Wrocław, 14 listopada 2005

Poufność Informacji i Techniki Kodowania
Politechnika Wrocławska
Wydział Informatyki i Zarządzania
IV rok studiów

Szyfry homofoniczne, poligramowe oraz polialfabetyczne

Autor dokumentu: **STAWARZ Paweł**
Indeks: **125939**
Termin seminarium: **17 października 2005**

Prowadząca: **dr inż. Teresa Mendyk-Krajewska**

Kryptologia

Nauka zajmująca się (jawnie) tajnym pismem (kryptografia), jego nieautoryzowanym deszyfrowaniem (kryptoanaliza) oraz zasadami, które mają z kolei utrudnić nieautoryzowane deszyfrowanie (bezpieczeństwo szyfrowania).

Pojęcia podstawowe

Słownik - zbiór znaków, zbiór słów

Słownik (zbiór znaków) tekstu jawnego – zbiór znaków V, używany do formułowania tekstu jawnego.

Słownik (zbiór znaków) tekstu zaszyfrowanego – zbiór znaków W, używany do zapisywania kryptogramu, zwanego też tekstem kodowanym lub tekstem zaszyfrowanym.

Poszczególne znaki ze zbioru W mogą być również logogramami, symbolami specjalnymi reprezentującymi słowo lub wyrażenie (np. &, %, $, £, ©) a także symbolami stenograficznymi.

Zbiory V i W mogą być rozłączne, mieć niepustą część wspólną lub być identyczne. We wszystkich praktycznych zastosowaniach są one niepustymi zbiorami skończonymi, chociaż teoretycznie można przyjąć, że są one zbiorami przeliczalnymi.

Przestrzeń tekstu jawnego – zbiór słów V^* zbudowany z alfabetu V.

Przestrzeń tekstu zaszyfrowanego – zbiór słów W^* zbudowany z alfabetu W.

e **- słowo puste (o długości 0).**

Z^n **- zbiór wszystkich słów o długości n, a więc** $Z^n \subseteq Z^*$.

$Z^{(n)} = \{e\} \cup Z \cup Z^2 \cup \mathbf{L} \cup Z^n$ **- zbiór wszystkich słów o dł. nie większej niż n.**

Szyfrowanie i deszyfrowanie

Szyfrowanie definiowane jest jako różnowartościowa relacja $X : V^* - - \rightarrow W^*$ umożliwiająca jednoznaczną deszyfrację, a więc $(x \rightarrow z) \wedge (y \rightarrow z) \Rightarrow (x = y)$.

Deszyfrowanie definiowane jest jako relacja odwrotna $X^{-1} : V^* \leftarrow - - W^*$, w której $x \leftarrow y \Leftrightarrow x \rightarrow y$.

Włókno $x \in V^*$ definiowane jest jako $H_x = \left\{ y \in W^* : x \xrightarrow{X} y \right\}$. Z reguły wymaga się, aby szyfrowanie X było totalne, tzn. aby H_x było niepuste dla każdego $x \in V^*$.

Jeśli H_x zawiera co najwyżej jeden element dla każdego $x \in V^*$, to szyfrowanie jest funkcyjne $X : V^* \to W^*$, albo surjektywne $X : V^* \leftrightarrow W^*$ (funkcja wzajemnie jednoznaczna). Szyfrowanie jest deterministyczne.

Wariantami lub homofonami x nazywane są elementy należące do H_x (zakładając, że istnieje więcej niż jeden), którymi są różne słowa tekstu zaszyfrowanego przypisane do tego samego słowa tekstu jawnego w relacji szyfrowania $X : V^* \dashrightarrow W^*$.

Szyfrowania $X : V^* \dashrightarrow W^*$ dokonuje się przy pomocy niedeterministycznego operatora Hilberta h, gdzie $X(x) = hH_x$.

Tekst zerowy/sztuczny jest elementem różnym od słowa pustego e, będącym homofonem dla $e \in V^*$.

Jeśli zbiór wszystkich par w relacji $X : V^* \dashrightarrow W^*$ jest skończony, to szyfrowanie jest skończone i wówczas $X : V^{(n)} \dashrightarrow W^{(m)}$.

Systemy kryptograficzne (kryptosystemy)

System kryptograficzny (kryptosystem) stanowi system szyfrowania wraz z odpowiadającym mu systemem deszyfrowania.

