Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

Навчально-науковий Фізико-Технічний Інститут

Кафедра Математичних методів криптографічного захисту інформації

Комп'ютерний Практикум №2

«Статистичні критерії на відкритий текст»

Варіант 13

Виконали студенти:

Бондаренко Андрій
Яценко Артем
група: ФІ-22мн

Перевірила:

аспірант
Ядуха Дарія Вікторівна

Київ, 2022 р.

Зміст

0.1. Мета роботи

Засвоєння статистичних методів розрізнення змістовного тексту від випадкової послідовності, порівняння їх, визначення похибок першого та другого роду.

0.2. Завдання

У цьому комп'ютерному практикумі необхідно реалізувати критерії для розрізнення відкритих текстів від випадкової послідовності символів алфавіту, в ролі якої буде виступати результат перетворення відкритого тексту.

Для отримання випадкової послідовності нам потрібно реалізувати 4 методи спотворення тексту:

1. шляхом застосування шифру Віженера з *випадковим* ключем
2. шляхом застосування шифру афінної та афінної біграмної підстановки з випадковими ключами
3. y_i — рівномірно розподілена послідовність символів з $(Z_m)^l$
4. y_i Обчислюється відповідно до такого співвідношення:

$$y_i = (s_{i-1} + s_{i-2}) \mod m^l$$

Для нашого, 13-го варіанта, потрібно було реалізувати критерії $2.0-2.3, 4.0, 5.0$ для монограмного та біграмного випадків. Застосувати ці критерії для N різних текстів української мови довжини L та порахувати помилки першого та другого роду.

Далі, згенерувати випадковий текст довжини L = 10000, який точно не є зв'язним текстом українською мовою (наприклад, текст, який складається з величезної кількості літер а: «аааааааа . . . »). Застосувати один з варіантів спотворення (на вибір) до цього тексту, після чого застосувати один з реалізованих критеріїв (на вибір). Порівняти результати застосування критерію до різних текстів.

І, нарешті побудувати окремі таблиці для кожного зі способів спотворення, які містять ймовірності похибок першого та другого роду для кожного з критеріїв, що реалізовуються в роботі, для різних значень L та l.

0.3. Хід Роботи

У ході даної лабораторної роботи було реалізовано 6 критеріїв (2.0-2.3,4.0,5.0) та 4 методи спотворення тексту (віженера, афінний, рівномірно розподілена послідовність та фіббоначі).

Робота критеріїв була перевірена на N тескстах довжини L. для різних значень $L = [10, 100, 1000]$; $N = 10000$, а також для різних значень $L = [1, 2]$. Тобто для монограм та біграм.

При $l = 1$

$L = [100, 1000]$:

$h = 12$,
$k = 9$,
$k(4.0) = 0.001$,
$k(5.0) = 2$,
$j = 7$

$L = [10]$:

$h = 7$,
$k = 4$,
$k(4.0) = 0.001$,
$k(5.0) = 2$,
$j = 7$

При $l = 2$

$L = [100, 1000]$:

$h = 50$,
$k = 30$,
$k(4.0) = 0.001$,
$k(5.0) = 50$,
$j = 100$

$L = [10]$:

$h = 7$,
$k = 4$,
$k(4.0) = 0.001$,
$k(5.0) = 2$,
$j = 7$

0.4. Таблиці

Споворення за допомогою шифру Віженера (r = 1)					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
1000	2_0 (h=12,h=50)	0	0.45555	0.168	0.0101
	2_1 (h=12,h=50, k=h-3, k=h-20)	0	0.5	0	0.01665
	2_2 (h=12,h=50)	0.5	0	0.5	0
	2_3 (h=12,h=50)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.11645	0.38355
	5_0 (k = 2 j=7 ,k=50, j=100)	0.49765	5.00E-05	0	0.5

Споворення за допомогою шифру Віженера (r = 5)					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
1000	2_0 (h=12,h=50)	0	0.5	0.168	0
	2_1 (h=12,h=50, k=h-3, k=h-20)	0	0.5	0	0.05995
	2_2 (h=12,h=50)	0.5	0	0.5	0
	2_3 (h=12,h=50)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.11645	0
	5_0 (k = 2 j=7 ,k=50, j=100)	0.49765	0	0	0.5

