Тема 3. КЛАСИЧНІ ШИФРИ ЗАМІНИ

1. Загальна характеристика шифрів заміни

У загальному випадку шифр заміни здійснює перетворення, при якому літери або якісь інші фрагменти відкритого тексту замінюються відповідними фрагментами шифрованого тексту.

Найпростіший випадок шифрування заміною полягає в тому, що знаки відкритого тексту, записані в одному (первинному) алфавіті, заміняють знаками, що взято із іншого (вторинного) алфавіту, у відповідності з наперед установленим правилом. При цьому один і той же знак на протязі тексту замінюється однаково.

Якщо використовується один і той же вторинний алфавіт, то шифр заміни називають моноалфавітним. Якщо вторинних алфавітів декілька, то шифр називають багатоалфавітним.

Одним із перших моноалфавітних шифрів заміни вважається полібіанський квадрат. У II столітті до н.е. грецький письменник та історик Полібій винайшов з метою шифрування квадратну таблицю розміром 5x5, заповнену літерами грецького алфавіту у випадковому порядку.

При шифруванні чергову літеру відкритого тексту знаходили у цьому квадраті, а у криптограму записували літеру, розташовану рядком нижче в тому ж стовпчику. Якщо літера знаходилась у нижньому рядку таблиці, то для криптограми брали саму верхню літеру з того ж стовпчика.

Таким чином, основним для будь-якого шифру заміни є поняття алфавіту, який являє собою фіксовану послідовність всіх використовуваних знаків. При цьому фіксується як порядок слідування знаків, так і їх загальна кількість. Знаки алфавіту нумеруються по порядку, починаючи з нуля, тобто $0 \le j < m$. Таким чином, m являє собою загальну кількість знаків в алфавіті і називається його обсягом.

Знаки відкритого тексту теж доцільно нумерувати, починаючи з нуля, тобто $0 \le i < n$. Тут n являє собою загальну кількість знаків у повідомленні. [8]

2. Моноалфавітна звичайна заміна Цезаря (шифр Цезаря)

Як повідомляє історик Гай Светоній, римський імператор Гай Юлій Цезар користувався у своєму військовому та особистому листуванні шифром, суть якого полягала у заміні кожної літери повідомлення на одну із інших літер 26-значного латинського алфавіту.

Щоб зрозуміти зашифроване повідомлення Цезаря, треба було кожну літеру в ньому замінити третьою, що йде після неї в алфавіті. При досягненні кінця алфавіту виконувався циклічний перехід до його початку.

Такий метод шифрування можна відобразити шифрувальною таблицею, в якій указано замінюючи знаки для кожного знаку криптограми. Використання таблиці очевидне: при шифруванні для кожного знаку відкритого тексту шукаємо відповідний знак криптограми, при дешифруванні – навпаки.

Номер (код)	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Знак відкритого тексту	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Знак кріптограми	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W

Отже, шифр Цезаря пов’язаний із використанням первинного алфавіту (для відкритого тексту) і вторинного алфавіту (для криптограми), циклічно зміщеного відносно первинного на три знаки вперед.

Зокрема, Цезар використовував свій шифр у листуванні з Цицероном (близько 50 р. до н.е.). А його відоме послання VENI VIDI VICI – «Пришёл, увидел, победил» своєму другові Амінтію у зашифрованому вигляді мало такий вигляд: SBKF SFAF SFZF.

Загальний випадок шифру Цезаря для первинного алфавіту деякого обсягу m ($0 \le j < m$) полягає в тому, що вторинний алфавіт циклічно зміщується відносно первинного на K знаків вперед ($0 \le K < m$). Значення K є ключем цього шифру.

Наприклад, для алфавіту “АБВГДЕЖЗИК” обсягом $m=10$ шифрувальна таблиця для $K=6$ має наступний вигляд:

Номер (код)	0	1	2	3	4	5	6	7	8	9
Знак відкритого	А	Б	В	Г	Д	Е	Ж	З	И	К

тексту										
Знак криптограми	Е	Ж	З	И	К	А	Б	В	Г	Д

Таким чином, відкритому тексту “ЖАБА” відповідає криптограма “БЕЖЕ” і навпаки.

