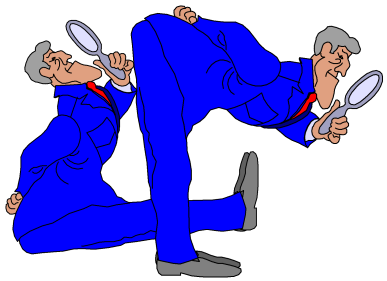
# 2. Proč jednoduše, když to jde i složitě?

Jan Paseka

Ústav matematiky a statistiky
Masarykova univerzita

27. září 2021

# O čem to bude



1. Úvod do problematiky
   - Polyabecední šifrování
2. Zneprůhlednění četností
3. Vigenerova šifra
4. Kryptoanalýza

# Polyabecední šifrování I

*Slovo, věta a z šifer se zvedá*
*poznaný život, náhlý smysl.*
Gottfried Benn

# Polyabecední šifrování I

*Slovo, věta a z šifer se zvedá*
*poznaný život, náhlý smysl.*
Gottfried Benn

**FI** V této kapitole se budeme zabývat převážně **polyabecedními** šifrováními. U těchto šifrování se písmena zprávy nešifrují pomocí téže abecedy. Zejména tedy nelze polyabecední šifrování popsat jednoduše pomocí abecedy zprávy a pod ní napsané abecedy kryptogramu.

# Polyabecední šifrování I

*Slovo, věta a z šifer se zvedá*
*poznaný život, náhlý smysl.*
Gottfried Benn

**FI** V této kapitole se budeme zabývat převážně **polyabecedními** šifrováními. U těchto šifrování se písmena zprávy nešifrují pomocí téže abecedy. Zejména tedy nelze polyabecední šifrování popsat jednoduše pomocí abecedy zprávy a pod ní napsané abecedy kryptogramu.

Přiřazení písmene zprávy k nějakému písmenu kryptogramu nesmí být prováděno svévolným způsobem. Šifrování musí splňovat silný požadavek **jednoznačnosti**; v opačném případě by nebylo možné žádné dešifrování.

# Polyabecední šifrování II

Jinak řečeno: kdyby nebylo šifrování jednoznačné, nenáchazel by se příjemce principiálně v žádné lepší situaci než Mr. X!

# Polyabecední šifrování II

Jinak řečeno: kdyby nebylo šifrování jednoznačné, nenáchazel by se příjemce principiálně v žádné lepší situaci než Mr. X!

Typickým příkladem algoritmu, u kterého není šifra jednoznačná, je **homofonní** šifra. Takovéto algoritmy budou předvedeny v následujícím odstavci.

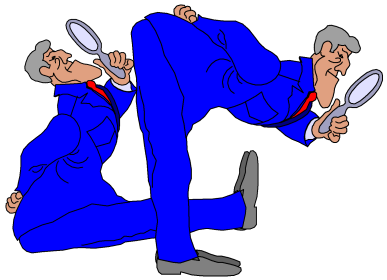
# Polyabecední šifrování II

Jinak řečeno: kdyby nebylo šifrování jednoznačné, nenáchazel by se příjemce principiálně v žádné lepší situaci než Mr. X!

Typickým příkladem algoritmu, u kterého není šifra jednoznačná, je **homofonní** šifra. Takovéto algoritmy budou předvedeny v následujícím odstavci.

Největší část kapitoly bude však věnována zkoumání takovýchto polyabecedních šifer, které vzniknou z kombinací monoabecedních algoritmů; takovýmto charakteristickým příkladem je ***Vigenerevo šifrování***.

# O čem to bude



1 Úvod do problematiky
2 Zneprůhlednění četností
3 Vigenerova šifra
4 Kryptoanalýza

- Zneprůhlednění

# Zneprůhlednění I

Jak můžeme dosáhnout toho, že všechna písmena kryptogramu se v něm vyskytují se stejnou četností? Zcela jednoduše: Šifrovací předpis přiřadí každému písmeno nikoliv znak, nýbrž ***množinu*** znaků (v našem příkladu to budou dvojice čísel), a to tak, že jsou splněny následující podmínky:

# Zneprůhlednění I

Jak můžeme dosáhnout toho, že všechna písmena kryptogramu se v něm vyskytují se stejnou četností? Zcela jednoduše: Šifrovací předpis přiřadí každému písmeno nikoliv znak, nýbrž ***množinu*** znaků (v našem příkladu to budou dvojice čísel), a to tak, že jsou splněny následující podmínky:

– Aby bylo dešifrování jednoznačné, musí být množiny příslušející různým písmenů zprávy ***disjunktní***.

# Zneprůhlednění I

Jak můžeme dosáhnout toho, že všechna písmena kryptogramu se v něm vyskytují se stejnou četností? Zcela jednoduše: Šifrovací předpis přiřadí každému písmeno nikoliv znak, nýbrž ***množinu*** znaků (v našem příkladu to budou dvojice čísel), a to tak, že jsou splněny následující podmínky:

– Aby bylo dešifrování jednoznačné, musí být množiny příslušející různým písmenů zprávy ***disjunktní***.

– Počet písmen kryptogramu, které patří k nějakému písmenu zprávy, odpovídá ***četnosti tohoto písmene***.

# Zneprůhlednění I

Jak můžeme dosáhnout toho, že všechna písmena kryptogramu se v něm vyskytují se stejnou četností? Zcela jednoduše: Šifrovací předpis přiřadí každému písmeno nikoliv znak, nýbrž ***množinu*** znaků (v našem příkladu to budou dvojice čísel), a to tak, že jsou splněny následující podmínky:

– Aby bylo dešifrování jednoznačné, musí být množiny příslušející různým písmenů zprávy ***disjunktní***.

– Počet písmen kryptogramu, které patří k nějakému písmenu zprávy, odpovídá ***četnosti tohoto písmene***.

