасности за проширења извора. Добијене вредности су приказане у табели 2.

Проширење	Ефикасност
1	0,7219
2	0,9255
3	0,9917
4	0,9745
5	0,9783

Табела 2. Ефикасност извора {0,8, 0,2}

Међутим, табела 2 показује да се прва три проширења понашају према очекивањима, док четврто и пето проширење имају мању ефикасност кода од трећег проширења. Овакви случајеви, када се ефикасност смањује при проширивању извора, зову се “нерегуларна проширења”.

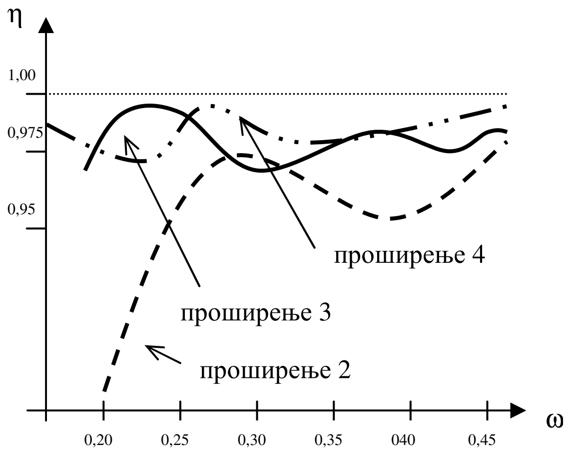
Да би се открила остала нерегуларна проширења, одређена је ефикасност Хафмановог кода (бинарног кода, бинарног извора) са вероватноћама {ω, 1-ω}, где ω варира од 0,05 до 0,5 с кораком од 0,05 за првих пет проширења. Резултати су приказани у табели 3. Подвучена су нерегуларна проширења.

ω	проширење				
	прво	**друго**	**треће**	**четврто**	**пето**
0,05	0,2864	0,4992	0,6610	0,7800	0,8519
0,10	0,4690	0,7271	0,8805	0,9522	0,9767
0,15	0,6098	0,8544	0,9663	0,9837	0,9926
0,20	0,7219	0,9255	0,9917	0,9745	0,9783
0,25	0,8113	0,9615	0,9859	0,9913	0,9921
0,30	0,8813	0,9738	0,9699	0,9882	0,9913
0,35	0,9341	0,9692	0,9840	0,9836	0,9963
0,40	0,9710	0,9710	0,9894	0,9896	0,9931
0,45	0,9928	0,9928	0,9928	0,9929	0,9946

Табела 3. Ефикасности за опсег Хафманових бинарних кодова

3. АНАЛИЗА

На претходним примерима види се да с проширењима извора ефикасност кода понекад опада. Разлог оваквог понашања лежи у начину генерисања Хафмановог кода када вероватноћа симбола варира. Од вероватноће симбола зависи Хафманов код, облик њему придруженог стабла али и расподела средње дужине кодне речи. Свако кодно стабло је оптимално за један скуп вероватноћа, а померањем од тих вредности, кодовање се погоршава. У многим случајевима кодовање по једном кодном стаблу побољшава се све док код не "прескочи" на друго кодно стабло чиме добија и другу формулу за дужину кодне речи. Тако се долази до дисконтинуитета у кривој која представља ефикасност у зависности од вероватноће симбола.



Слика 1. Ефикасност Хафмановог кода, параметар је проширење извора

Слика 1 приказује ефикасност бинарног Хафмановог кода за друго, треће и четврто проширење у зависности од вероватноће симбола. Свака крива је комбинација неколико конвексних функција које одговарају различитим кодним стаблима.

Друго проширење бинарног Хафмановог кода

Најједноставнији случај је пример другог проширења извора. На слици 1 се види да постоји дисконтинуитет негде око $\omega = 0,4$. Анализирајући понашање кодовања за овај случај, налазимо да за мање вредности вероватноће појављивања симбола, дужине кодних речи су из скупа {3,3,2,1}, док за веће вредности ω све кодне речи имају исту дужину и то 2.

	P(A)=0,35 P(B)=0,65		P(A)=0,45 P(B)=0,55	
	кодне речи	L_i	кодне речи	L_i
BB	0	$(1-\omega)^2$	11	2
BA	10	$2\ \omega\,(1-\omega)$	10	2
AB	111	$3\ \omega\,(1-\omega)$	01	2
AA	110	$3\ \omega^2$	00	2
	$L= -\ \omega^2+3\omega+1$		**L=2**	

Табела 4.
Дужине кодних речи Хафмановог кода са другим проширењем извора за вероватноће 0,35 и 0,45

Код за вероватноће симбола {0,35; 0,65} и {0,45; 0,55} дат је у табели 4, у којој се такође налазе и формуле за средњу дужину кодних речи у функцији ω. Комбинујући ове формуле добијамо да је $L= -\ \omega^2+3\omega+1$. а с обзиром да до преласка долази када је L=2, произилази да је ω=0,38196 тражена вероватноћа