Недоліки моноалфавітної звичайної заміни Цезаря:

- ✓ вона не маскує частот появи різних знаків відкритого тексту і тому її легко зламати на підставі аналізу частот появи знаків у криптограмі;
- ✓ у вторинному алфавіті зберігається той же самий алфавітний порядок знаків, що і в первинному;
- ✓ мала кількість можливих ключів (рівна обсягу алфавіту). [8]

3. Шифр Цезаря з ключовим словом

Шифр Цезаря з ключовим словом теж є моноалфавітним. Він має ту особливість, що порядок знаків у вторинному алфавіті у порівнянні з первинним є дещо іншим завдяки використанню ключового слова. Крім того, практично необмеженою стає кількість ключів.

Для шифрування як ключ вибирається деяке число $0 \le K < m$, а також ключове слово. Всі знаки ключового слова мають бути різні.

Ключове слово записують під знаками алфавіту, починаючи зі знаку, числовий код якого співпадає з числом *K*. Знаки алфавіту, що залишились, записують за ключовим словом в алфавітному порядку.

Припустимо, що для алфавіту “АБВГДЕЖЗИК” обсягом $m=10$ ключем вибрано число K = 3 і ключове слово БЕДА. Отримуємо шифрувальну таблицю:

Номер (код)	0	1	2	3	4	5	6	7	8	9
Знак відкритого тексту	А	Б	В	Г	Д	Е	Ж	З	И	К
Ключове слово				Б	Е	Д	А			
Знак криптограми	З	И	К					В	Г	Ж

В цій таблиці первинний алфавіт (для відкритого тексту) вказано в другому рядку, вторинний (для криптограми) – в третьому та четвертому. У відповідності з цією таблицею відкритий текст ЖАЖДА шифрується як АЗАЕЗ.

Вимога про відмінність всіх знаків ключового слова не є обов'язковою. В цьому випадку просто записують ключове слово без повторення однакових знаків. Наприклад, ключове слово ЖАЖДА записують як ЖАД.

В шифрі Цезаря з ключовим словом недолік моноалфавітної звичайної заміни Цезаря, пов'язаний з відсутністю маскування частот появи різних знаків відкритого тексту, зберігається. [8]

4. Шифр Гронсфельда

Шифр Гронсфельда являє собою модифікацію моноалфавітної звичайної заміни Цезаря більш складним ключем, який являє собою послідовність чисел. Кожне з чисел ключа має бути меншим обсягу алфавіту *m*. Цей ключ записують під відкритим текстом. Якщо ключ коротший відкритого тексту, то його повторюють циклічно. Кріптограму отримують як і в шифрі Цезаря, але здійснюють відлік такої кількості літер, яка указана відповідним числом ключа.

Таким чином, шифр Гронсфельда є багатоалфавітним, оскільки використовує декілька вторинних алфавітів (по кількості різних чисел у складі ключа).

Наприклад, для алфавіту "АБВГДЕЖЗИК" обсягом *m=10* шифрувальна таблиця для чисел ключа в межах від *K=1* до *K=6* має наступний вигляд:

Номер (код)	0	1	2	3	4	5	6	7	8	9
Знак відкритого тексту	А	Б	В	Г	Д	Е	Ж	З	И	К
Знак криптограми *K=1*	К	А	Б	В	Г	Д	Е	Ж	З	И
Знак криптограми *K=2*	И	К	А	Б	В	Г	Д	Е	Ж	З
Знак криптограми *K=3*	З	И	К	А	Б	В	Г	Д	Е	Ж
Знак криптограми *K=4*	Ж	З	И	К	А	Б	В	Г	Д	Е
Знак криптограми *K=5*	Е	Ж	З	И	К	А	Б	В	Г	Д
Знак криптограми *K=6*	Д	Е	Ж	З	И	К	А	Б	В	Г

Для прикладу зашифруємо відкритий текст "ЗАДВИЖКА", використовуючи ключ (5, 1, 3). Розмістимо ключ під відкритим текстом:

З	А	Д	В	И	Ж	К	А
5	1	3	5	1	3	5	1

Щоб зашифрувати першу літеру відкритого тексту З, треба використати перше число ключа 5. Це означає, що відповідний знак криптограми треба взяти із того рядка шифрувальної таблиці, для якого К=5. Отримуємо першу літеру криптограми В. Щоб зашифрувати другу літеру відкритого тексту А, треба використати друге число ключа 1. Це означає, що відповідний знак криптограми треба взяти із того рядка шифрувальної таблиці, для якого К=1. Отримуємо другу літеру криптограми К. Остаточно отримуємо криптограму ВКБЗЗГДК. [8]