V následujícím příkladu homofonní šifry jsou znaky kryptogramu tvořeny 100 dvojicemi číslic 00, 01, . . . , 99:

# Zneprůhlednění II

## Příklad 2.1

| **Homofonní šifra** | |
|---|---|
| *Písmeno* | *Zašifrování (němčina)* |
| *a* | *10 21 52 59 71* |
| *b* | *20 34* |
| *c* | *28 06 80* |
| *d* | *04 19 70 81 87* |
| *e* | *09 18 33 38 40 42 53 54 55 60 66 75 85 86 92 93 99* |
| *f* | *00 41* |
| *g* | *08 12 97* |
| *h* | *07 24 47 89* |
| *i* | *14 39 46 50 65 76 88 94* |
| *j* | *57* |
| *k* | *23* |
| *l* | *16 03 84* |
| *m* | *27 11 49* |
| *n* | *30 35 43 62 63 67 68 72 77 79* |
| *o* | *02 05 82* |
| *p* | *31* |
| *q* | *25* |
| *r* | *17 36 51 69 74 78 83* |
| *s* | *15 26 45 56 61 73 96* |
| *t* | *13 32 90 91 95 98* |
| *u* | *29 01 58* |
| *v* | *37* |
| *w* | *22* |
| *x* | *44* |
| *y* | *48* |
| *z* | *64* |

*Při zašifrování se přiřadí písmenu zprávy náhodně jeden z příslušných znaků kryptogramu.*

# Zneprůhlednění II

## Příklad 2.1

| **Homofonní šifra** | |
|---|---|
| *Písmeno* | *Zašifrování (němčina)* |
| *a* | *10 21 52 59 71* |
| *b* | *20 34* |
| *c* | *28 06 80* |
| *d* | *04 19 70 81 87* |
| *e* | *09 18 33 38 40 42 53 54 55 60 66 75 85 86 92 93 99* |
| *f* | *00 41* |
| *g* | *08 12 97* |
| *h* | *07 24 47 89* |
| *i* | *14 39 46 50 65 76 88 94* |
| *j* | *57* |
| *k* | *23* |
| *l* | *16 03 84* |
| *m* | *27 11 49* |
| *n* | *30 35 43 62 63 67 68 72 77 79* |
| *o* | *02 05 82* |
| *p* | *31* |
| *q* | *25* |
| *r* | *17 36 51 69 74 78 83* |
| *s* | *15 26 45 56 61 73 96* |
| *t* | *13 32 90 91 95 98* |
| *u* | *29 01 58* |
| *v* | *37* |
| *w* | *22* |
| *x* | *44* |
| *y* | *48* |
| *z* | *64* |

*Při zašifrování se přiřadí písmenu zprávy náhodně jeden z příslušných znaků kryptogramu.*

*Příjemce pak může pomocí výše uvedené tabulky jednoduše dešifrovat následující text:*

# Zneprůhlednění II

## Příklad 2.1

| Homofonní šifra | |
|---|---|
| *Písmeno* | *Zašifrování (němčina)* |
| *a* | *10 21 52 59 71* |
| *b* | *20 34* |
| *c* | *28 06 80* |
| *d* | *04 19 70 81 87* |
| *e* | *09 18 33 38 40 42 53 54 55 60 66 75 85 86 92 93 99* |
| *f* | *00 41* |
| *g* | *08 12 97* |
| *h* | *07 24 47 89* |
| *i* | *14 39 46 50 65 76 88 94* |
| *j* | *57* |
| *k* | *23* |
| *l* | *16 03 84* |
| *m* | *27 11 49* |
| *n* | *30 35 43 62 63 67 68 72 77 79* |
| *o* | *02 05 82* |
| *p* | *31* |
| *q* | *25* |
| *r* | *17 36 51 69 74 78 83* |
| *s* | *15 26 45 56 61 73 96* |
| *t* | *13 32 90 91 95 98* |
| *u* | *29 01 58* |
| *v* | *37* |
| *w* | *22* |
| *x* | *44* |
| *y* | *48* |
| *z* | *64* |

*Při zašifrování se přiřadí písmenu zprávy náhodně jeden z příslušných znaků kryptogramu.*

*Příjemce pak může pomocí výše uvedené tabulky jednoduše dešifrovat následující text:*

**0009744959950537108948310213996471374883669642775 2.**

# Zneprůhlednění II

## Příklad 2.1

| Homofonní šifra | |
|---|---|
| *Písmeno* | *Zašifrování (němčina)* |
| *a* | *10 21 52 59 71* |
| *b* | *20 34* |
| *c* | *28 06 80* |
| *d* | *04 19 70 81 87* |
| *e* | *09 18 33 38 40 42 53 54 55 60 66 75 85 86 92 93 99* |
| *f* | *00 41* |
| *g* | *08 12 97* |
| *h* | *07 24 47 89* |
| *i* | *14 39 46 50 65 76 88 94* |
| *j* | *57* |
| *k* | *23* |
| *l* | *16 03 84* |
| *m* | *27 11 49* |
| *n* | *30 35 43 62 63 67 68 72 77 79* |
| *o* | *02 05 82* |
| *p* | *31* |
| *q* | *25* |
| *r* | *17 36 51 69 74 78 83* |
| *s* | *15 26 45 56 61 73 96* |
| *t* | *13 32 90 91 95 98* |
| *u* | *29 01 58* |
| *v* | *37* |
| *w* | *22* |
| *x* | *44* |
| *y* | *48* |
| *z* | *64* |

*Při zašifrování se přiřadí písmenu zprávy náhodně jeden z příslušných znaků kryptogramu.*

*Příjemce pak může pomocí výše uvedené tabulky jednoduše dešifrovat následující text:*

**0009744959950537108948310213996471374883669642775 2.**

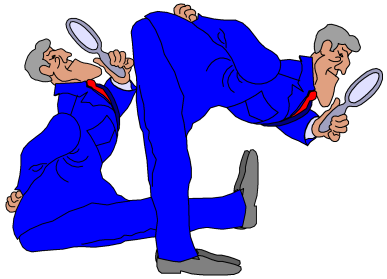
*Protože znaky byly vybrány náhodně, vyskytuje se každý znak (v našem případě dvojice číslic) stejně často (odtud jméno "homofonní"). Případný kryptoanalytik je tak postaven před podstatně těžší úlohu než při zkoumání monoabecedního šifrování.*

# Zneprůhlednění III

Avšak neměli bychom příliš jásat, neboť **kryptoanalýza** je také zde možná. Analýza je založena na pozorování, že četnosti písmen kryptogramu jsou sice stejné, ale že z uvažování nad ***dvojicemi*** znaků kryptogramu můžeme stejně dobře získat novou informaci. Budeme diskutovat dva příklady

- Uvažuje-li Mr. X ekvivalent písmena **c**, tedy dvojici 28, zjistí, že v úvahu jako ***přirozený následník*** přichází pouze určitá písmena kryptogramu. Toto jsou dvojice 07, 24, 23, 47, 89, tedy ekvivalenty písmene **h** a písmene **k**.
- Zkoumáme-li ekvivalent písmene **e**, tedy např. 99, zjistíme, že jisté znaky kryptogramu se vyskytují jako předchůdci a následníci 99 – a to prakticky stejně početně. Musí se pak jednat o ekvivalenty písmene **i**.

# O čem to bude



1 Úvod do problematiky

2 Zneprůhlednění četností

3 Vigenerova šifra
- Střídavá šifra

4 Kryptoanalýza

# Střídavá šifra I

Vigenerovo zašifrování bylo veřejnosti zpřístupněno v roce 1586 francouzským diplomatem Blaisem de **Vigenere** (1523–1596). Základní idea je používat střídavě různá monoabecední šifrování.

