Київський національний університет імені Тараса Шевченка факультет радіофізики, електроніки та комп'ютерних систем

Лабораторна робота № 1 Тема: « Дослідження кількості інформації при різних варіантах кодування »

> Роботу виконав студент 3 курсу КІ - МА Співак Богдан Олегович

<u>Мета:</u> Дослідити імовірнісні параметри української мови для оцінки кількості інформації текстів. Дослідити вплив різних методів кодування інформації на її кількість.

Теоретичні відомості

Відносна частома появи символу - імовірність появи певного символу в певному місці тексту - відношення числа появи символу в тексті до загальної кількості символів.

Середня ентропія нерівноймовірного алфавіту:

$$H = \sum_{i=1}^{m} p_i \log_2 \frac{1}{p_i} = -\sum_{i=1}^{m} p_i \log_2 p_i$$

де m - кількість символів алфавіту, p - імовірність появи символу Ентропія вимірюється в **БІТАХ** (як представлення кількості можливих варіантів).

Кількість інформації в тексті - середня ентропія вихідного алфавіту помножена на кількість символів тексту. (**HINT**: результат обрахунку для порівняння значення з розміром файлів треба перевести з бітів в байти)

1. Дослідження кількості інформації в тексті

1. Оберіть 3 текстових файла різного тематичного та лінгвістичного спрямування

https://uaf.ua/article/42189 http://svatovo.ws/stati_uk_glasses.html https://onlyart.org.ua/children/kazky/ukrayinska-narodna-kazka-tsap-ta-baran/

https://github.com/pilmenchik/CS

- 2. Переконайтесь, що тексти, які ви використовуєте є унікальними і не повторюються у ваших колег! Використовуйте наявні електронні засоби зв'язку та документообігу, щоб уникнути дублювання! Вдруге аналіз того самого тексту не зараховується!
- 3. Створіть програму (будь-якою зручною для вас мовою), яка в якості вхідних даних приймає текстовий файл, та аналізуючи його вміст:
 - а. обраховує частоти (імовірності) появи символів в тексті
 - обраховує середню ентропію алфавіту для даного тексту
 - с. виходячи з ентропії визначає кількість інформації та порівнює її з розмірами файлів
- d. виводить на екран значення частот, ентропії та кількості інформації Перший файл:

```
Введите название файла: text1
Размер файла:3277 bytes
Колличество информации:1084,3208192468628 bytes
Энтропия:4,8624251984164255
Вероятность появления символа:
/n - 0,0016816143497757848%
r - 0,0016816143497757848%
  - 0,13957399103139012%
 - 0,0005605381165919282%
 - 0,0005605381165919282%
 - 0,011771300448430493%
 - 0,0011210762331838565%
 - 0,006726457399103139%
 - 0,0005605381165919282%
 - 0,0005605381165919282%
 - 0,0005605381165919282%
 - 0,0016816143497757848%
 - 0,0016816143497757848%
 - 0,0033632286995515697%
 - 0,014013452914798207%
 - 0,0011210762331838565%
 - 0,0005605381165919282%
 - 0,004484304932735426%
 - 0,002802690582959641%
 - 0,0033632286995515697%
 - 0,0005605381165919282%
M - 0,0005605381165919282%
 - 0,004484304932735426%
 - 0.0011210762331838565%
 - 0,0005605381165919282%
 - 0,0011210762331838565%
 - 0,0072869955156950675%
 - 0,007847533632286996%
Ч - 0,0011210762331838565%
 - 0,0005605381165919282%
 - 0,06390134529147982%
 - 0,010089686098654708%
 - 0,04035874439461883%
 - 0,01569506726457399%
 - 0,029708520179372196%
e - 0,04988789237668161%
ж - 0,004484304932735426%
```

```
0,004484304932735426%
 - 0,020179372197309416%
 - 0,03195067264573991%
 - 0,007847533632286996%
 - 0,030269058295964126%
 - 0,029708520179372196%
 - 0,02298206278026906%
 - 0,05661434977578475%
 - 0,07286995515695067%
п - 0,016816143497757848%
 - 0,04540358744394619%
 - 0,030269058295964126%
 - 0,04091928251121076%
  - 0,02914798206278027%
 - 0,006726457399103139%
 - 0,005044843049327354%
ц - 0,010089686098654708%
ч - 0,0072869955156950675%
ш - 0,002802690582959641%
щ - 0,0016816143497757848%
ь - 0,013452914798206279%
ю - 0,003923766816143498%
я - 0,011210762331838564%
 - 0,003923766816143498%
 - 0,04988789237668161%
 - 0,008968609865470852%
 - 0,0016816143497757848%
  - 0,0005605381165919282%
```