System szyfrowania M jest niepustym, z reguły skończonym, zbiorem różnowartościowych relacji $\{c_0, c_1, c_2, \ldots, c_{q-1}\}$ postaci $c_i : V^{(n_i)} \dashrightarrow W^{(m_i)}$. Mocą systemu szyfrowania (liczbą elem.) jest $q = |M|$.

Każdą z relacji tworzących $c_i : V^{(n_i)} \dashrightarrow W^{(m_i)}$ nazywamy krokiem szyfrowania. Liczba n_i oznacza maksymalną szerokość szyfrowania tekstu jawnego. Liczba m_i oznacza maksymalną szerokość szyfru c_i.

Krok szyfrowania nazywany jest endomorficznym, jeśli $V = W$. Wówczas w analizie teoretycznej tekst jawny zapisuje się zazwyczaj małymi literami, natomiast kryptogram wersalikami. Wielkie litery pisane kursywą rezerwuje się z kolei do zapisu znaków klucza.
Krok szyfrowania nazywany jest rozstawieniem, jeśli generuje słowa różnej długości.

Szyfrowanie $X=[c_{i1}, c_{i2}, c_{i3}, \mathbf{K}]$ **w ramach systemu szyfrowania** M **jest skończenie generowane, jeśli jest indukowane przez skończony lub nieskończony ciąg kroków szyfrowania** $(c_{i1}, c_{i2}, c_{i3}, \mathbf{K})$, $c_i \in M$ **w wyniku konkatenacji, tzn.: dla** $x \in V^*$ **oraz** $y \in W^*$**, relacja** $x \xrightarrow{X} y$ **zachodzi wtedy i tylko wtedy, gdy istnieją rozkłady** $x = x_1 * x_2 * x_3 * \mathbf{L} * x_k$ **oraz** $y = y_1 * y_2 * y_3 * \mathbf{L} * y_k$**, spełniające** $x_j \xrightarrow{c_{ij}} y_j$ **dla** $j = 1,2,\mathbf{K},k$.

Przykład szyfrowania skończenie generowanego:

$c_i : V^{(n_i)} \dashrightarrow V^{(n_i)}$ **jest cykliczną transpozycją** n_i **elementów (**$q = 4$**),**

$n_1 = 3, n_2 = 5, n_0 = 2, n_3 = 6$**:**

$$\frac{nea}{ean}\ \frac{rlyev}{lyevr}\ \frac{er}{re}\ \frac{yinven}{inveny} \quad (c_1, c_2, c_0, c_3).$$

Różnowartościowość poszczególnych kroków szyfrowania nie zapewnia jednak różnowartościowości uzyskanego szyfrowania, np.: dla różnowartościowego odwzorowania

$$V^1 \to W^{(4)}: \begin{matrix} a \to \cdot - \\ i \to \cdot\cdot \\ l \to \cdot - \cdot\cdot \end{matrix}$$ **możemy uzyskać w** $V^* \dashrightarrow W^*$**:** $\begin{matrix} ai \to \cdot - \cdot\cdot \\ l \to \cdot - \cdot\cdot \end{matrix}$

Szyfrowanie $X=[c_{i1}, c_{i2}, c_{i3}, \mathbf{K}]$**, skończenie generowane przez** M**, jest monoalfabetyczne, jeśli zawiera lub wykorzystuje tylko jeden krok szyfrowania (alfabet). Jeśli** $q = 1$**, to każde szyfrowanie utworzone za pomocą** M **jest monoalfabetyczne.**

Szyfrowanie jest polialfabetyczne, jeśli zawiera i wykorzystuje więcej niż jeden krok szyfrowania.

Szyfrowanie jest monograficzne (jednoznakowe), jeśli $n_i = 1$ **we wszystkich użytych krokach szyfrowania, w przeciwnym przypadku szyfrowanie jest poligraficzne (wieloznakowe). W szczególnym przypadku szyfrowania poligraficznego wszystkie kroki szyfrowania w** M **charakteryzuje jednakowa maksymalna szerokość szyfrowania** n **oraz jednakowa maksymalna szerokość szyfru** m**.**

Szyfrowanie jest blokowe, jeśli $c_i : V^n \dashrightarrow W^m$ **zachodzi dla wszystkich** $c_i \in M$**. Słowo z** V^n **stanowi wówczas blok szyfrowania.**

Systemy szyfrowania o kroku szyfrowania $c_i : V^n \dashrightarrow W^m$ **określają:**

- **szyfrowanie dwuznakowe (bigram) dla** $n = 2$**,**
- **szyfrowanie trójznakowe (trigram) dla** $n = 3$**,**
- **szyfrowanie czteroznakowe (tetragram) dla** $n = 4$**,**

- **szyfrowanie jednodzielne dla** $m = 1$,
- **szyfrowanie dwudzielne dla** $m = 2$,
- **szyfrowanie trójdzielne dla** $m = 3$.

Klucz służy do wybrania odpowiedniego kroku z kryptosystemu M **i pozwala na zmianę szyfrowania zgodnie z ustalonymi z góry regułami, na przykład codziennie, po każdej wiadomości czy też po każdym znaku. Kombinatoryczna złożoność metody szyfrowania jest zdeterminowana liczbą dostępnych w jej ramach kluczy.**

Słownik (zbiór znaków) klucza – zbiór znaków K**.**

Przestrzeń klucza – zbiór słów K^* **zbudowany z alfabetu** K**.**

Kroki szyfrowania

Podstawienie proste

Podstawienie proste to podstawienie o monograficznych krokach szyfrowania $c_i : V^{(1)} \dashrightarrow W^{(m_i)}$.

W przypadku monoalfabetycznym ustala się dowolny $c_s \in M$ **i szyfrowanie przebiega wg schematu** $X = [c_s, c_s, c_s, \mathbf{K}]$.

Jednodzielne podstawienia proste $V^{(1)} \dashrightarrow W$

Heterogeniczne szyfrowanie bez homofonów i wartości zerowych $V \dashrightarrow W$ **- przypadek podstawowy, w którym jako** W **często służy alfabet dziwnie wyglądających, niezwykłych grafemów, np.:**

- **język tajski (http://www.thai-language.com):** ก ข ฃ ค ฅ ฆ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ว ศ ษ ส ห ฬ อ ฮ
- **język telugu używany w Fiji, Malezji, Singapurze, południowych Indiach (http://www.omniglot.com/writing/telugu.htm):**

అ	ఆ	ఇ	ఈ	ఉ	ఊ	ఋ	ౠ	ఌ	ఎ	ఏ	ఐ	ఒ	ఓ	ఔ
a	ā	i	ī	u	ū	ṛ	ṝ	ḷ	e	ē	ai	o	ō	au
[ʌ]	[a:]	[i]	[i:]	[u]	[u:]	[ri/ru]	[ri:, ru:]		[e]	[e:]	[aj]	[o]	[o:]	[aw]

క	k	[k]	ఖ	kh	[kʰ]	గ	g	[g]	ఘ	gh	[gʰ]	ఙ	ṅ	[ŋ]
చ	c	[ʧ]	ఛ	ch	[ʧʰ]	జ	j	[ʤ]	ఝ	jh	[ʤʰ]	ఞ	ñ	[ɲ]
ట	ṭ	[ʈ]	ఠ	ṭh	[ʈʰ]	డ	ḍ	[ɖ]	ఢ	ḍh	[ɖʰ]	ణ	ṇ	[ɳ]
త	t	[t]	థ	th	[tʰ]	ద	d	[d]	ధ	dh	[dʰ]	న	n	[n]
ప	p	[p]	ఫ	ph	[pʰ]	బ	b	[b]	భ	bh	[bʰ]	మ	m	[m]
య	y	[j]	ర	r	[r]	ల	l	[l]	వ	v	[ʋ]	ళ	!	[ɭ]
శ	ś	[ɕ]	ష	ṣ	[ʂ]	స	s	[s]				హ	h	[h]

౦	౧	౨	౩	౪	౫	౬	౭	౮	౯
0	1	2	3	4	5	6	7	8	9

- **dysk szyfrujący Giovanniego Battisty Porty (1563)**

- **tajemne znaki Karola Wielkiego**

Heterogeniczne szyfrowanie z homofonami i wartościami zerowymi $V^{(1)} - - \rightarrow W$ **wykorzystywano już w XV wieku, przywiązując jednocześnie uwagę do częstości występowania znaków, a więc np. dla samogłosek, występujących zazwyczaj najczęściej, wybierano większą liczbę homofonów. W celu zwiększenia bezpieczeństwa wykorzystywano również wartości zerowe utrudniające możliwość rozpoznawania homofonów na podstawie rozpoznawania stałego wzorca liter, otaczających je w często występujących wyrazach.**