# Střídavá šifra I

Vigenerovo zašifrování bylo veřejnosti zpřístupněno v roce 1586 francouzským diplomatem Blaisem de **Vigenere** (1523–1596). Základní idea je používat střídavě různá monoabecední šifrování.

Tato idea je tak přirozená, že variace Vigenerova zašifrování byly vícenásobně znovuobjeveny.

# Střídavá šifra I

Vigenerovo zašifrování bylo veřejnosti zpřístupněno v roce 1586 francouzským diplomatem Blaisem de **Vigenere** (1523–1596). Základní idea je používat střídavě různá monoabecední šifrování.

Tato idea je tak přirozená, že variace Vigenerova zašifrování byly vícenásobně znovuobjeveny.

Dva z nejdůležitějších předchůdců byli Johannes **Trithemius** (1462 – 1516), jehož knihy *Poligraphia* (1518) a *Steganographia* (1531) byly uveřejněny posmrtně, a Giovanni Battista **Della Porta** (1538-1615), vynálezce přístroje *Camera obscura*, který v roce 1558 ve své knize *Magia naturalis* zveřejnil polyabecední algoritmus, který vykazuje velkou podobu s Vigenerovou šifrou.

# Střídavá šifra II

V této kapitole se budeme hlavně zabývat ***Vigenerovou šifrou***, která je nejznámější mezi všemi periodickými polyabecedními algoritmy, a to ze dvou důvodů:

# Střídavá šifra II

V této kapitole se budeme hlavně zabývat ***Vigenerovou šifrou***, která je nejznámější mezi všemi periodickými polyabecedními algoritmy, a to ze dvou důvodů:

- Vigenerova šifra je prototyp mnoha algoritmů, které byly profesionálně používány až do našeho století (Caesarova šifra je zvláštním případem Vigenerovy šifry pro klíčové slovo délky 1).

# Střídavá šifra II

V této kapitole se budeme hlavně zabývat ***Vigenerovou šifrou***, která je nejznámější mezi všemi periodickými polyabecedními algoritmy, a to ze dvou důvodů:

- Vigenerova šifra je prototyp mnoha algoritmů, které byly profesionálně používány až do našeho století (Caesarova šifra je zvláštním případem Vigenerovy šifry pro klíčové slovo délky 1).
- Při kryptoanalýze se seznámíme se dvěmi extrémně důležitými metodami, a to **Kasiského testem** a **Friedmanovým testem**.

# Střídavá šifra II

V této kapitole se budeme hlavně zabývat ***Vigenerovou šifrou***, která je nejznámější mezi všemi periodickými polyabecedními algoritmy, a to ze dvou důvodů:

- Vigenerova šifra je prototyp mnoha algoritmů, které byly profesionálně používány až do našeho století (Caesarova šifra je zvláštním případem Vigenerovy šifry pro klíčové slovo délky 1).
- Při kryptoanalýze se seznámíme se dvěmi extrémně důležitými metodami, a to **Kasiského testem** a **Friedmanovým testem**.

Abychom mohli použít Vigenerův algoritmus, potřebujeme dvě věci: **klíčové slovo** a **Vigenerův čtverec**.

# Střídavá šifra III

**Vigenerův čtverec**

| Zpráva: | 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 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Kryptogram: | 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 |

Tento čtverec se skládá z 26 abeced, které jsou napsány pod sebou takovým způsobem, že první abeceda je obyčejná abeceda, druhá abeceda je o jedno písmeno posunutá, třetí o dvě atd. Jinak řečeno: Vigenerův čtverec sestává z 26 posouvacích šifer v přirozeném pořadí.

# Střídavá šifra III

**Vigenerův čtverec**

| Zpráva: | 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 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Kryptogram: | 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 |

Tento čtverec se skládá z 26 abeced, které jsou napsány pod sebou takovým způsobem, že první abeceda je obyčejná abeceda, druhá abeceda je o jedno písmeno posunutá, třetí o dvě atd. Jinak řečeno: Vigenerův čtverec sestává z 26 posouvacích šifer v přirozeném pořadí.

Klíčovým slovem může být libovolná posloupnost písmen; pro náš demonstranční případ vybereme slovo "VENUSE".

# Střídavá šifra IV

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Klíčové slovo: | V | E | N | U | S | E | V | E | N | U | S | E |
| Zpráva: | **p** | **o** | **l** | **y** | **a** | **b** | **e** | **c** | **e** | **d** | **n** | **i.** |

# Střídavá šifra IV

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Klíčové slovo: | V | E | N | U | S | E | V | E | N | U | S | E |
| Zpráva: | **p** | **o** | **l** | **y** | **a** | **b** | **e** | **c** | **e** | **d** | **n** | **i.** |

**Při šifrování** určí písmeno klíčového slova, které stojí nad určitým písmenem zprávy příslušnou abecedu tj. ***řádku*** ve Vigenerově čtverci a pomocí této abecedy bude písmeno zprávy šifrováno.

# Střídavá šifra IV

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Klíčové slovo: | V | E | N | U | S | E | V | E | N | U | S | E |
| Zpráva: | **p** | **o** | **l** | **y** | **a** | **b** | **e** | **c** | **e** | **d** | **n** | **i.** |

**Při šifrování** určí písmeno klíčového slova, které stojí nad určitým písmenem zprávy příslušnou abecedu tj. ***řádku*** ve Vigenerově čtverci a pomocí této abecedy bude písmeno zprávy šifrováno.

Celkem tedy máme

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Klíčové slovo: | V | E | N | U | S | E | V | E | N | U | S | E |
| Zpráva: | **p** | **o** | **l** | **y** | **a** | **b** | **e** | **c** | **e** | **d** | **n** | **i** |
| Kryptogram: | **K** | **S** | **Y** | **S** | **S** | **F** | **Z** | **G** | **R** | **Y** | **F** | **M.** |

# Střídavá šifra V

Je jasné, že takováto šifrovací metoda staví Mr. X před podstatně větší problémy, než je tomu při monoabecedním šifrování.

# Střídavá šifra V

Je jasné, že takováto šifrovací metoda staví Mr. X před podstatně větší problémy, než je tomu při monoabecedním šifrování.

Četnost písmen je daleko rovnoměrnější, což lze poznat i na našem krátkém příkladu.

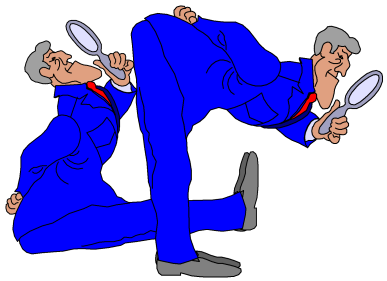
# Střídavá šifra V

Je jasné, že takováto šifrovací metoda staví Mr. X před podstatně větší problémy, než je tomu při monoabecedním šifrování.