Другий файл:

```
Введите название файла: text2
Размер файла:6301 bytes
Колличество информации:2102,179791408984 bytes
Энтропия:4,813233638028583
Вероятность появления символа:
/n - 0,0028620492272467086%
r - 0,0028620492272467086%
 - 0,1534058385804236%
 - 0,0008586147681740126%
 - 0,0005724098454493417%
 - 0,0031482541499713796%
 - 0,0031482541499713796%
 - 0,013737836290784202%
 - 0,0014310246136233543%
 - 0,01030337721808815%
 - 0,0005724098454493417%
 - 0,0005724098454493417%
 - 0,0005724098454493417%
 - 0,0005724098454493417%
 - 0,0014310246136233543%
 - 0,00028620492272467084%
 - 0,00028620492272467084%
 - 0,00028620492272467084%
 - 0,0011448196908986834%
 - 0,0011448196908986834%
 - 0,0005724098454493417%
 - 0,0005724098454493417%
 - 0,0005724098454493417%
 - 0,0005724098454493417%
 - 0,0014310246136233543%
 - 0,00028620492272467084%
 - 0,0008586147681740126%
 - 0,0008586147681740126%
 - 0,0017172295363480253%
И - 0,00028620492272467084%
 - 0,00028620492272467084%
 - 0,0011448196908986834%
 - 0,0005724098454493417%
 - 0,0011448196908986834%
 - 0,0008586147681740126%
 - 0,0011448196908986834%
П - 0,0031482541499713796%
```

```
- 0,0031482541499713796%
 - 0,0005724098454493417%
 - 0,0005724098454493417%
 - 0,00028620492272467084%
  - 0,0017172295363480253%
 - 0,0011448196908986834%
 - 0,061247853463079566%
 - 0,013451631368059531%
 - 0,03949627933600458%
 - 0,010875787063537493%
 - 0,02346880366342301%
 - 0,03663423010875787%
 - 0,007727532913566113%
 - 0,024613623354321695%
 - 0,04836863194046938%
 - 0,0045792787635947334%
 - 0,0363480251860332%
 - 0,03205495134516314%
 - 0,02718946765884373%
 - 0,04378935317687464%
 - 0,08070978820835718%
п - 0,027761877504293073%
 - 0,03692043503148254%
 - 0,028048082427017746%
 - 0,042072123640526614%
 - 0,034058385804235834%
 - 0,0014310246136233543%
 - 0,005437893531768746%
ц - 0,004865483686319404%
ч - 0,006868918145392101%
ш - 0,006296508299942759%
ш - 0,0031482541499713796%
 - 0,012020606754436176%
 - 0,009158557527189467%
 - 0,015455065827132226%
 - 0,007441327990841442%
 - 0,04264453348597596%
 - 0,006010303377218088%
  - 0,00028620492272467084%
```

Третій файл:

```
Введите название файла: text3
Размер файла:4958 bytes
Колличество информации:1691,9851971426353 bytes
Энтропия:4,801660722646712
Вероятность появления символа:
/n - 0,011706278822277403%
/r - 0,011706278822277403%
  - 0,17878680383114579%
 - 0,006385242993969493%
 - 0,028024122029088328%
  - 0,001418942887548776%
 - 0,011351543100390209%
 - 0,0046115643845335225%
 - 0,000709471443774388%
 - 0,000354735721887194%
 - 0,000709471443774388%
 - 0,000709471443774388%
 - 0,001064207165661582%
 - 0,002837885775097552%
 - 0,008513657325292657%
 - 0,000709471443774388%
 - 0,000354735721887194%
 - 0,000709471443774388%
3 - 0,000709471443774388%
 - 0,000354735721887194%
 - 0,000354735721887194%
 - 0,000354735721887194%
 - 0,0031926214969847464%
 - 0,0017736786094359701%
 - 0,001418942887548776%
 - 0,001418942887548776%
 - 0,000354735721887194%
 - 0,000709471443774388%
Ц - 0,006739978715856687%
Щ - 0,000354735721887194%
Я - 0,001064207165661582%
a - 0,08194395175594182%
 - 0,01596310748492373%
в - 0,03795672224192976%
 - 0,009932600212841433%
д - 0,02908832919474991%
```

- 0,02305782192266761%

```
0,011351543100390209%
 0,01312522170982618%
 0,056048244058176655%
 0,01596310748492373%
 0,024831500532103583%
 0,024831500532103583%
 0,01844625753813409%
 0,03228095069173466%
- 0,06278822277403334%
- 0,022703086200780417%
- 0,0297978006385243%
 0,02092940759134445%
 0,0418588151826889%
- 0,027314650585313942%
- 0,005675771550195104%
- 0,0031926214969847464%
- 0,0070947144377438804%
- 0,00886839304717985%
 0,0024831500532103584%
 0.0177367860943597%
- 0.002128414331323164%
- 0,012061014544164597%
- 0,001418942887548776%
- 0,03263568641362185%
- 0,0031926214969847464%
- 0,01312522170982618%
 0,000709471443774388%
```

- 4. Проведіть стиснення кожного вхідного файлу за допомогою 5 різних алгоритмів стиснення (zip, rar, gzip, bzip2, xz, або будь-які інші на ваш вибір, можна використовувати готові програмні засоби для стиснення).
- 5. Порівняйте результуючі обсяги архівів з обчисленою кількістю інформації та наведіть у звіті висновки щодо кореляції цих величин для обраних вами файлів (яка відмінність, що вийшло більше і чому)

Назва	Розмір файлу	Кількість інформації	7z	rar	bz2	XZ	zip
text1	3277	1084	1315	1311	1044	1248	1348
Text2	6301	2102	2342	2410	1921	2272	2423
text3	4958	1692	1837	1868	1456	1768	1885

Найбільш близькі до кількості інформації алгоритми- це bz2, xz, тому ці алгоритми можна вважати кращими за інші наведені у таблиці.