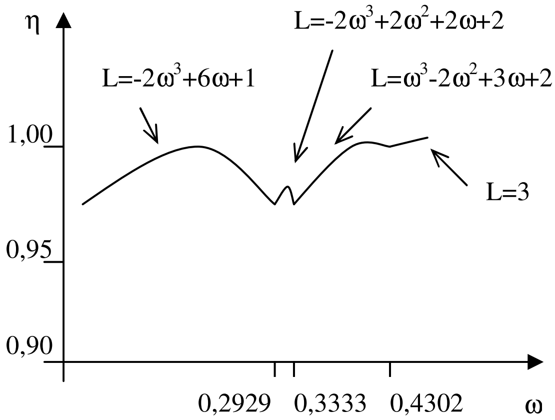
Треће проширење бинарног Хафмановог кода

Посматрајући поново слику 1, примећује се да за треће проширење извора, да постоји неколико различитих скупова кодних речи (а не само два, као за случај другог проширења). Бинарни Хафманов код је другачији за различите вредности вероватноће симбола извора. Из табеле 3 се види да ефикасност кода опада за вредност око $\omega\cong0,30$.

	P(A)=0,25 P(B)=0,75		P(A)=0,30 P(B)=0,60		P(A)=0,35 P(B)=0,65		P(A)=0,45 P(B)=0,55	
	кодне речи	L_i	кодне речи	L_i	кодне речи	L_i	кодне речи	L_i
BBB	0	$(1-\omega)^3$	11	$2(1-\omega)^3$	10	$2(1-\omega)^3$	111	3
BBA	110	$3\omega\,(1-\omega)^2$	00	$2\omega\,(1-\omega)^2$	111	$3\omega\,(1-\omega)^2$	110	3
BAB	101	$3\omega\,(1-\omega)^2$	101	$3\omega\,(1-\omega)^2$	110	$3\omega\,(1-\omega)^2$	101	3
ABB	100	$3\omega\,(1-\omega)^2$	100	$3\omega\,(1-\omega)^2$	011	$3\omega\,(1-\omega)^2$	100	3
BAA	11111	$5\omega^2\,(1-\omega)$	0111	$4\omega^2\,(1-\omega)$	001	$3\omega^2\,(1-\omega)$	011	3
ABA	11110	$5\omega^2\,(1-\omega)$	0110	$4\omega^2\,(1-\omega)$	000	$3\omega^2\,(1-\omega)$	010	3
AAB	11101	$5\omega^2\,(1-\omega)$	0101	$4\omega^2\,(1-\omega)$	0101	$4\omega^2\,(1-\omega)$	001	3
AAA	11100	$5\omega^3$	0100	$4\omega^3$	0100	$4\omega^3$	000	3
	$L=-2\omega^3+6\omega+1$		**$L=-2\omega^3-2\omega^2-2\omega-2$**		**$L=\omega^3-2\omega^2+3\omega+2$**		**L=3**	

Табела 5.
Дужине кодних речи Хафмановог кода са трећим проширењем извора за вероватноће 0,35 и 0,45

Испитујући понашање кода за ω=0,25, за ω=0,30 и за ω=0,35 добијамо резултате који су приказани у табели 5. За сваки тип стабла је израчуната средња дужина кодне речи. С обзиром да до преласка с једног кода на други долази за ону вероватноћу за коју су одговарајуће дужине једнаке, изједначавајући L, добијамо следеће вредности за граничне вероватноће: ω=0,2929, ω=0,3333 и ω=0,4302.



Слика 2. Треће проширење Хафмановог кода

На слици 2 се примећује се да за вероватноће бликске ω=0,30 постоје чак два дисконтиуитета, чиме је прецизније дато понашање кода од оног који показује слика 1, за случај трећег проширења.

4. ЗАКЉУЧАК

Посматрајући разна проширења извора Хафмановог кода, закључујемо да веће проширење извора не мора обавезно да побољша ефикасност кода. У свим случајевима је слично: кодовање са једним Хафмановим стаблом је оптимално само за опсег вероватноћа изворних симбола, док се погоршава све више и више како вероватноћа одступа од тог оптимума. Тек неко друго Хафманово стабло даје боље перформансе, и код се мења, користећи ново стабло.
Ефикасност кода може бити релативно слаба у области преласка са једног стабла на друго. У свим случајевима ефикасност је у оквиру познатих теоријских граница.

5. ЛИТЕРАТУРА

[1] N. Abramson, *"Information Theory and Coding"*, McGraw-Hill, 1963

[2] R. M. Capocelli, R. Giancarlo, I. J. Taneja, *"Bounds on the redunddancy of Huffman codes"*, IEEE Trans. Inform. Theory, Vol. IT-32, no 6, pp 854-857, Nov. 1986

[3] R. G. Gallager, *"Variations on a theme by Huffman"*, IEEE Trans. Inform. Theory, Vol. IT-24, no 6, pp 668-674, Nov. 1978

[4] D. A. Huffman, *"A method for the construction of minimum redundancy codes"*, Proc. IRE, Vol. 40, pp 1098-1101, 1952

[5] O. Johnsen, *"On the redundancy of binary Huffman codes"*, IEEE Trans. Inform. Theory, Vol. IT-26, no 2, pp 220-223, Mar. 1980