Četnost písmen je daleko rovnoměrnější, což lze poznat i na našem krátkém příkladu.

Např. písmeno zprávy **e** bylo zašifrováno do **Z** a **R**, písmeno kryptogramu **S** vzniklo ze tří různých písmen zprávy (**o, y, a**).

# O čem to bude



1. Úvod do problematiky
2. Zneprůhlednění četností
3. Vigenerova šifra
4. Kryptoanalýza
   - Dvě metody
   - Kasiského test
   - Friedmanův test
   - Určení klíčového slova
   - Závěrečné poznámky

# Dvě metody I

Přirozeně lze i pomocí dnešních metod prolomit text zašifrovaný pomocí Vigenerovy šifry. Totiž dostatečně dlouhý text vykazuje mnoho statisticky zachytitelných pravidelností, které umožňují zjištění klíčového slova.

# Dvě metody I

Přirozeně lze i pomocí dnešních metod prolomit text zašifrovaný pomocí Vigenerovy šifry. Totiž dostatečně dlouhý text vykazuje mnoho statisticky zachytitelných pravidelností, které umožňují zjištění klíčového slova.

První uveřejněný útok byl publikován v roce 1863 pruským majorem dělostřelectva Friedrichem Wilhelmem **Kasiským** (1805–1881).

# Dvě metody I

Přirozeně lze i pomocí dnešních metod prolomit text zašifrovaný pomocí Vigenerovy šifry. Totiž dostatečně dlouhý text vykazuje mnoho statisticky zachytitelných pravidelností, které umožňují zjištění klíčového slova.

První uveřejněný útok byl publikován v roce 1863 pruským majorem dělostřelectva Friedrichem Wilhelmem **Kasiským** (1805–1881).

Druhá metoda se připisuje plukovníkovi Williamu Fredericku **Friedmanovi** (1891–1969). Obě metody slouží k určení délky klíčového slova.

# Dvě metody I

Přirozeně lze i pomocí dnešních metod prolomit text zašifrovaný pomocí Vigenerovy šifry. Totiž dostatečně dlouhý text vykazuje mnoho statisticky zachytitelných pravidelností, které umožňují zjištění klíčového slova.

První uveřejněný útok byl publikován v roce 1863 pruským majorem dělostřelectva Friedrichem Wilhelmem **Kasiským** (1805–1881).

Druhá metoda se připisuje plukovníkovi Williamu Fredericku **Friedmanovi** (1891–1969). Obě metody slouží k určení délky klíčového slova.

Protože oba testy mají i mimo speciální analýzu Vigenerovy šifry svůj dalekosáhlý a zásadní význam, představíme podrobně obě metody.

# Dvě metody II

Předpokládejme, že Mr. X zachytil následující text, o kterém ví (nebo se domnívá), že je zašifrován pomocí Vigenerovy šifry:

## Kryptogram

UEQPCVCKAHVNR ZURNLAO
KIRVGJTDVRVRI CVIDLMY
IYSBCCOJQSZNY MBVDLOK
FSLMWEFRZAVIQ MFJTDIH
CIFPSEBXMFFTD MHZGNMW

KAXAUVUHJHNUU LSVSJIP
JCKTIVSVMZJEN ZSKAHZS
UIHQVIBXMFFIP LCXEQXO
CAVBVRTWMBLNG NIVRLPF
VTDMHZGNMWKRX VRQEKVR

LKDBSEIPUCEAW JSBAPMB
VSZCFUEGITLEU OSJOUOH
UAVAGZEZISYRH VRZHUMF
RREMWKULKVKGH AHFEUBK
LRGMBJIHLII FW MBZHUMP

LEUWGRBHZOLCK CWTHWDS
ILDAGVNEMJFRV QSVIQMU
VSWMZCTHIIWGD JSXEOWS
JTKIHKEQ

# Kasiského test I

Ačkoliv tato působivá metoda analýzy polyabecedních algoritmů byla poprvé publikována Kasiským, musíme se zmínit, že anglický matematik Charles **Babbage** (1792–1871), který je mimo jiné znám svou koncepcí předchůdce moderního počítače, provedl neuveřejněná zkoumání i v kryptografii. Mj. vyvinul Kasiského test už v roce 1854, tj. devět let před Kasiským.

# Kasiského test I

Ačkoliv tato působivá metoda analýzy polyabecedních algoritmů byla poprvé publikována Kasiským, musíme se zmínit, že anglický matematik Charles **Babbage** (1792–1871), který je mimo jiné znám svou koncepcí předchůdce moderního počítače, provedl neuveřejněná zkoumání i v kryptografii. Mj. vyvinul Kasiského test už v roce 1854, tj. devět let před Kasiským.

Test je založen na následující myšlence: vyskytují-li si ve zprávě dvě posloupnosti stejných písmen (např. v němčině slovo **ein**), mohou obecně odpovídající posloupnosti v kryptogramu dopadnout různě.

# Kasiského test I

Ačkoliv tato působivá metoda analýzy polyabecedních algoritmů byla poprvé publikována Kasiským, musíme se zmínit, že anglický matematik Charles **Babbage** (1792–1871), který je mimo jiné znám svou koncepcí předchůdce moderního počítače, provedl neuveřejněná zkoumání i v kryptografii. Mj. vyvinul Kasiského test už v roce 1854, tj. devět let před Kasiským.

Test je založen na následující myšlence: vyskytují-li si ve zprávě dvě posloupnosti stejných písmen (např. v němčině slovo **ein**), mohou obecně odpovídající posloupnosti v kryptogramu dopadnout různě.

Jsou-li ale obě počáteční písmena posloupností zašifrována pomocí téhož písmene klíčového slova, jsou i obě písmena kryptogramu stejná.

# Kasiského test II

V tomto případě bude také druhé písmeno posloupnosti v zprávě zašifrováno pomocí téhož písmene klíčového slova; tedy obdržíme i v kryptogramu stejné písmeno.

# Kasiského test II

V tomto případě bude také druhé písmeno posloupnosti v zprávě zašifrováno pomocí téhož písmene klíčového slova; tedy obdržíme i v kryptogramu stejné písmeno.

To tedy znamená: Budou-li obě počáteční písmena posloupností zprávy zašifrována pomocí téhož písmene klíčového slova, pak sestávají obě posloupnosti v kryptogramu ze stejných písmen.

# Kasiského test II

V tomto případě bude také druhé písmeno posloupnosti v zprávě zašifrováno pomocí téhož písmene klíčového slova; tedy obdržíme i v kryptogramu stejné písmeno.