Szyfr książkowy jest jedną z metod wykorzystujących homofony, która używana jest do dzisiaj. Szyfrem posługują się dwie osoby (zarówno nadawca jak i odbiorca) dysponujące identycznym egzemplarzem niewinnie wyglądającej książki. Nadawca wybiera kolejne litery tekstu jawnego, podając stronę x**, wiersz** y **oraz znak** z**, tworząc w ten sposób grupę szyfru** $x - y - z$ **dla każdego znaku wiadomości.**

Permutacje

W przypadku wzajemnie jednoznacznego odwzorowania $V \leftrightarrow W$**, zbiór** W **nazywany jest N-znakowym alfabetem nieuporządkowanym tekstu zaszyfrowanego, który odpowiada zbiorowi** V **nazywanego N-znakowym alfabetem standardowym tekstu jawnego. W celu zdefiniowania podstawienia wystarczy utworzyć listę odpowiadających sobie par znaków tekstu jawnego i kryptogramu, np. dla** $V = W = Z_{26}$**:**

do szyfrowania należy wykorzystać listę par uporządkowaną alfabetem standardowym

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
O	M	K	G	H	J	N	P	I	L	S	T	F	C	A	W	B	Q	D	R	V	X	E	U	Z	Y

do deszyfrowania należy wykorzystać listę par uporządkowaną alfabetem tekstu zaszyfrowanego:

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
o	q	n	s	w	m	d	e	i	f	c	j	b	g	a	h	r	t	k	l	x	u	p	v	z	y

W endomorficznym przypadku, odwzorowanie $V \leftrightarrow V$ **jest permutacją** V**. Do zapisu permutacji oprócz notacji podstawień, można również wykorzystać notację cykliczną (dla powyższego przykładu):**

$$(a\ o)(b\ m\ f\ j\ l\ t\ r\ q)(c\ k\ s\ d\ g\ n)(e\ h\ p\ w)(i)(u\ v\ x)(y\ z)$$

- **podczas szyfrowania, za znaleziony znak tekstu jawnego podstawiany jest znak następujący po nim cyklicznie**
- **podczas deszyfrowania, za znaleziony znak tekstu zaszyfrowanego podstawiany jest znak poprzedzający go cyklicznie**

Dla powyższej permutacji słowo *klopoty* **zaszyfrowane będzie słowem** *stawarz***.**

Permutacje samoodwrotne (inwolucyjne) opisywane były już przez najstarsze źródła jako proces szyfrowania i deszyfrowania stanowiący odbicie (inwolucję). Permutacje te zawierają cykle co najwyżej dwuelementowe. Na przykład w hebrajskim Starym Testamencie użyto podstawienia bustrofedonicznego, korzystającego z alfabetu odwróconego (inwersyjnego):

a	b	c	d	e	f	g	h	i	l
z	v	t	s	r	q	p	o	n	m

$$(a\ z)(b\ y)(c\ t)(d\ s)(e\ r)(f\ q)(g\ p)(h\ p)(i\ n)(l\ m)$$

Permutacje monocykliczne są permutacjami z jednym cyklem, np cykl standardowego alfabetu: Z_{26}:

$$(a\ b\ c\ d\ e\ f\ g\ h\ i\ j\ k\ l\ m\ n\ o\ p\ q\ r\ s\ t\ u\ v\ w\ x\ y\ z)$$

Wielodzielne podstawienia proste $V^{(1)} \dashrightarrow W^m$

Dwudzielne podstawienie proste $V^{(1)} \dashrightarrow W^2$ - podstawienie korzystające z bigramów było już znane w starożytności, kiedy to Polibiusz opisał piątkowe podstawienie dwudzielne dla liter greckich. We współczesnej formie dwudzielnego podstawienia prostego alfabet Z_{25} wpisuje się (wierszami bądź kolumnami bądź przy pomocy hasła) w tabelę (zwaną kwadratem bądź szachownicą Polibiusza) o wymiarach 5×5, uzyskując szyfr $Z_{25} \to Z_5 \times Z_5$:

	1	2	3	4	5
1	a	b	c	d	e
2	f	g	h	i	k
3	l	m	n	o	p
4	q	r	s	t	u
5	v	w	x	y	z

	1	2	3	4	5
1	m	l	e	k	o
2	a	b	c	d	f
3	g	h	i	n	p
4	q	r	s	t	u
5	v	w	x	y	z

Dla powyższego szyfru słowo *babcia* zaszyfrowane będzie słowem 121112132411 w pierwszym przypadku, albo 222122233322 w drugim przypadku.
Powyższy sposób szyfrowania jest powszechnie znanym i używanym w więzieniach całego świata, którego szybkość transmisji wynosi od 8 do 15 słów na minutę.
Podstawienie dwudzielne pozostawia na ogół dużo miejsca dla homofonów, np.:

	1	2	3	4	5	6	7	8	9
1 4 7	a	b	c	d	e	f	g	h	i
2 5 8	j	k	l	m	n	o	p	q	r
3 6 9	s	t	u	v	w	x	y	z	

W tym przypadku słowo *babcia* może zostać zaszyfrowane na $3^6 = 729$ sposobów, np.: 427112734917.

Homofony jednak, jak wiadomo, najlepiej wykorzystuje się w celu zrównoważenia częstości występowania znaków w kryptogramie. Poniżej znajdują się oszacowania częstości występowania 26 liter angielskiego alfabetu, uzyskane przez Bekera i Pipera:

A	B	C	D	E	F	G	H	I	J	K	L	M
0,082	0,015	0,028	0,043	0,127	0,022	0,020	0,061	0,070	0,002	0,008	0,040	0,024
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0,067	0,075	0,019	0,001	0,060	0,063	0,091	0,028	0,010	0,023	0,001	0,020	0,001

Korzystając z powyższej wiedzy można zastosować podstawienie dwudzielne z homofonami wyrównując przy tym rozkład:

	1	2	3	4	5	6	7	8	9	
4 5 6 7 8 9 0	e	t	a	o	n	i	r	s	h	70%
2 3	b	c	d	f	g	j	k	l	m	20%
1	p	q	u	v	w	x	y	z		10%

Dla powyższej propozycji szyfrowania tekst jawny *shesellsseashells* **można zaszyfrować słowem** $884901486128384858815368997138 2898$.

Trójdzielne podstawienie proste $V^{(1)} - - \rightarrow W^3$ **- podstawianie korzystające z trójznaków (trigramów), np. przy** $|W| = 3$ **daje** $3^3 = 27 > 26$ **możliwości.**

$|M| = 2$

Pięciodzielne podstawienie proste $V^{(1)} - - \rightarrow W^5$ **- podstawienie korzystające z pięcioznaków, np. wykorzystane przez Francisa Bacona przy** $|W| = 2$.

Ośmiodzielne podstawienie proste $V^{(1)} - - \rightarrow W^8$ **- np. kod ASCII z bitem parzystości.**

Rozstawianie $V^{(1)} - - \rightarrow W^{(m)}$

Oprócz homofonów oraz wartości zerowych można wykorzystywać rozstawianie alfabetów, będące odwzorowaniem V **na** $W^1 \cup W^2$**. Przykład takiego szyfru został opracowany już po 1590 roku przez Matteo Argentiego: dla alfabetu łacińskiego rozszerzonego o /et/, /con/, /non/, /che/, krok szyfrowania** $Z_{24}^{(1)} - - \rightarrow Z_{10}^1 \cup Z_{10}^2$ **wygląda następująco:**

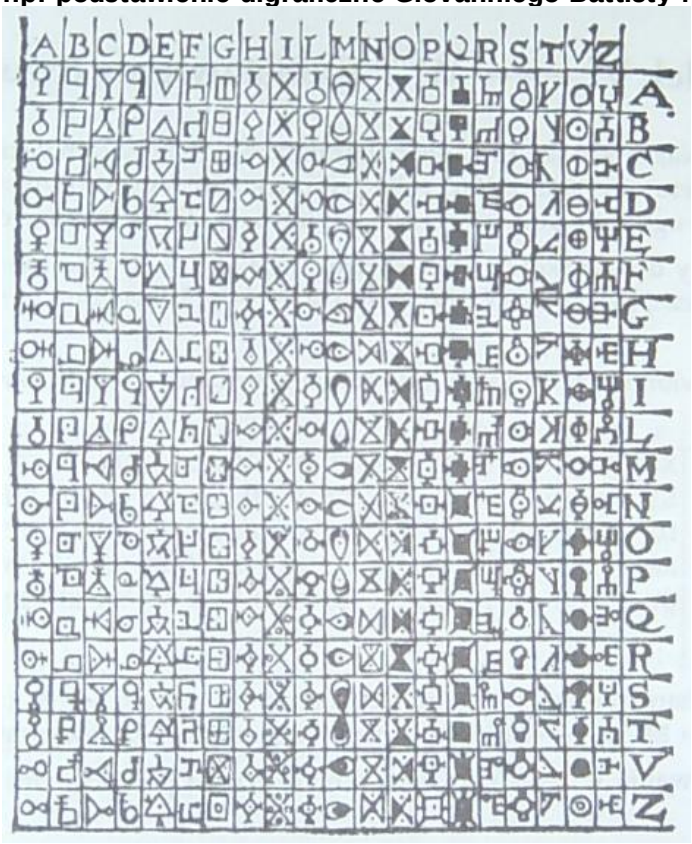
a	b	c	d	e	f	g	h	i	l	m	n	
1	86	02	20	62 82	22	06	60	3	24	26	84	
o	p	q	r	s	t	v	z	et	con	non	che	*e*
9	66	68	28	42	80 40	04	88	08	64	00	44	5 7