Споворення за допомогою шифру Віженера (r = 10)					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
1000	2_0 (h=12,h=50)	0	0.5	0.168	0
	2_1 (h=12,h=50, k=h-3, k=h-20)	0	0.5	0	0.1265
	2_2 (h=12,h=50)	0.5	0	0.5	0
	2_3 (h=12,h=50)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.11645	0
	5_0 (k = 2 j=7 ,k=50, j=100)	0.49765	0	0	0.5

Споворення за допомогою шифру Віженера (r = 1)					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
100	2_0 (h=12,h=50)	0.0442	0.04115	0.5	0
	2_1 (h=12,h=50, k=h-3, k=h-20)	0	0.28265	0	0.4714
	2_2 (h=12,h=50)	0.5	0	0.5	0
	2_3 (h=12,h=50)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.23725	0.26275
	5_0 (k = 2 j=7 ,k=50, j=100)	0.018	0.18965	0	0.5

Споворення за допомогою шифру Віженера (r = 5)					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
100	2_0 (h=12,h=50)	0.0442	0.2116	0.5	0
	2_1 (h=12,h=50, k=h-3, k=h-20)	0	0.4797	0	0.49185
	2_2 (h=12,h=50)	0.5	0	0.5	0
	2_3 (h=12,h=50)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.23725	0.01665
	5_0 (k = 2 j=7 ,k=50, j=100)	0.018	0.0331	0	0.5

Споворення за допомогою шифру Віженера (r = 10)					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
100	2_0 (h=12,h=50)	0.0442	0.26285	0.5	0
	2_1 (h=12,h=50, k=h-3, k=h-20)	0	0.4911	0	0.4937
	2_2 (h=12,h=50)	0.5	0	0.5	0
	2_3 (h=12,h=50)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.23725	0.005
	5_0 (k = 2 j=7 ,k=50, j=100)	0.018	0.0205	0	0.5

Спотворення за допомогою шифру Віженера (r = 1)					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
10	2_0 (h=7)	0.49905	5.00E-05	0.5	0
	2_1 (h=7)	0.39075	0.00525	0.4999	0
	2_2 (h=7)	0.5	0	0.5	0
	2_3 (h=7)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.5	0
	5_0 (k = 2 j=7 ,k=50, j=100)	0	0.4996	0	0.5

Спотворення за допомогою шифру Віженера (r = 5)					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
10	2_0 (h=7)	0.49905	0	0.5	0
	2_1 (h=7)	0.39075	0.00515	0.4999	0
	2_2 (h=7)	0.5	0	0.5	0
	2_3 (h=7)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.5	0
	5_0 (k = 2 j=7 ,k=50, j=100)	0	0.49945	0	0.5

Спотворення за допомогою шифру Віженера (r = 10)					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
10	2_0 (h=7)	0.49905	0	0.5	0
	2_1 (h=7)	0.39075	0.0039	0.4999	0
	2_2 (h=7)	0.5	0	0.5	0
	2_3 (h=7)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.5	0
	5_0 (k = 2 j=7 ,k=50, j=100)	0	0.49955	0	0.5

Спотворення за допомогою шифру афінної підстановки					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
1000	2_0 (h=12,h=50)	0	0.2179	0.168	0.0002
	2_1 (h=12,h=50, k=h-3, k=h-20)	0	0.24605	0	0.0016
	2_2 (h=12,h=50)	0.5	0	0.5	0
	2_3 (h=12,h=50)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.11645	0.1851
	5_0 (k = 2 j=7 ,k=50, j=100)	0.49765	0.2563	0	0.5

Спотворення за допомогою шифру афінної підстановки					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
100	2_0 (h=12,h=50)	0.0442	0.01475	0.5	0
	2_1 (h=12,h=50, k=h-3, k=h-20)	0	0.14425	0.47965	0
	2_2 (h=12,h=50)	0.5	0	0.5	0
	2_3 (h=12,h=50)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.23725	0.1332
	5_0 (k = 2 j=7 ,k=50, j=100)	0.018	0.3576	0	0.5