To tedy znamená: Budou-li obě počáteční písmena posloupností zprávy zašifrována pomocí téhož písmene klíčového slova, pak sestávají obě posloupnosti v kryptogramu ze stejných písmen.

Kdy může nastat případ, že dvě písmena jsou zašifrována pomocí téhož písmene klíčového slova? Právě tehdy, když se klíčové slovo mezi tato písmena $n$-krát vejde pro vhodné přirozené $n$.

# Kasiského test II

V tomto případě bude také druhé písmeno posloupnosti v zprávě zašifrováno pomocí téhož písmene klíčového slova; tedy obdržíme i v kryptogramu stejné písmeno.

To tedy znamená: Budou-li obě počáteční písmena posloupností zprávy zašifrována pomocí téhož písmene klíčového slova, pak sestávají obě posloupnosti v kryptogramu ze stejných písmen.

Kdy může nastat případ, že dvě písmena jsou zašifrována pomocí téhož písmene klíčového slova? Právě tehdy, když se klíčové slovo mezi tato písmena $n$-krát vejde pro vhodné přirozené $n$.
Když nyní Mr. X najde v kryptogramu dvě posloupnosti skládající se se stejných písmen, pak se může domnívat, že jejich vzdálenost je ***několikanásobek délky klíče***.

# Kasiského test III

Tato pravděpodobnost se řídí pravidlem ***"čím delší, tím milejší"***: stejná písmena nevypovídají, že víme něco o délce klíče, a také dvojice složené ze stejných písmen by mohly vzniknout náhodně.

# Kasiského test III

Tato pravděpodobnost se řídí pravidlem ***"čím delší, tím milejší"***: stejná písmena nevypovídají, že víme něco o délce klíče, a také dvojice složené ze stejných písmen by mohly vzniknout náhodně.

Z posloupností třech nebo více písmen již může Mr. X dostatečně spolehlivě usuzovat na délku klíčového slova. V našem případě pak zjistí:

# Kasiského test III

Tato pravděpodobnost se řídí pravidlem ***"čím delší, tím milejší"***: stejná písmena nevypovídají, že víme něco o délce klíče, a také dvojice složené ze stejných písmen by mohly vzniknout náhodně.

Z posloupností třech nebo více písmen již může Mr. X dostatečně spolehlivě usuzovat na délku klíčového slova. V našem případě pak zjistí:

| Posloupnost | Odstup | Rozklad na součin prvočinitelů odstupu |
|---|---|---|
| **JTD** | 50 | $2 \cdot 5 \cdot 5$ |
| **VIQM** | 265 | $5 \cdot 53$ |
| **TDMHZGNMWK** | 90 | $2 \cdot 3 \cdot 3 \cdot 5$ |
| **MWK** | 75 | $3 \cdot 5 \cdot 5$ |

# Kasiského test IV

Největší společný faktor je 5. Optimistický kryptoanalytik by mohl říci, že "že délka klíčového slova je 5" (ve skutečnosti funguje Kasiského test v praxi velmi dobře).

# Kasiského test IV

Největší společný faktor je 5. Optimistický kryptoanalytik by mohl říci, že "že délka klíčového slova je 5" (ve skutečnosti funguje Kasiského test v praxi velmi dobře).

Pokud je ale kryptoanalytik opatrný, mluví pouze o silné indicii pro délku klíčového slova 5. Jsou dva důvody pro jeho opatrnost:

# Kasiského test IV

Největší společný faktor je 5. Optimistický kryptoanalytik by mohl říci, že "že délka klíčového slova je 5" (ve skutečnosti funguje Kasiského test v praxi velmi dobře).

Pokud je ale kryptoanalytik opatrný, mluví pouze o silné indicii pro délku klíčového slova 5. Jsou dva důvody pro jeho opatrnost:

1. Mohl by nastat případ, že se vyskytují dvě posloupnosti kryptogramu ze tří nebo více stejných písmen, které mají vzdálenost nedělitelnou pěti. Pak bychom získali jako největší společný dělitel jedničku!

# Kasiského test IV

Největší společný faktor je 5. Optimistický kryptoanalytik by mohl říci, že "že délka klíčového slova je 5" (ve skutečnosti funguje Kasiského test v praxi velmi dobře).

Pokud je ale kryptoanalytik opatrný, mluví pouze o silné indicii pro délku klíčového slova 5. Jsou dva důvody pro jeho opatrnost:

1. Mohl by nastat případ, že se vyskytují dvě posloupnosti kryptogramu ze tří nebo více stejných písmen, které mají vzdálenost nedělitelnou pěti. Pak bychom získali jako největší společný dělitel jedničku!
   (V našem případě tento případ skutečně nastane: posloupnost **KAH** se vyskytne dvakrát a sice s odstupem $128 = 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2$.)

# Kasiského test V

To znamená, že nesmíme počítat největší společný dělitel "slepě" pomocí počítače, ale musíme jej určit "citem". Musíme tedy zřejmé chyby vypustit.

2. Právě proto bychom mohli přijít na myšlenku, že délka klíče by nemusela být pouze 5, nýbrž 10, 15 nebo 30 (neboť faktory 2 a 3 se vyskytují také dostatečně často).

# Kasiského test V

To znamená, že nesmíme počítat největší společný dělitel "slepě" pomocí počítače, ale musíme jej určit "citem". Musíme tedy zřejmé chyby vypustit.

2. Právě proto bychom mohli přijít na myšlenku, že délka klíče by nemusela být pouze 5, nýbrž 10, 15 nebo 30 (neboť faktory 2 a 3 se vyskytují také dostatečně často).

   Jinak řečeno: Kasiského test nám poskytuje délku klíčového slova až na ***násobek***.

# Kasiského test V

To znamená, že nesmíme počítat největší společný dělitel "slepě" pomocí počítače, ale musíme jej určit "citem". Musíme tedy zřejmé chyby vypustit.

2. Právě proto bychom mohli přijít na myšlenku, že délka klíče by nemusela být pouze 5, nýbrž 10, 15 nebo 30 (neboť faktory 2 a 3 se vyskytují také dostatečně často).

   Jinak řečeno: Kasiského test nám poskytuje délku klíčového slova až na ***násobek***.

Z výše uvedeného důvodu budeme prezentovat druhou metodu; tato určuje ***řádový odhad*** délky klíčového slova. Kombinace těchto obou metod je prakticky úplně spolehlivá.