5. Гомофонічна заміна

Гомофонічна заміна одному знакові відкритого тексту ставить у відповідність декілька різних символів криптограми. Цей метод застосовується для спотворення статистичних властивостей криптограм. Наприклад, для алфавіту “АБВГДЕЖЗИК” обсягом m=10 шифрувальна таблиця для кожного знаку алфавіту може містити набори по три різні символи у вигляді довільних двозначних чисел, вибраних випадковим способом:

Номер (код)	0	1	2	3	4	5	6	7	8	9
Знак відкритого тексту	А	Б	В	Г	Д	Е	Ж	З	И	К
Символи криптограми	17 31 48	23 44 63	14 89 42	55 52 11	37 88 25	97 51 15	47 67 33	76 19 59	27 64 73	77 38 45

Для прикладу зашифруємо відкритий текст “ЖАДАЖА”. Для шифрування першої літери відкритого тексту Ж застосовуємо відповідний перший символ криптограми 47. Аналогічно для літер відкритого тексту А і Д, які теж зустрічаються перший раз, використовуються відповідні перші символи криптограми 17 і 37. Четверта і п’ята літери криптограми А і Ж зустрічаються вдруге, тому для них відповідними символами криптограми будуть 31 і 67.

Остання шоста літера криптограми А зустрічається втретє, тому вона буде шифруватись символом 48. Остаточно криптограма має вигляд 471737316748. Бачимо, що при використанні шифру гомофонічної заміни кожний знак відкритого тексту замінюється відповідними символами криптограми по черзі. Після того, як весь набір символів для даного знаку вичерпався, здійснюється його повторне використання з початку. [8]

6. Шифруюча таблиця Трисемуса

Багато хто з істориків вважають Іоганна Трисемуса, аббата з Німеччини, одним із засновників сучасної криптології. У 1508 році Трисемус написав свій твір "Поліграфія" – перший друкований твір з криптології. В ньому він вперше описав свій шифр, відомий під назвою шифруюча таблиця Трисемуса. В цьому шифрі використовується прямокутна таблиця певного розміру (по можливості, якнайближча до квадратної) для запису знаків алфавіту і ключове слово, записане без повторення однакових знаків. Для заповнення шифрувальної таблиці використовується підхід, запозичений із шифру Цезаря з ключовим словом. Спочатку в таблицю по рядкам вписується ключове слово. Далі таблиця доповнюється рештою знаків в алфавітному порядку. Наприклад, для алфавіту "АБВГДЕЖЗИКЛМ" обсягом $m = 12$ шифрувальна таблиця може мати розмір 3х4. Виберемо ключове слово ЖАД. За таких умов шифрувальна таблиця має такий вигляд:

Ж	А	Д	Б
В	Г	Е	З
И	К	Л	М

Спосіб шифрування запозичений із полібіанського квадрата. При шифруванні чергову літеру відкритого тексту знаходили у шифрувальній таблиці, а у криптограму записували літеру, розташовану рядком нижче в тому ж стовпчику. Якщо літера знаходилась у нижньому рядку таблиці, то для криптограми брали саму верхню літеру з того ж стовпчика. Наприклад, ГЛАЗ шифрується як КДГМ. [8]

7. Біграмний шифр Плейфера

Даний шифр називається біграмним тому, що шифруються одночасно не один, а два сусідні знаки відкритого тексту. Шифрувальна таблиця Плейфера являє собою прямокутну матрицю (по можливості, якнайближчу до квадратної). Її розміри мають бути достатніми для розміщення всіх знаків алфавіту відкритого тексту. Матриця заповнюється знаками алфавіту випадковим способом. Наприклад, для алфавіту "АБВГДЕЖЗИКЛМ" обсягом $m=12$ шифрувальна таблиця може мати наступний вигляд:

Ж	А	Д	Б
В	Г	Е	З
И	К	Л	М

Процес шифрування складається з таких кроків.

1. Відкритий текст розбивається на пари знаків (біграми). В тексті повинна бути парна кількість знаків і не повинно бути біграм з однаковими знаками. Якщо ці умови не виконуються, то текст модифікують з утворенням незначних орфографічних помилок. *Наприклад*, можна замінити один із цих знаків іншим, вставити між ними дефіс або виключити один із них взагалі. Після цього кожна дана біграма відкритого тексту за допомогою шифрувальної таблиці перетворюється в результуючу біграму криптограми.