2. Дослідження способів кодування інформації на прикладі Base64

- 1. Ознайомтесь зі стандартом RFC4648
- 2. Для практичного засвоєння методу кодування, створіть програму, що кодує довільний файл в Base64 (шляхом реалізації алгоритму вручну, а не виклику бібліотечної функції)
- a. перевірте коректність роботи програми, порівнявши результат з існуючими програмними засобами (наприклад, openssl enc -base64)
 - 3. Закодуйте в Base64 обрані вами текстові файли

а.Обрахуйте кількість інформації в base64-закодованому варіанті файлу b.Порівняйте отримане значення з кількістю інформації вихідного файлу с.Зробіть висновки з отриманого результату

Введите название файла: text1.txt

Колличество информации:1084,3208192468628 bytes

Колличество информации в закодированом файле:2688,258721898126 bytes

Введите название файла: text1.txt

Колличество информации:1084,3208192468628 bytes

Колличество информации в закодированом файле:2688,258721898126 bytes

0J/RgNC10LFQuNC00LXQvdGCINC70JDQpCwg0YFQu9C10L0g0JLQuNC60L7QvdC60L7QvNGDINC70INCQOL3QtNGA0ZbQuSDQn9Cw0LLQtdC70LrQv1DQv9GA0L7QstGN0YLQsNCyINC00LXQu9C10LPQsNGC0ZbQ siDQmtC+0L3Qs9GA0LXRgdGDINCV0YHRgtC+0L3RgdGM0LrQvtGXINGE0YPRgtCx0L7Qu9GM0L3QvtGXINCw0YHQvtGG0ZbQsNGG0ZbR1yDQstGW0LQg0ZbQvNC10L3R1iDQo9CE0KTQkC4g0JfQsNGF0ZbQtCwg0LIg0Y/Q ıtC+0LzRgyDRgtCw0LrQvtC2INGD0LfRj9CyINGD0YfQsNGB0YLRjCDQstGW0YbQtS3Qv9GA0LXQt9C40LTQtdC90YIg0ZYg0LPQtdC90LXRgNCw0LvRjNC90LjQuSDRgdC10LrRgNC10YLQsNGAINCj0JDQpCDQrtGA0ZbQ ISDQ19Cw0L/R1tGB0L7RhtGM0LrQuNC5LCDQstGM0LTQsdGD0LLRgdGPINC90LAgwqvQkC4g0JvQtSDQmtC+0Log0JDRgNC10L3R1sK7INCyINC10LDQu9C70ZbQvdC90ZYuIA0K0KMg0YHQstC+0ZTQvNGDINCy0LjRgdG 0YPQv9GWINC/0LXRgNC10LQg0JrQvtC90LPRgNC10YHQvtC8INCV0KTQkCDQkNC90LTRgNGW0Lkg0J/QsNCy0LXQu9C60L4g0L/QtdGA0LXQtNCw0LIg0LLR1tGC0LDQvdC90Y8g0LLR1tC0INC/0YDQtdC30LjQtNC10L30 gtCwINCj0ITQpNCQINCQ0LvQtdC60YHQsNC90LTQtdGA0LAg0KfQtdGE0LXRgNGW0L3QsCDRgtCwINCy0ZbQtNC30L3QsNGH0LjQsiDRg9GB0L/R1tGF0Lgg0LXRgdGC0L7QvdGB0YzQutC+0Zcg0LDRgdC+0YbR1tCw0Ybf ltGXLiDCq9CU0YDRg9C20L3QsCDRhNGD0YLQsdC+0LvRjNC90LAg0YHR1tC84oCZ0Y8g0KPQhNCk0JAg0L/R1tC0INC60LXRgNGW0LLQvdC40YbRgtCy0L7QvCDQkNC70LXQutGB0LDQvdC00LXRgNCwINCn0LXRhNC10YDR ltC90LAg0LTQvtC70LDR1CDQutGA0LjQt9GDLCDQstC40LrQu9C40LrQsNC90YMg0L/QsNC90LTQtdC80ZbR1NGOLiDQ19Cw0YDQsNC3INC80Lgg0L3QsCDRiNC70Y/RhdGDINC00L4g0L/QvtCy0L3QvtGG0ZbQvdC90L7Q s9C+INCy0ZbQtNGA0L7QtNC20LXQvdC90Y8g0YTRg9GC0LHQvtC70YMg0ZYg0LIg0JXRgdGC0L7QvdGW0ZcsINGWINCyINCj0LrRgNCw0ZfQvdGWLCDR1iDQsiDQhNCy0YDQvtC/0ZYg0LIg0YbR1tC70L7QvNGDwrssIOKA LCDQt9Cy0LXRgNC90YPQstGB0Y8g0LLR1tC9INC00L4g0LTQtdC70LXQs9Cw0YLR1tCyINC60L7QvdCz0YDQtdGB0YMuIA0K0JfQsCDRgdC70L7QstCw0LzQuCDQkNC90LTRgNGW0Y8g0J/QsNCy0LXQu9C60LAsINCV0KTQ cDQv9GW0LQg0LrQtdGA0ZbQstC90LjRhtGC0LLQvtC8INC/0YDQtdC30LjQtNC10L3RgtCwINCQ0LnQstCw0YDQsCDQn9C+0YXQu9Cw0LrQscDQtNC+0YHRj9Cz0LvQsCDQs9Cw0YDQvdC40YUg0YDQtdC30YPQu9GW0YLQ sNGC0ZbQsiDigJQg0LfQvtC60YDQtdC80LAsINGDINGA0L7Qt9Cy0LjRgtC60YMg0LLQuNGJ0L7R1yDQu9GN0LPQuCwg0LzQsNGB0L7QstC+0LPQviwg0LTQuNGC0Y/Rh9C+LdG00L3QsNGG0YzQutC+0LPQviDRgtCwINC2 ₹ZÞQvdC+ÐYFQvtCz0L4gÐYTRg9GC0LHQvtC7ÐYMuIMKr0KHQsNC8ÐLUgÐYTRg9GC0LHQvtC7INC3ÐLDRgNCwÐLcg0ZQgÐL3QsNC5ÐLHR1tC7ÐYzR1CDQv9C+ÐL/Rg9C7ÐY/RgNC9ÐLjQvCDQstC4ÐLTQvtC8INGBÐL/QvtGA 0/LRgyDRgdC10/DQtdC0INC10/HRgtC+0L3RgdGM0LrQuNGFINC00ZbQstGH0LDRgtC+0LosINGWINCyINGG0YzQvtC80YMg0LLQtdC70LfQutCwINC30LDRgdC70YPQs9CwINGC0LDQutC+0LYg0LPQtdC90LXRgNCw0Lvi INC90L7Qs9C+INGB0LXQutGA0LXRgtCw0YDRjyDQldCk0JAg0YLQsCDQs9C+0LvQvtCy0Lgg0JrQvtC80ZbRgtC10YLRgyDQttGN0L3QvtGH0L7Qs9C+INGE0YPRgtCx0L7Qu9GDINCj0ITQpNCQINCQ0L3QvdC1INCg0LXQ ıS4g0]rRgNGW0Lwg0YLQvtCz0L4sINGP0Log0ZYg0LzQuCDQsiDQo9C60YDQsNGX0L3R1iwg0L3QsNGI0ZYg0LXRgdGC0L7QvdGB0YzQutGWINC60L7Qu9C10LPQuCDRgNC+0LHQu9GP0YLRjCDQsNC60YbQtdC90YIg0L3Q :CDRgNC+0LFQstC40YLQutGDINGB0L/QvtGA0YLQuNCy0L3QvtGXINGW0L3RnNGA0LDRgdGC0YDRg9C60YLRg9GA0LgsING30L4g0LTQvtC30LLQvtC70Y/R1CDQt9Cw0LvRg9GH0LDRgtC4INC00L4g0LFQsNC90Y/RgtGN INGEØYPRgtCx0L7Qu9C+0Lwg0LTQtdC00LDQu9GWINCx0ZbQu9GW0YjQtSDQtNGW0YLQu9Cw0YXR1tCywrssIOKA1CDQt9Cw0LFQvdCw0YFQuNCyINC/0YDQtdC30LjQtNC10L3RgiDQo9CQ0KQuIA0K0JfQsCDQv9GW0LTI zdGD0LzQutCw0LzQuCDQmtC+0L3Qs9GA0LXRgdGDINCV0YHRgtC+0L3RgdGM0LrQvtGXINGE0YPRgtCx0L7Qu9GM0L3QvtGXINCw0YHQvtGG0ZbQsNGG0ZbR1yDQkNC50LLQsNGA0LAg0J/QvtGF0LvQsNC60LAg0LHRg9C7 DL4g0L/QtdGA0LXQvtCx0YDQsNC90L4g0L3QsCDQvdC+0LLQuNC5INC/0YDQtdC30LjQtNC10L3RgtGB0YzQutC40Lkg0YLQtdGA0LzR1tC9LiDQktGW0YLQsNGO0YfQuCDRgdCy0L7Qs9C+INC10YHRgtC+0L3RgdGM0LrQ rtCz0L4g0LrQvtC70LXQs9GDINC3INC90L7QstC40Lwg0LzQsNC90LTQsNGC0L7QvCDQtNC+0LLR1tG40Lgg0LLR1tC0INC60L7QvdCz0YDQtdGB0YMsINCQ0L3QtNG40ZbQuSDQn9Cw0LLQtdC70LrQviDQv9GW0LTQutGA 8LXRgdC70LjQsiwg0YnQviDQo9CQ0KQg0YDQsNC00LAg0L/RgNC+0LTQvtCy0LbRg9Cy0LDRgtC4INCw0LrRgtC40LLQvdGDINGB0L/RltCy0L/RgNCw0YbRjiDQtyDQv9Cw0YDRgtC90LXRgNCw0LzQuCDRgtCwINC00L7Q :dGA0LfOVNC4INC00YDRg9C30Y/QVNC4INC3INCV0KTOKC4g0J/R1tC0INGH0LDRg5DQstGW0LFQuNGC0YMg0LTQV1DQotCw0LVQu9GW0L3QvdCwINC/0YDQtdC30LfQtNC10L3Rg1DQo9CQ0KQg0YLQsNC60L7Qt1DQt9GD ØYHRgtGAØZÞQstGBØY8g0Lcg0LLQtdGC0LXRgNCw0L3QvtC8INC10YHRgtC+0L3RgdGM0LrQvtCz0L4g0YTRg9GC0LHQvtC70YMg0JzQsNGA0YLR1tC90L7QvCDQoNC10LnQvNC+0LwsINGP0LrQuNC5INGUINGA0LXQutC+ 8YDQtNGB0LzQtdC90L7QvCDQt9Cx0ZbRgNC90L7R1yDQ1dGB0YLQvtC90ZbR1yDQt9CwINC60ZbQu9GM0LrR1tGB0YLRjiDQt9GW0LPRgNCw0L3QuNGFINC80LDRgtGH0ZbQsiAoMTU3KSwg0L/R1tC30L3R1tGI0LUg0L/I zNCw0YbRjtCy0LDQsiDQs9C+0LvQvtCy0L3QuNC8INGC0YDQtdC90LXRgNC+0Lwg0L3QsNGG0ZbQvtC90LDQu9GM0L3QvtGXINC60L7QvNCw0L3QtNC4LCDQsCDQvdC40L3R1iDQvtGH0L7Qu9G00ZQg0LfQsdGN0YDQvdGD