# Kasiského test VI

Kryptogram

UEQPCVCKAHVNR ZURNLAO
KIRVGJTDVRVRI CVIDLMY
IYSBCCOJQSZNY MBVDLOK
FSLMWEFRZAVIQMFJTDIH
CIFPSEBXMFFTD MHZGNMW

KAXAUVUHJHNUU LSVSJIP
JCKTIVSVMZJEN ZSKAHZS
UIHQVIBXMFFIP LCXEQXO
CAVBVRTWMBLNG NIVRLPF
VTDMHZGNMWKRX VRQEKVR

LKDBSEIPUCEAW JSBAPMB
VSZCFUEGITLEU OSJOUOH
UAVAGZEZISYRH VRZHUMF
RREMWKULKVKGH AHFEUBK
LRGMBJIHLII FW MBZHUMP

LEUWGRBHZOLCK CWTHWDS
ILDAGVNEMJFRV QSVIQMU
VSWMZCTHIIWGD JSXEOWS
JTKIHKEQ

# Friedmanův test I

Tento postup byl vyvinut Williamem Friedmanem v roce 1925. V tomto testu se ptáme na to,***s jakou šancí se náhodně vybraný pár písmen ze zprávy sestává ze stejných písmen***. Odpověď je pak dána indexem koincidence.

# Friedmanův test I

Tento postup byl vyvinut Williamem Friedmanem v roce 1925. V tomto testu se ptáme na to,***s jakou šancí se náhodně vybraný pár písmen ze zprávy sestává ze stejných písmen***. Odpověď je pak dána indexem koincidence.

Představme si nejprve libovolnou posloupnost písmen délky $n$. Buď $n_1$ počet písmen **a**, $n_2$ počet písmen **b**, . . . , $n_{26}$ počet písmen **z**.

# Friedmanův test I

Tento postup byl vyvinut Williamem Friedmanem v roce 1925. V tomto testu se ptáme na to,***s jakou šancí se náhodně vybraný pár písmen ze zprávy sestává ze stejných písmen***. Odpověď je pak dána indexem koincidence.

Představme si nejprve libovolnou posloupnost písmen délky $n$. Buď $n_1$ počet písmen **a**, $n_2$ počet písmen **b**, ..., $n_{26}$ počet písmen **z**.

Zajímáme se o ***počet dvojic***, kdy jsou obě písmena rovna **aa**. (Nepožadujeme, aby se uvažované dvojice skládaly za sebou následujících písmen.) Pro počet prvního **a** máme právě $n_1$ možností, pro výběr druhého **a** zbývá $n_1 - 1$ možností. Protože nezáleží na pořadí písmen, je počet hledaných dvojic roven

$$\frac{n_1 \cdot (n_1 - 1)}{2}.$$

# Friedmanův test II

Je tedy počet dvojic, kdy jsou obě písmena stejná, roven

$$\frac{n_1 \cdot (n_1 - 1)}{2} + \frac{n_2 \cdot (n_2 - 1)}{2} + \cdots + \frac{n_{26} \cdot (n_{26} - 1)}{2} = \sum_{i=1}^{26} \frac{n_i \cdot (n_i - 1)}{2}.$$

# Friedmanův test II

Je tedy počet dvojic, kdy jsou obě písmena stejná, roven

$$\frac{n_1 \cdot (n_1 - 1)}{2} + \frac{n_2 \cdot (n_2 - 1)}{2} + \cdots + \frac{n_{26} \cdot (n_{26} - 1)}{2} = \sum_{i=1}^{26} \frac{n_i \cdot (n_i - 1)}{2}.$$

Šance obdržení dvojice složené ze stejných písmen je určena následujícím výrazem:

$$\mathbf{I} = \frac{\sum_{i=1}^{26} n_i \cdot (n_i - 1)}{n \cdot (n - 1)}$$

# Friedmanův test II

Je tedy počet dvojic, kdy jsou obě písmena stejná, roven

$$\frac{n_1 \cdot (n_1 - 1)}{2} + \frac{n_2 \cdot (n_2 - 1)}{2} + \cdots + \frac{n_{26} \cdot (n_{26} - 1)}{2} = \sum_{i=1}^{26} \frac{n_i \cdot (n_i - 1)}{2}.$$

Šance obdržení dvojice složené ze stejných písmen je určena následujícím výrazem:

$$\mathbf{I} = \frac{\sum_{i=1}^{26} n_i \cdot (n_i - 1)}{n \cdot (n - 1)}$$

a nazývá se Friedmanův **index koincidence**. Friedman sám značil toto číslo jako $\kappa$, proto se občas pro metodu, kterou v dalším předvedeme, používá název **kappa-test**.

# Friedmanův test III

Přibližme se nyní tomuto indexu koincidence z jiné strany. Předpokládejme, že bychom věděli, že se v našem textu vyskytuje písmeno **a** s pravděpodobností $p_1$, písmeno **b** s pravděpodobností $p_2$, ..., písmeno **z** s pravděpodobností $p_{26}$.

# Friedmanův test III

Přibližme se nyní tomuto indexu koincidence z jiné strany. Předpokládejme, že bychom věděli, že se v našem textu vyskytuje písmeno **a** s pravděpodobností $p_1$, písmeno **b** s pravděpodobností $p_2$, ..., písmeno **z** s pravděpodobností $p_{26}$.

(Konkrétní hodnoty pro pravděpodobnosti $p_i$ můžeme určit, jestliže víme, ze kterého jazyka text pochází.)

# Friedmanův test III

Přibližme se nyní tomuto indexu koincidence z jiné strany. Předpokládejme, že bychom věděli, že se v našem textu vyskytuje písmeno **a** s pravděpodobností $p_1$, písmeno **b** s pravděpodobností $p_2$, ..., písmeno **z** s pravděpodobností $p_{26}$.

(Konkrétní hodnoty pro pravděpodobnosti $p_i$ můžeme určit, jestliže víme, ze kterého jazyka text pochází.)

Představme si nyní dvě libovolně vybraná písmena našeho textu. Pravděpodobnost, že první písmeno je rovno **a** je $p_1$; tedy přibližná pravděpodobnost, že obě písmena jsou rovna **a** je ${p_1}^2$ (pokud $n$ je dostatečně velké, lze takto vzniklou chybu zanedbat). Odpovídající vztahy platí i pro ostatní písmena.

# Friedmanův test III

Přibližme se nyní tomuto indexu koincidence z jiné strany. Předpokládejme, že bychom věděli, že se v našem textu vyskytuje písmeno **a** s pravděpodobností $p_1$, písmeno **b** s pravděpodobností $p_2$, ..., písmeno **z** s pravděpodobností $p_{26}$.

(Konkrétní hodnoty pro pravděpodobnosti $p_i$ můžeme určit, jestliže víme, ze kterého jazyka text pochází.)

Představme si nyní dvě libovolně vybraná písmena našeho textu. Pravděpodobnost, že první písmeno je rovno **a** je $p_1$; tedy přibližná pravděpodobnost, že obě písmena jsou rovna **a** je ${p_1}^2$ (pokud $n$ je dostatečně velké, lze takto vzniklou chybu zanedbat). Odpovídající vztahy platí i pro ostatní písmena.