Oczywiście rozstawienie powinno brać pod uwagę częstość występujących znaków, tzn. powinny one otrzymywać kody o jak najmniejszej liczbie cyfr, np. dla alfabetu angielskiego wykorzystano zdanie „a sin to err" zawierające najczęściej występujące znaki w tym języku, a następnie przechodząc kolumny od lewej do prawej przydzielono im liczby od 0 do 7, a pozostałym literom przypisano liczby od 80 do 99:

s 0	u 82	b 87	w 91	a 5	y 97
c 80	d 83	e 3	f 92	g 95	h 98
i 1	j 84	k 88	l 93	m 96	n 7
o 2	p 85	q 89	r 4	t 6	v 99
x 81	z 86	. 90	/ 94		

Podstawienie dwuznakowe $V^2 \dashrightarrow W^{(m)}$

Grafemy $V^2 \to W^1$ - przyporządkowanie dwuliterowym ciągom znakowym różnych symboli, np. podstawienie digraficzne Giovanniego Battisty Porty z 1563:



Dwudzielny szyfr bigramowy $V^2 \dashrightarrow V^2$. Do jego reprezentacji używa się zazwyczaj macierzy. W przypadku $V^2 \leftrightarrow V^2$ jest to permutacja bigramów. Przykład samoodwrotnej permutacji dwuznakowej $V^2 \leftrightarrow V^2$:

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	...
a	XZ	KJ	YJ	HP	PL	EL	VB	CI	DW	XN	ZL	YP	VN	HH	CC	...
b	LP	QT	HE	RS	UR	CR	ZH	GV	WC	HL	YN	KT	WT	MC	KH	...
c	DX	MN	AO	NH	SF	GI	WL	XX	AH	GR	BZ	HS	ZU	YM	WU	...
d	KM	YZ	RY	FP	TR	CR	XE	JK	NY	PO	GJ	JR	PE	MO	VB	...
e	QU	HP	QG	JQ	YQ	OB	SA	NL	PX	OP	VS	AF	XK	XR	UQ	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

Szyfr Playfaira jest szczególną postacią dwudzielnego szyfru bigramowego. Opracowany w 1854 roku przez Charles`a Wheatstone`a spopularyzowany później przez jego przyjaciela Lyon`a Playfair`a. Krok szyfrowania wygląda następująco: na podstawie hasła w kwadrat 5×5 wpisywany jest permutowany alfabet Z_{25}:

Charles Wheatstone

p	*a*	*l*	*m*	*e*
r	*s*	*t*	*o*	*n*
b	*c*	*d*	*f*	*g*
h	*i*	*k*	*q*	*u*
v	*w*	*x*	*y*	*z*

Lyon Playfair

Tekst jawny: $kryptografii$,

dzielimy na bigramy: $kr-yp-to-gr-af-ix-ix$, **tak aby każda z par składała się z dwóch różnych liter (w razie potrzeby należy wykorzystać** x**), a następnie dla każdego bigramu wykonujemy następujący krok szyfrowania:**