Введите название файла: text2.txt

Колличество информации:2102,179791408984 bytes

Колличество информации в закодированом файле:5140,612831886515 bytes

Введите название файла: text2.txt

Колличество информации:2102,179791408984 bytes

Колличество информации в закодированом файле:5140,612831886515 bytes

0J/QsNGA0YMg0YDQsNC30ZbQsiDRgdC/0L7RgdGC0LXRgNGW0LPQsNCyLCDRj9C6INGG0LUg0YDQvtCx0LvRj9GC0Ywg0L/RgNC+0YTQtdGB0ZbQvtC90LDQu9C4LiDQntGB0Ywg0YfQvtCz0L4g0LfQsNC/0LDQvCfRj9G DLDQsi4g033QsNGB0LDQvNC/0LXRgNC10LQg0LLQvtC90Lgg0LZQuNGO0YLRjCDRgdC60LvQviAo0LnQtNC10YLRjNGB0Y8g0L/RgNC+INC30LDQtNC90ZQpINC90LUg0LfQvdGw0LzQsNGO0YFQuCDQudC+0LPQviwg0L/v tGC0ZbQvCDQsdGA0LjQt9C60LDRjtGC0Ywg0L3QsCDQvdGM0L7Qs9C+INC3ĬMKr0L/RiNC40LrQsNC70LrQuMK7LiDQn9C+0YLRItC8INGA0L7Qt9C60L7Rh9GD0Y7RgtGMINC/0LvRitCy0LrRgyAo0L3Qt5DQt9C90Zb vMCu0977RP9C4INC3INC90LXR1yDQv9GA0L7Qt5C-097DQuNC5INGI0LDRgCkuINCS0ZDQtNC80ZDbgNGO0977RgtGMLCDQvtcX0VDRltc30LDRjtGC0Yug0ZYg0L3QvJ0V9C9C0L7m3D0977NgtGMLCDQvtcX0VDRltc30LDRjtGC0Yug0ZYg0L3QvNC20YDRltcY0DLDRjtGH01gQgZFR1yDQt9CwINC00L7QtV9C9C17RjtDRhNC10L3RgyDQsNC60YPRgNCw0YLQvdC+INCy0LjQs9Cw0L3Rj9G00YLRjCDQstC+0LTRgywg0YDQvtC30LPQu9Cw0LTQttGD0Y7RgtGML1DQodC60LDQttGDINCy0ZbQtNGA0LDQt9GDLCDRidG-INC3INC00LXRiNC10LLQvtGOINC/0LvRltCy0LrQvtGOINC80L7QttC90LAg0L3QsNCy0ZbRgtGMINC90LUg0L/RgNC+0LHRg9Cy0LDRgtC4INGG0LUg0LLQuNC60L7QvdCw0YLQuCwg0L7RgdC60ZbQu9GM0LrQuCDRgyDfgNC10LFRg9C70YZRgtCw0YLRliDQv9C70ZbQstC60LAg0LRltC00YXQvtC00LjRgtC40LZQt5DQs1DQvtC00VLQwtCw0YUDQoNCtCf0VDQvtC00L7QstC20YPRINC80L4SINGPOLog0LfQvdCP0YLQuCDRgdGC0LXQutC3 OLAGOLZÖYTC20L3QSCDQV9C+0LTQUNCYOLJRGTC40YHRJYDQS1DQUTC10YDR1TCYOL3QUNGGOYLQSTGWINC/OL4GOJXQUTGBOL/QU9GDOLDRGTCWOYbR1TGXINCYOLDRINC+0LPQV1DQSNCYOYLQV14gOJ/RITCOOL3RITCO OLDRINC8OL4gOZFRhSDQTNC+0LTQVTC80YMsINCWOLrRg9GAOLDRgTC9OL4uINCJINCYOLDQVdC9OZYgOLLR1TCOOLrRgNC4OLLQSNGUOLZQV1DQSOCWOYDRJ9GHOYMgOLLQVTCOOYMgOLIGOLTRg9GIOZYsINGFOLDQUSDQ O9C7OZTRgTGMOYHRJy4gOJnQTNC1OLZQV1DQSTC4OULZRITGAOY7R1NC8OL4gOYHRgTC1OLrQu9CwINGWINC3INC/OYDQUNC/OYPRgdC6OL7QVCDQs1A1INGBOLwgOLcgOYPRgdGWOYUgOLHQVTC6OZDQS1DQV9GAOLJRgdGC CONTROL OF TOWN CONTROL OF THE CONTR 017QmtCjINCv01rQntCHINCa01vQldCH0KLQmMCc01XQnNCeINCr91vQntCS01rQoykg91fQsNCy0LTQsNc90L3RjyDQstGB0LUg0YtQtC70L4g0YDQsNC30LggMi0zINC+0LHRgNC+0LHQnNGC01Lgg0LvQtdC30L7QvCwg