Tedy pravděpodobnost toho, že obě písmena jsou si rovna je

$$ {p_1}^2 + {p_2}^2 + \cdots + {p_{26}}^2 = \sum_{i=1}^{26} {p_i}^2. $$

# Friedmanův test IV

Toto číslo závisí přirozeným způsobem na pravděpodobnostech $p_1, p_2, \ldots, p_{26}$. Uvažme dva případy.

# Friedmanův test IV

Toto číslo závisí přirozeným způsobem na pravděpodobnostech $p_1, p_2, \ldots, p_{26}$. Uvažme dva případy.

- Pro text v německém jazyce máme

$$\sum_{i=1}^{26} {p_i}^2 = 0.0762.$$

To znamená, že náhodně zvolená dvojice písmen se skládá ze dvou stejných písmen s šancí 7,62%.

# Friedmanův test IV

Toto číslo závisí přirozeným způsobem na pravděpodobnostech $p_1, p_2, \ldots, p_{26}$. Uvažme dva případy.

- Pro text v německém jazyce máme

$$\sum_{i=1}^{26} {p_i}^2 = 0.0762.$$

To znamená, že náhodně zvolená dvojice písmen se skládá ze dvou stejných písmen s šancí 7,62%.

Jinak řečeno, každá třináctá dvojice písmen sestává ze stejných písmen.

# Friedmanův test V

- Představme si obráceně zcela náhodný text, tj. text, ve kterém jsou písmena divoce promíchána. Pak každé písmeno se zde vyskytne se stejnou pravděpodobností

$$p_i = \frac{1}{26}.$$

V tomto případě pak

$$\sum_{i=1}^{26} {p_i}^2 = \sum_{i=1}^{26} \frac{1}{26^2} = \frac{1}{26} = 0.0385.$$

Šance, že v takovémto textu najdeme dvě stejná písmena, se nám zmenšila na polovinu, a tedy každá šestadvacátá dvojice písmen sestává ze stejných písmen.

# Friedmanův test VI

1. Pokud známe pravděpodobnosti $p_1, p_2, \ldots, p_{26}$ (jako je tomu např. s němčinou či angličtinou), pak víme, že součet čtverců pravděpodobností je přibližně roven indexu koincidence:

$$I \approx \sum_{i=1}^{26} {p_i}^2.$$

Obecně lze pak dokázat, že index koincidence (nebo stejně platně i $\sum_{i=1}^{26} {p_i}^2$) je tím větší, čím je text ***nepravidelnější***, a menší, čím je text ***pravidelnější***. Hodnota 0.0385 je absolutní minimum pro index koincidence. Totiž

$$0 \leq \sum_{i=1}^{26} \left(p_i - \frac{1}{26}\right)^2 = \sum_{i=1}^{26} {p_i}^2 - \frac{1}{26}.$$

# Friedmanův test VII

2. Vraťme se na okamžik k *mono*abecedním šifrováním. Protože monoabecední šifrování je pouze permutace písmen, zůstává rozdělení četností zachováno. (Četnosti jednotlivých písmen jsou permutovány zároveň s písmeny).

# Friedmanův test VII

2. Vraťme se na okamžik k *mono*abecedním šifrováním. Protože monoabecední šifrování je pouze permutace písmen, zůstává rozdělení četností zachováno. (Četnosti jednotlivých písmen jsou permutovány zároveň s písmeny).

   Např. četnost 0.17 už nepatří písmenu **e**, nýbrž jeho ekvivalentu v kryptogramu.

# Friedmanův test VII

2. Vraťme se na okamžik k *mono*abecedním šifrováním. Protože monoabecední šifrování je pouze permutace písmen, zůstává rozdělení četností zachováno. (Četnosti jednotlivých písmen jsou permutovány zároveň s písmeny).

   Např. četnost 0.17 už nepatří písmenu **e**, nýbrž jeho ekvivalentu v kryptogramu.

   Máme tedy, že při monoabecedním šifrování index koincidence zůstává zachován, zatímco při polyabecedním šifrování klesá, vzhledem k tomu, že polyabecední šifrování bylo vytvořeno za tím účelem, aby se navzájem vyrovnaly četnosti jednotlivých písmen.

# Friedmanův test VIII

Z toho lze odvodit ***test***, který nám ukáže, zda byl kryptogram vytvořen monoabecedním šifrováním nebo ne:

# Friedmanův test VIII

Z toho lze odvodit ***test***, který nám ukáže, zda byl kryptogram vytvořen monoabecedním šifrováním nebo ne:

Nejprve vypočteme index koincidence kryptogramu. Je-li tento index přibližně 0.0762, je šifrování pravděpodobně monoabecední. Je-li index koincidence zřetelně menší, můžeme vycházet z toho, že text byl šifrovan polyabecedním šifrováním.

# Friedmanův test VIII

Z toho lze odvodit ***test***, který nám ukáže, zda byl kryptogram vytvořen monoabecedním šifrováním nebo ne:

Nejprve vypočteme index koincidence kryptogramu. Je-li tento index přibližně 0.0762, je šifrování pravděpodobně monoabecední. Je-li index koincidence zřetelně menší, můžeme vycházet z toho, že text byl šifrovan polyabecedním šifrováním.

Nyní použijeme index koincidence k výpočtu délky klíčového slova pro text zašifrovaný Vigenerovou šifrou. Cílem je určit index koincidence textu bez jeho znalosti.

# Friedmanův test VIII

Z toho lze odvodit ***test***, který nám ukáže, zda byl kryptogram vytvořen monoabecedním šifrováním nebo ne:

Nejprve vypočteme index koincidence kryptogramu. Je-li tento index přibliženě 0.0762, je šifrování pravděpodobně monoabecední. Je-li index koincidence zřetelně menší, můžeme vycházet z toho, že text byl šifrovan polyabecedním šifrováním.

Nyní použijeme index koincidence k výpočtu délky klíčového slova pro text zašifrovaný Vigenerovou šifrou. Cílem je určit index koincidence textu bez jeho znalosti.

Protože bylo použito polyabecedního algoritmu, je index koincidence menší než 0.0762. Ale o kolik menší? Odpověď je, že to závisí na délce klíčového slova.

# Friedmanův test IX

Předpokládejme, že klíčové slovo má délku $l$ a skládá se z navzájem různých písmen.

# Friedmanův test IX

Předpokládejme, že klíčové slovo má délku $l$ a skládá se z navzájem různých písmen.