- **jeśli obie litery bigramu znajdują się w tym samym wierszu, to zamieniamy je sąsiadującymi z nimi literami z prawej strony (jeśli jedna z liter znajduje się na końcu wiersza, to wiersz taki traktujemy cyklicznie, czyli bierzemy pierwszą jego literę),**
- **jeśli obie litery bigramu znajdują się w tej samej kolumnie, to zamieniamy je literami leżącymi pod nimi (analogicznie jak przy wierszach, jeśli litera znajduje się na końcu kolumny, to bierzemy literę z jej początku),**
- **jeśli obie litery bigramu znajdują się w różnych wierszach i kolumnach, to pierwsza litera bigramu zastępowana jest literą z tego samego wiersza, ale z kolumny litery drugiej, natomiast druga litera bigramu podobnie zastępowana jest literą z tego samego wiersza, ale z kolumny litery pierwszej**

Tekst zaszyfrowany:

$$ht-vm-on-bn-mc-kw-kw \Leftrightarrow htvmonbnmckwkw$$

Szyfr Delastelle`a jest kolejną szczególną postacią dwudzielnego szyfru bigramowego, opublikowaną w 1901 roku przez Felixa Marie Delastelle`a. Podobnie jak w przypadku szyfru Playfaira na podstawie hasła w kwadrat 5×5 wpisywany jest permutowany alfabet Z_{25}:

	1	2	3	4	5
1	*b*	*o*	*r*	*d*	*e*
2	*a*	*u*	*x*	*c*	*f*
3	*g*	*h*	*i*	*j*	*k*
4	*l*	*m*	*n*	*p*	*q*
5	*s*	*t*	*v*	*y*	*z*

Krok szyfrowania jest tutaj samoodwrotny i wygląda następująco:

- **dla każdej litery bigramu wykorzystuje się dwudzielne podstawienie proste, następnie transpozycję i odwrotne dwudzielne podstawienie proste:**

$$\begin{pmatrix} k & 3 & 5 \\ r & 1 & 3 \end{pmatrix} \Rightarrow \begin{pmatrix} k & r \\ 3 & 1 \\ 5 & 3 \end{pmatrix} \Rightarrow \begin{pmatrix} 3 & 1 & g \\ 5 & 3 & v \end{pmatrix}$$

Zgodnie z tym szyfrowaniem słowo *kryptografii* **zostanie zaszyfrowane słowem** *gvypsugruehihi*.

Szyfrowanie polialfabetyczne

Przyjmijmy następującą odpowiedniość między znakami a liczbami w Z_{26}**:**

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
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

Szyfrowanie polialfabetyczne wymaga, aby zbiór M **dostępnych kroków szyfrowania posiadał przynajmniej dwa elementy. Poszczególne kroki szyfrowania mogą się znacznie różnić, ale wymaga się aby wszystkie kroki szyfrowania miały jednakową szerokość szyfrowania.**

Szyfr Vigenere`a
Nazwa szyfru pochodzi on nazwiska Blavise`a de Vigenere`a, żyjącego w XVI wieku. Szyfr Vigenere`a koduje jednocześnie m znaków korzystając z m przekształceń $c_i : V \to V$ **takich, że** $c_i = x_i + k_i$**, gdzie** $k_i \in K \subseteq V = Z_{26}$ **i wszystkie działania wykonywane są w** Z_{26}**. Szyfrowanie przebiega wg schematu** $X = [c_1, c_2, ..., c_m, c_1, c_2, ...]$

Przykład
Przyjmijmy $m = 7$ **oraz niech słowem kluczowym będzie** *PSOTNIK* **tzn.** $K = \{15,18,14,19,13,8,10\}$ **przy czym ważna jest tutaj kolejność. Przypuśćmy, że tekstem jawnym jest ciąg:** *rumcajszakochalsiewhaneczce*
Przekształcamy symbole tekstu jawnego w reszty modulo 26, zapisujemy je w grupach po 6 znaków, po czym „dodajemy" słowo kluczowe modulo 26:

r	*u*	*m*	*c*	*a*	*j*	*s*	*z*	*a*	*k*	*o*	*c*	*h*	*a*
17	20	12	2	0	9	18	25	0	10	14	2	7	0
l	*s*	*i*	*e*	*w*	*h*	*a*	*n*	*e*	*c*	*z*	*c*	*e*	
11	18	8	4	22	7	0	13	4	2	25	2	4	

6	12	0	21	13	17	2	14	18	24	7	15	15	10
G	*M*	*A*	*V*	*N*	*R*	*C*	*O*	*S*	*Y*	*H*	*P*	*P*	*K*
0	10	22	23	9	15	10	2	22	16	18	15	12	
A	*K*	*W*	*X*	*J*	*P*	*K*	*C*	*W*	*Q*	*S*	*P*	*M*	

tekst zaszyfrowany: *GMAVNRCOSYHPPKAKWXJPKCWQSPM*
Aby odszyfrować ten tekst należy wykonać „odejmowanie" słowa kluczowego modulo 26 na każdej 6-znakowej grupie tego tekstu.