0L7RgdC+0LHQu9C40LLQviDRj9C60ZbRgdC90L4g0LrggOYHQutC70LAsINC60L7Qu9C4INC/0L7Rh9C90LXRgtC1INGG0Y4g0L/RgNC+0YbQtdC00YPRgNGDINCy0ZbQtNGA0LDQt9GDINCy0YHQtSDQt9GA0L7Q
t9GD0LzRltGU0YLQtS4g0JTQsNC70ZYg0LHQtdGA0LXQvNC+INC/0LvRltCy0LrRgywg0YLRg9GCINC90LUg0L7QsdGW0LnRgtC40YHRjyDQsdC10Lcg0LTQvtC/0L7QvNC+0LPQuCDQtNGA0YPQs9CwLCDR1iDQv9C+0YfQ
uNC90LDRINC80L4g0ZfRIyDQstGM0LTQutC70LXRjtCy0LDRgtC4INC/0YQQuCDRhtGM0L7QvNGDINGW0L3RiNC+0V4g0YDRg9C60L7RjiDQsNCx0L4g0LTRgNGD0LMg0L/RiNC40LrQsNGUINC90LAg0L/Qu9GM0LLQutGD ICjQsdC10Lcg0LfRg9C/0LjQvdC60LggKSDR1iDRgtCw0Log0YDQvtC3J9GU0LTQvdGD0ZTRgtC1INGX0Zcg0LTQv1DQutGW0L3RhtGPLCDQv9GA0L7Qt9C+0YDRgyDQv9C70ZbQstC69YMg0LzQvtC20L3QsCDQstC40LrQ uNC90YPRgtC4Lg0KDQrQ1NGA0YPQsyDCq9C/0YjQuNC60LDR1MK7INC90LAg0L/Qu9GW0LLQutGDLCDRidC+0LEg0LLQvtC90LAg0LHRg9C70LAg0L/QvtGB0YLR1tC50L3Qv1DQstC+0LvQvtC20L7Rjiwg0LAg0JLQuCDC q9C/0YjQuNC60LDR1NGC0LXCuyDQvdCwINGB0LrQu9C+INC60ZbQu9GM0LrQsCDRgNCw0LfR1tCyLCDQt9Cy0LXRgNGF0YMg0LLQvdC40Lcg0YLRgNC40LzgNG00YfQuNGB0Ywg0LfQsCDQutGA0LDR1yDQv9C70ZbQstC6 Lgg0L/RgNC40LrQu9Cw0LTQsNGU0YLQtSDR19GXINC00L4g0YHQutC70LAo0LrQu9C10LnQutC+0Y4g0YHRgtC+0YDQvtC90L7RjikuINCU0LDQu9GWINC80L7Qttd90Lwg8uLQuMC90YhQL0DQtyDQt9Cw0L/QsNGA0LrQu Cwg0ZYg0YPQt9GP0LLR1NC4INCyINGA0YPQutC4INGA0LXQt9C40L3QvtCy0YMg0LPRg9C80LrRgyDQstGW0LQg0YbQtdC90YLRgNGDINC20LXQvdC10LzQv1DQstCH0ETRgyPQtMGAULAGUXQbUC4INGA0LXQt9C40L3QvtCy0YMg0LPRg9C80LrRgyDQstGW0LQg0YbQtdC90YLRgNGDINC20LXQvdC10LzQv1DQstCH0ETRgyPQtMGA0LXQt9C40L3QvtCy0YMg0LPRg9C80LrRgyDQstGW0LQg0YbQtdC90YLRgNGDINC20LXQvdC10LzQv1DQstCH0ETRgyPQtMGA0LXQt9C40L3QvtCy0YMg0LPRg9C80LrRgyDQstGW0LQg0YbQtdC90YLRgNGDINC20LXQvdC10LzQv1DQstCH0ETRgyPQtMGA0LXQt9C40L3QvtCy0YMg0LPRg9C80LrRgyDQstGW0LQg0YbQtdC90YLRgNGDINC20LXQvdC10LzQv1DQstCH0ETRgyPQtMGA0LXQt9C40L3QvtCy0YMg0LPRg9C80LrRgyDQstGW0LQg0YbQtdC90YLRgNGDINC20LXQvdC10LzQv1DQstCH0ETRgyPQtMGA0LXQt9C40L3QvtCy0YMg0LPRg9C80LrRgyDQstGW0LQg0YbQtdC90YLRgNGDINC20LXQvdC10LzQv1DQstCH0ETRgyPQtMGA0LXQt9C40L3QvtCy0YMg0LPRg9C80LxQt9C40L3QvtCy0YMg0LPRg9C80LxQt9C40L3QvtCy0YMg0LPRg9C80LxQt9C40L3QvtCy0YMg0LPRg9C80LxQt9C40L3QvtCy0YMg0LPRg9C80LxQt9C40L3QvtCy0YMg0LPRg9C80LxQt9C40L3QvtCy0YMg0LPRg9C80LxQt9C40L3QvtCy0YMg0LPRg9C80LxQt9C40L3QvtCy0YMg0LPRg9C80LxQt9C40L3QvtCy0YMg0LYRgyDQxtQt9C40L3QvtCy0YMg0LYRgyDQxt9C40L3QvtCy0YMg0LYRgyDQxtQt9C40L3QvtCy0YMg0LYRgyDQxtQt9C40L3QvtCy0YMg0