2. Якщо обидва знаки даної біграми відкритого тексту знаходяться в різних рядках та стовпчиках матриці, то вони вважаються протилежними кінцями діагоналі відповідного прямокутника. Результуючу біграму криптограми знаходять на кінцях другої діагоналі цього ж прямокутника. Знак, який знаходиться на лівому кінці першої діагоналі, замінюється знаком який знаходиться на лівому кінці другої діагоналі. Знак, який знаходиться на правому кінці першої діагоналі, замінюється знаком, який знаходиться на правому кінці другої діагоналі. Наприклад, даним біграмам відкритого тексту АЛ, МА, ДИ і КЕ відповідають результуючі криптограми КД, БК, ЛЖ і ГЛ.

3. Якщо знаки даної біграми знаходяться в одному й тому ж рядку, то кожний із знаків замінюється тим, що стоїть справа від нього (за останнім

знаком у рядку йде перший). Наприклад, даним біграмам ЖБ і ЛИ відповідають результуючі біграми АЖ і МК.

4. Якщо знаки даної біграми знаходяться в одному й тому ж стовпчику, то кожний із знаків замінюється тим, що стоїть нижче його (за останнім нижнім знаком йде самий верхній). Наприклад, даним біграмам ЖВ, ЛД і МЗ відповідають результуючі біграми ВИ, ДЕ і БМ. Наприклад, відкритий текст БАЗА перетворюється на криптограму ЖДБГ. [8]

8. Біграмний двотабличний шифр

Цей шифр винайдений у 1854 році англійцем Чарльзом Уитстоном. Такий метод використовує дві прямокутні таблиці однакового розміру *(по можливості, якнайближчі до квадрату)*, в кожній з яких випадковим способом розміщено один і той же алфавіт. Відкритий текст розбивають на пари знаків – біграми. Перший знак біграми відкритого тексту фіксується у першій таблиці, другий знак біграми – у другій. Між зафіксованими знаками вибудовується уявний прямокутник. Одна діагональ цього прямокутника з'єднує знаки біграми відкритого тексту, друга діагональ дає результуючу біграму до криптограми. Перший знак результуючої біграми теж прочитується із першої таблиці, другий знак біграми – із другої таблиці. Якщо знаки відкритого тексту потрапили в один і той же рядок, то і біграма криптограми береться з того ж рядка. Перший знак біграми криптограми береться із першої таблиці у стовпчику, номер якого такий же, як і номер стовпчика другого знаку біграми відкритого тексту. Другий знак біграми криптограми береться із другої таблиці у стовпчику, номер якого такий же, як і номер стовпчика першого знаку біграми відкритого тексту. При використанні алфавіту, який складається із десяти цифр, крапки та пропуску, шифрувальна таблиця може бути такою:

Таблиця 1			Таблиця 2		
2	7	.	0	2	7
6	0	3	_	.	4
1	4	9	6	8	5
_	5	8	3	1	9

У відповідності з цією таблицею біграму «78» буде зашифровано як «42», біграму «42» – як «78», біграму «59» – як «81» і т.д. А для повідомлення «2.718_3.14» одержимо криптограму «6252330465». Перевагою біграмного двотабличного шифру у порівнянні з біграмним шифром Плейфейра є можливість використання біграм з однаковими знаками. [8]

9. Координатні заміни

В координатних замінах знаки алфавіту використовуються для позначень координат шифрувальної таблиці. Якщо алфавіт має N знаків і мова йде про двохкоординатну заміну, то шифрувальна таблиця має форму квадрата розміром $N*N$. В окремих комірках таблиці випадковим способом розміщують всі N можливих пар знаків, а вертикалі та горизонаталі таблиці позначають знаками, розташованими в алфавітному порядку. Відкритий текст розбивають на пари знаків – біграми. Перший знак біграми відкритого тексту використовується як індекс рядка, другий знак біграми – як індекс стовпчика. На їх перетині знаходиться результуюча біграма до криптограми. *Наприклад,* при використанні алфавіту «0, 1, 2» шифрувальна таблиця може мати такий вигляд:

	0	1	2
0	10	21	01
1	00	20	11
2	12	02	22

Для такої шифрувальної таблиці повідомлення «1020110221» перетворюється на криптограму «0012200102».

Суть ускладненого координатного методу, полягає в тому, що таблиці з числами ставиться у відповідність таблиця з алфавітом, тоді кожній літері у відповідність ставляють числа, приклад:

	0	1	2		0	1	2
0	10	21	01	0	А	Б	В
1	00	20	11	1	Г	Д	Е
2	12	02	22	2	Ж	Й	К

Тоді, наприклад, слово ДАВАЙ перетвориться на «2010011002». [8]