Спотворення за допомогою шифру афінної підстановки					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
10	2_0 (h=7)	0.49905	0	0.5	0
	2_1 (h=7, k=h-3)	0.39075	0.0019	0.4999	0
	2_2 (h=7)	0.5	0	0.5	0
	2_3 (h=7)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.5	0
	5_0 (k = 2 j=7 ,k=50, j=100)	0	0.49985	0	0.5

Спотворення за допомогою рівномірно розподіленої послідовності					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
1000	2_0 (h=12,h=50)	0	0.5	0.168	0
	2_1 (h=12,h=50, k=h-3, k=h-20)	0	0.5	0	0.2901
	2_2 (h=12,h=50)	0.5	0	0.5	0
	2_3 (h=12,h=50)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.11645	0
	5_0 (k = 2 j=7 ,k=50, j=100)	0.49765	0	0	0.5

Спотворення за допомогою рівномірно розподіленої послідовності					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
100	2_0 (h=12,h=50)	0.0442	0.29625	0.5	0
	2_1 (h=12,h=50, k=h-3, k=h-20)	0	0.4959	0.47965	0
	2_2 (h=12,h=50)	0.5	0	0.5	0
	2_3 (h=12,h=50)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.23725	0.0001
	5_0 (k = 2 j=7 ,k=50, j=100)	0.018	0.0161	0	0.5

Спотворення за допомогою рівномірно розподіленої послідовності					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
10	2_0 (h=7)	0.49905	0	0.5	0
	2_1 (h=7, k=h-3)	0.39075	0.0039	0.4999	0
	2_2 (h=7)	0.5	0	0.5	0
	2_3 (h=7)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.5	0
	5_0 (k = 2, j=7)	0	0.49975	0	0.5

Спотворення за допомогою співвідношення Фібоначчі					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
1000	2_0 (h=12,h=50)	0	0.5	0.168	0
	2_1 (h=12,h=50, k=h-3, k=h-20)	0	0.5	0	0.0046
	2_2 (h=12,h=50)	0.5	0	0.5	0
	2_3 (h=12,h=50)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.11645	0.0018
	5_0 (k = 2 j=7 ,k=50, j=100)	0.49765	0	0	0.5

Спотворення за допомогою співвідношення Фібоначчі					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
100	2_0 (h=12,h=50)	0.0442	0.233	0.5	0
	2_1 (h=12,h=50, k=h-3, k=h-20)	0	0.4871	0.47965	0
	2_2 (h=12,h=50)	0.5	0	0.5	0
	2_3 (h=12,h=50)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.23725	0.04765
	5_0 (k = 2 j=7 ,k=50, j=100)	0.018	0.04565	0	0.5

Спотворення за допомогою співвідношення Фібоначчі					
L	Номер критерію	**FP** (l=1)	**FN** (l=1)	**FP** (l=2)	**FN** (l=2)
10	2_0 (h=7)	0.49905	0	0.5	0
	2_1 (h=7, k=h-3)	0.39075	0.00055	0.4999	0
	2_2 (h=7)	0.5	0	0.5	0
	2_3 (h=7)	0	0.5	0	0.5
	4_0 (k=0.001)	0.5	0	0.5	0
	5_0 (k = 2, j=7)	0	0.5	0	0.5

0.5. Висновок

Порівнюючи критерії між собою ми дійшли висновку:

Критерій 2.0: При l = 1: працює краще, хоча не розрізняє випадковий текст спотворений Афінною підстановкою та рівномірно розподілений текст. При l = 2: критерій добре працює на випадковому тексті та погано працює із змістовним текстом.

Критерій 2.1: При обох значеннях l критерій працює добре майже для всіх спотворень. При цьому потрібно було знайти компроміс у параметрах спотвореного тексту шифром Віженера та Фібоначчі. Але для Афінної заміни виникають складнощі.

Критерій 2.2: Майже завжди показує однакове значення помилки першого роду, і низькі показники помилки другого роду.

Критерій 2.3: Аналогічно критерію 2.2

Критерій 4.0: Працює майже ідеально для всіх випадків, хоча для Афінної підстваноки при l = 1 практично не відрізняється.

Крієтрій 5.0: Погано працює для змістовного тексту при l = 1. З рештою працює добре.