LYRgyDQxtQt9C40L3QvtCy0YMg0LYRgyDQxtQt9C40L3QvtCy0YMg0LYRgyDQxtQt9C40L3QvtCy0YMg0LYRgyDQxtQt9C40L3QvtCy0YMg0LYRgyDQxtQt9C40L3QvtCy0YMg0LYRgyDQxtQt9C40L3QvtCy0YMg0LYRgyDQxtQt9C40L3QvtCy0YMg0LYRgyDQxtQt9C40L3QvtCy0YMg0LYRgyDQxtQty0YMg0LYRgyDQxtQty0YMg0LYRgyDQxtQty0YMg0LYRgyDQxtQty0YMg0LYRgyDQxtQty0YMg0LYRgyDQxtQty0YMg0LYRgyDQxtQty0YMg0LYRgyDQxtQty0YMg0LYRgyDQxtQty0YMg0LYRgyDQxtQty0YMg0LYRgyDQxtQty0YMg0LYRgyDQxty0YMg0LYRgyDqxty0YMg0LYRgyDqxty0YMg0LYRgyDqxty0YMg0LYRgyDqxty0YMg0LYRgyDqxty0YMg0LYRgyDqxty0YMg0LYRgyDqxty0YMg0LYRgyDqxty0YMg0LYRgyDqxty0YMg0LYRgyDqxty0YMg0LYRgyDqxty0YMg0LYRgyDqxty0YMg0LYRgyDqxt vtGHOLDREtC60YMg0YHQuNC70YZQvdC+INC90LUg0L3QsNGC0LjRgdC60LDRjtGHOLggNGP0LrRidC+INC60YPQtNC4INC30LDQsdGA0LDQu9Cw0YHRjyDQs9GA0YQg9GNLCPQsjBR9C3H3N283H3VBR9GF0DLZQvdGA 0LLQsNGO0YHRjyDQv9GW0YHQu9GPINGC0LDRgNGW0LvQutC4IC0g0YLQvimg0LLQvtC90LAg0LZQvtC20LUg0L/QvtGA0LLQsNGC0Lgg0L/Qu9GW0LLQutGDKSDR1IDQv9GA0LjfgtGA0Ljfg ΘΥΖΦΥΑΘΜΘΥΤΟΣΕΝΟ ΣΤΟ ΔΕΙΕΡΕΝΟΘΟΥ/ΝΕΤΟ 4INC YOLF ON DICTOR HOLD LIDOK SCHOLL VONTON BELLOW TO A STANDAY OF THE PROPERTY OF THE 1tc20ljQcCDRgtc+0l3Rg9Cy0lDQxdC90Y,QvCwg0l3QsNC/07DQsNCy0LvRj9G00YLRj0CDRgtC10l3RgtC40ljQvdCw0lFQvdGw0YLQtdC/0LvQvtC20l4g0l/QvtCy0ZbRgtGA0Y8pl1DQryDRgNGw0Lbf gywg0lbQvtC00l3QvtC20l4g0YDQsNC30YMg0l3Qt5DQt9Cw0lTQuNGA0LDQu9C+0YHRjyDR11DQstC40LPQu9GP0LTQsNGUIMKr0lrQu9Cw0YHQvdC+wrsuINCf0l7RgtGA0ZbQsdC90l4g0l7QsdGA0ZbQt9Cw0YLQuCD(WOLQgOLrRg9GCOL7QvCDQtNC+INGBOLrQu9CwINCZ0YDQsNC00YPRgdGW0LIg0LTQviAzMCDR1iDQvdCiINC30LDQsdGD0LLQsNC50YLQt5DQstC40YHRg9Cy0LDRgtC4INC70LXQt9C+ICjRj9C60YrQviDQstC+0L3C iDQt9Cw0YLRg9C/0LjRgtGM0YHRjywg0LHRg9C00LUg0LfQsNC00LjRgCkuINCd0LAg0L3QtSDRgNGD0YXQvtC80LjRhSDRgdGC0LXQutC70LDRhSDQv9C+0YLRgNGW0LHQvdC+INC+0LHRgNGW0LfQsNGC0Lgg0L/Qu9GN VIDQI-SUMPITURGS-D-BIJNG CHMP IN SYMBOLING SEED BELLOWED INCOME SHEET OF THE WALLE SHEET OF THE SEED BELLOWED TO BE SHEET OF THE SEED BELLOWED THE SHEET OF THE SEED BELLOWED THE SHEET OF THE SEED BELLOWED THE SHEET OF THE SHEE ĿŢĢṣSDQvtcx@YjQuNCy@LrQscDQvdC1INC9@LDQvNC+@LrQvdGD@LvQsCDR1iDQv9C+INGC@ZbQusDQttc1INGB@YXQtdC8@ZYuINCh@LrQu9C+INC/@YDQuNGB@L/Rg9GJ@LXQvdC4@Lkg@L3QsCAx@YHQvCDR1iDQv9G LDRntGO0ZTQvNC+Lg0KDQrQn9C+INC/0YDQvtc/0YPRgdC60LDR1NC80L7RgdGC0ZYg0YHQstGW0YLQu9CwINC80L7QttGDINGB0LrQsNC30LDRgtC4LCDRidC+INC+0L/RgtC40LzQsNC70YzQvdC1INC00LvRjyDQktcQ