Warto zauważyć, że w szyfrze Vignere`a mamy 26^m możliwych słów kluczowych o długości m, zatem nawet dla stosunkowo małych wartości m wyczerpujące przeszukiwanie kluczy trwałoby dość długo. W powyższym przykładzie możliwych do zastosowania kluczy jest około $8 \cdot 10^9$.

Do szyfrowania i deszyfrowania pomocna może być również tablica regularna Trithemiusa:

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
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	A
C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	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
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
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	X
Z	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

Szyfr Hilla

Został wymyślony w 1929 r. przez Lester S. Hilla. Szyfr ten opiera się na przekształceniu liniowym tekstu jawnego w tekst zaszyfrowany – tzn. tekst jawny dzielony jest na grupy m-znakowe, a następnie każda taka grupa przekształcana jest do m-znakowego tekstu zaszyfrowanego. Szyfr koduje jednocześnie m znaków korzystając z m różnych przekształceń liniowych $c_i : V^m \to V$, takich że $c_i = \sum_{j=1}^{m} x_j \cdot k_{j,i}$, gdzie $k_{i,j}$ jest elementem macierzy K o wymiarach $m \times m$, odwracalnej nad Z_{26} i wszystkie działania wykonywane są w Z_{26}. Szyfrowanie $X : V^m \to V^m$ szyfruje m-znakowe słowo ze zbioru V^m w inne m-znakowe słowo z tego samego zbioru, zgodnie ze wzorem $X(x) = xK$, gdzie $x = [x_1, x_2, ..., x_m]$. Deszyfrowanie wykonywane jest przy pomocy macierzy odwrotnej K^{-1}.

Macierz o wyrazach rzeczywistych ma macierz odwrotną wtedy i tylko wtedy, gdy jej wyznacznik jest różny od zera. Pamiętając jednak, że działamy w Z_{26}, macierz K ma macierz odwrotną modulo 26 wtedy i tylko wtedy, gdy $NWD(\det K, 26) = 1$.

Ważne własności wyznaczników: $\det I_m = 1 \quad \det(AB) = \det A \cdot \det B$

$$1 = \det I_m = \det\left(KK^{-1}\right) = \det K \cdot \det K^{-1}$$

Ogólnie:

$$(y_1, y_2, ..., y_m) = (x_1, x_2, ..., x_m)\begin{pmatrix} k_{1,1} & k_{1,2} & \mathbf{L} & k_{1,m} \\ k_{2,1} & k_{2,2} & \mathbf{L} & k_{2,m} \\ \mathbb{M} & \mathbb{M} & & \mathbb{M} \\ k_{m,1} & k_{m,2} & \mathbf{L} & k_{m,m} \end{pmatrix}$$

Deszyfracja:

$$y = xK \ / \cdot K^{-1}$$

$$yK^{-1} = (xK)K^{-1} = x\left(KK^{-1}\right) = xI_m = x$$

Przykład

$$K = \begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix} \Rightarrow K^{-1} = \begin{pmatrix} 7 & 18 \\ 23 & 11 \end{pmatrix}$$

Tekst jawny: $july$

Do zaszyfrowania mamy dwa elementy tekstu jawnego: $ju\ (9,20)$ **oraz** $ly\ (11,24)$

Szyfrujemy:

$$(9,20)\begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix} = (99+60, 72+140) = (159, 212) = (3,4)$$

$$(11,24)\begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix} = (121+72, 88+168) = (193, 256) = (11,22)$$

W rezultacie słowo $july$ **zostało zaszyfrowane jako** $DELW$ **.**

Deszyfracja przy pomocy macierzy odwrotnej:

$$(3,4)\begin{pmatrix} 7 & 18 \\ 23 & 11 \end{pmatrix} = (21+92, 54+44) = (113, 98) = (9,20)$$

$$(11,22)\begin{pmatrix} 7 & 18 \\ 23 & 11 \end{pmatrix} = (77+506, 198+242) = (583, 440) = (11,24)$$