

Міністерство освіти і науки України
Національний технічний університет України
"Київський політехнічний інститут імені Ігоря Сікорського"
Фізико-технічний інститут

КРИПТОГРАФІЯ
КОМП'ЮТЕРНИЙ ПРАКТИКУМ №1

**Експериментальна оцінка ентропії на символ джерела
відкритого тексту**

Виконали:
Студенти ФБ-01
Сотнікова П.О.
Струкало В.В.

Мета роботи

Засвоєння понять ентропії на символ джерела та його надлишковості, вивчення та порівняння різних моделей джерела відкритого тексту для наближеного визначення ентропії, набуття практичних навичок щодо оцінки ентропії на символ джерела.

Частоти букв у тексті з пробілами

Letters Frequency	
' '	0.16763719112538508
'а'	0.06363574916431802
'б'	0.015144687684341613
'в'	0.036055171396735926
'г'	0.014219907255685916
'д'	0.026395670839614603
'е'	0.07060488628170676
'ж'	0.00903478731074261
'з'	0.013541939765353608
'и'	0.05865852559480894
'й'	0.00826055253326342
'к'	0.02672338926394442
'л'	0.03767123287671233
'м'	0.026970202202267812
'н'	0.05374070098970964
'о'	0.09874156125057351
'п'	0.02093813495444714
'р'	0.03296642524742741
'с'	0.04704295569246903
'т'	0.05520314445828144
'у'	0.023155354919053548
'ф'	0.0005345906796880121
'х'	0.00821446712984204
'ц'	0.0031348315527298947
'ч'	0.01374369142033165
'ш'	0.0060720079307858684
'щ'	0.0029801894212492627
'ъ'	0.0002888018614406502
'ы'	0.015204086648751393
'ь'	0.01749299501867995
'э'	0.0027098217211771647
'ю'	0.005929655240217605
'я'	0.017345521727731534
'ё'	7.168840532214721e-06

Частоти букв у тексті без пробілів

Frequency of Letters	
'а'	0.07645193716710612
'б'	0.01819481543729445
'в'	0.04331665352213879
'г'	0.017083784984232717
'д'	0.03171173742769997
'е'	0.08482465281836313
'ж'	0.010854386109535544
'з'	0.016269275394059
'и'	0.07047230482836855
'й'	0.009924221079065258
'к'	0.03210545807551808
'л'	0.04525818846669185
'м'	0.032401978938406094
'н'	0.06456403435704804
'о'	0.11862803118759682
'п'	0.025155058264503993
'р'	0.03960583641645309
'с'	0.05651736861726539
'т'	0.06632101274793635
'у'	0.027818824522398398
'ф'	0.0006422568067532934
'х'	0.009868854112965836
'ц'	0.003766184071785117
'ч'	0.016511659667872026
'ш'	0.007294905377854936
'щ'	0.003580397141095946
'ъ'	0.00034696632088971024
'ы'	0.018266177304711483
'ь'	0.0210160699543161
'э'	0.0032555776066460046
'ю'	0.007123882971458944
'я'	0.02083889566279795
'ё'	8.612639171021176e-06

Таблиці частот біграм наведено в таблиці

Ентропія

Entropy	4.3547600517319545				
H1 w/ spaces	4.3547600517319545	surplus:	0.14402125624927598		
H1 w/o spaces	4.448167298042461	surplus:	0.11819592347630059		
H2 w/ spaces	3.9436132206445955	surplus:	0.2248369484551599		
H2 w/o spaces	4.126240277770246	surplus:	0.18201469192597075		
H2 w/ spaces crossing	3.9438360468685425	surplus:	0.2247931493688765		
H2 w/o spaces crossing	4.127100273584213	surplus:	0.18184420647348276		

«найчастіші» букви, біграми

Текст з пробілами

' '	0.16763719112538508
'о'	0.09874156125057351
'е'	0.07060488628170676
'а'	0.06363574916431802
'и'	0.05865852559480894
'т'	0.05520314445828144
'н'	0.05374070098970964
'с'	0.04704295569246903
'л'	0.03767123287671233
'в'	0.036055171396735926

З перехрестям

'о '	0.024167185554171855
'и '	0.01963545421773612
'е '	0.01909779117782002
' с'	0.016759725044241988
' н'	0.01515492888510192
'то'	0.015027937995674116
'а '	0.014897874746018221
' п'	0.014852813462672872

'в'	0.014689978370583993
'и'	0.01298277020384086

Без перехрестя

'о'	0.024089352428393523
'и'	0.019550452251425574
'е'	0.019124418299796813
'с'	0.01676279740447008
'н'	0.01519999016844727
'то'	0.01519794192829521
'п'	0.014765763256210264
'а'	0.014671544209215442
'в'	0.014581421642524742
'и'	0.013020662646653995

Текст без пробілів

'о'	0.11862803118759682
'е'	0.08482465281836313
'а'	0.07645193716710612
'и'	0.07047230482836855
'т'	0.06632101274793635
'н'	0.06456403435704804
'с'	0.05651736861726539
'л'	0.04525818846669185
'в'	0.04331665352213879
'р'	0.03960583641645309

3 перехрестя

'то'	0.01862544739584551
'ст'	0.014121037109401434

'не'	0.011954343169377392
'на'	0.011405595016480901
'ос'	0.011069702088811074
'от'	0.01081255329070487
'ен'	0.010533257706158898
'по'	0.010415141511813464
'го'	0.01024288872839304
'он'	0.010144458566438514

Без перехрестя

'то'	0.018549141197893596
'ст'	0.014338796200600423
'не'	0.011786997391603918
'на'	0.011390816477188839
'ос'	0.010992174811752548
'от'	0.010869137260691964
'ен'	0.01043850583197992
'го'	0.01029086077070722
'по'	0.010224420493134505
'он'	0.010138294207392096

[illegible]

$$0,394387022 < R(30) < 0.54773628018$$

[illegible]

Висновки: на лабораторній роботі познайомилися з поняттями ентропії та надлишковості, вивчали та порівнювали різні моделі джерела відкритого тексту для наближеного визначення ентропії. Також визначали частоту символів та біграм в тексті та оцінювали ентропію.