Введите название файла: text3.txt

Колличество информации:1691,9851971426353 bytes

Колличество информации в закодированом файле:4039,1580925923054 bytes

DIREGO, INSER TO SETTIMENT TO SCHOLL IT SCHIFFCE A BENTH LOSE POSCOPE, SOLIC PRESENTATION OF CHARGE LIGHT CONTROL DOWN THE C

- 4. Закодуйте в Base64 стиснені кращим з алгоритмів текстові файли
- Обрахуйте кількість інформації в base64-закодованому варіанті стисненого файлу b.Порівняйте отримане значення з кількістю інформації вихідного файлу та base64закодованого файлу

с.Зробіть висновки з отриманого результату

Введите название файла: text1.txt.bz2

Колличество информации:617,9651961688024 bytes

Колличество информации в закодированом файле:1394,25092325769 bytes

Введите название файла: text2.txt.bz2

Колличество информации:1165,1855861901656 bytes

Колличество информации в закодированом файле:2591,818911882706 bytes

Введите название файла: text3.txt.bz2

Колличество информации:888,4820161082945 bytes

Колличество информации в закодированом файле:1966,1155966983108 bytes

	text1.txt	text1.txt.bz2	text2.txt	text2.txt.bz2	text3.txt	text3.txt.bz2
Кількість	1084	617	2102	1165	1692	888
інформації						
К-ть	2688	1394	5140	2591	4039	1966
інформації						
Base64						

Вихідні коди розроблених програм завантажте в свій репозиторій на GitHub.

В Moodle завантажте звіт, що містить:

- результати проведеного аналізу кількості інформації обраних текстів (самі тексти в вигляді посилань або в додатках)
- посилання на програму в GitLab gl.vlabs.knu.ua
- приклад роботи створеної програми для підрахунку кількості інформації
- приклад роботи створеної програми для кодування в Base64

Висновок:

В данній лабораторній роботі було досліджене поняття середньої ентропії та кількості інформації. Було досліджено декілька алгоритмів стиснення, а саме зміна кількості інформації. Також було досліджено та реалізовано алгоритм кодування Ваse64, та його вплив на кількість інформації.