Rozepišme náš kryptogram do $l$ sloupců. Pak se v prvním sloupci nacházejí písmena číslo $1, l+1, 2l+1, \ldots$, tedy všechna ta písmena, která byla zašifrována pomocí prvního písmene klíčového slova. Podobně se v druhém, ..., $l$-tém sloupci nacházejí všechna ta písmena, která byla zašifrována pomocí druhého, ..., $l$-tého písmene klíčového slova.

| Písmeno $S_i$ klíč. slova | $S_1$ | $S_2$ | $S_3$ | $S_l$ |
|---|---|---|---|---|
| | 1 | 2 | 3 | $l$ |
| | $l+1$ | $l+2$ | $l+3$ | $2l$ |
| | $2l+1$ | $2l+2$ | ... | $3l$ |
| | $3l+1$ | ... | | |
| | ... | | | |

# Friedmanův test X

Podrobnějším studiem výše uvedeného schématu lze vypočítat index koincidence.
***První pozorování***: Každý sloupec byl získán pomocí monoabecedního šifrování (dokonce pomocí posouvací šifry). Šance, že zde vybereme dvojici stejných písmen, je tedy rovna ***0,0762***. Uvažme nyní dvojice písmen, která stojí v různých sloupcích. Protože příslušné šifrovací abecedy byly vybrány "náhodně", může se takováto dvojice skládat ze stejných písmen pouze náhodným způsobem.

# Friedmanův test X

Podrobnějším studiem výše uvedeného schématu lze vypočítat index koincidence.
***První pozorování***: Každý sloupec byl získán pomocí monoabecedního šifrování (dokonce pomocí posouvací šifry). Šance, že zde vybereme dvojici stejných písmen, je tedy rovna ***0,0762***. Uvažme nyní dvojice písmen, která stojí v různých sloupcích. Protože příslušné šifrovací abecedy byly vybrány "náhodně", může se takováto dvojice skládat ze stejných písmen pouze náhodným způsobem.

Pravděpodobnost pro tento jev je podstatně nižší než 0,0762, tj. přibližně ***0,0385***. (Je to přesně $\frac{1}{26}$, jestliže je klíčové slovo náhodná posloupnost písmen. Pokud ne, je tato pravděpodobnost o něco vyšší.)

# Friedmanův test XI

***Druhé pozorování***: Vypočtěme nyní počet dvojic písmen ze stejných sloupců a z různách sloupců. Má-li náš kryptogram celkem $n$ písmen, pak v každém sloupci je právě $n/l$ písmen.

# Friedmanův test XI

***Druhé pozorování***: Vypočtěme nyní počet dvojic písmen ze stejných sloupců a z různách sloupců. Má-li náš kryptogram celkem $n$ písmen, pak v každém sloupci je právě $n/l$ písmen.

(Vzdáme se uvažování zaokrouhlovacích chyb; budeme předpokládat, že text je tak dostatečně dlouhý, že zaokrouhlovací chyby se neprojeví.)

# Friedmanův test XI

***Druhé pozorování***: Vypočtěme nyní počet dvojic písmen ze stejných sloupců a z různách sloupců. Má-li náš kryptogram celkem $n$ písmen, pak v každém sloupci je právě $n/l$ písmen.

(Vzdáme se uvažování zaokrouhlovacích chyb; budeme předpokládat, že text je tak dostatečně dlouhý, že zaokrouhlovací chyby se neprojeví.)

Pro výběr jednoho písmene máme přesně $n$ možností. Je-li toto písmeno zvoleno, pak je pevně určen i sloupec, ve kterém leží.

# Friedmanův test XI

***Druhé pozorování***: Vypočtěme nyní počet dvojic písmen ze stejných sloupců a z různách sloupců. Má-li náš kryptogram celkem $n$ písmen, pak v každém sloupci je právě $n/l$ písmen.

(Vzdáme se uvažování zaokrouhlovacích chyb; budeme předpokládat, že text je tak dostatečně dlouhý, že zaokrouhlovací chyby se neprojeví.)

Pro výběr jednoho písmene máme přesně $n$ možností. Je-li toto písmeno zvoleno, pak je pevně určen i sloupec, ve kterém leží.

V tomto sloupci máme k dispozici ještě zbývajících $n/l - 1$ písmen, tedy právě tolik možností pro výběr druhého písmene.

# Friedmanův test XII

Je tedy počet dvojic písmen, která se nacházejí ***v tom samém sloupci*** roven

$$n \cdot (\frac{n}{l} - 1)/2 = \frac{n \cdot (n - l)}{2l}.$$

# Friedmanův test XII

Je tedy počet dvojic písmen, která se nacházejí ***v tom samém sloupci*** roven

$$n \cdot (\frac{n}{l} - 1)/2 = \frac{n \cdot (n - l)}{2l}.$$

Protože máme k dispozici právě $n - n/l$ písmen mimo určený sloupec, je počet dvojic písmen ***z různých sloupců*** roven

$$n \cdot (n - \frac{n}{l})/2 = \frac{n^2 \cdot (l - 1)}{2l}.$$

# Friedmanův test XII

Je tedy počet dvojic písmen, která se nacházejí ***v tom samém sloupci*** roven

$$n \cdot (\frac{n}{l} - 1)/2 = \frac{n \cdot (n - l)}{2l}.$$

Protože máme k dispozici právě $n - n/l$ písmen mimo určený sloupec, je počet dvojic písmen ***z různých sloupců*** roven

$$n \cdot (n - \frac{n}{l})/2 = \frac{n^2 \cdot (l - 1)}{2l}.$$

Na základě výše zmíněného pak máme, že ***očekávaný počet A dvojic stejných písmen*** je roven

$$A = \frac{n \cdot (n - l)}{2l} \cdot 0,0762 + \frac{n^2 \cdot (l - 1)}{2l} \cdot 0,0385.$$

# Friedmanův test XIII

***Pravděpodobnost, že získáme dvojici složenou ze stejných písmen***, je rovna

$$\frac{A}{n \cdot (n-1)/2} = \frac{(n-l)}{l \cdot (n-1)} \cdot 0,0762 + \frac{n \cdot (l-1)}{l \cdot (n-1)} \cdot 0,0385,$$

# Friedmanův test XIII

***Pravděpodobnost, že získáme dvojici složenou ze stejných písmen***, je rovna

$$\frac{A}{n \cdot (n-1)/2} = \frac{(n-l)}{l \cdot (n-1)} \cdot 0,0762 + \frac{n \cdot (l-1)}{l \cdot (n-1)} \cdot 0,0385,$$

tj. po úpravě

$$\frac{A}{n \cdot (n-1)/2} = \frac{1}{l \cdot (n-1)} \cdot [0,0377 \cdot n + l \cdot (0,0385 \cdot n - 0,0762)].$$

# Friedmanův test XIV

Zároveň víme, že ***index koincidence*** $\mathbf{I}$ ***je aproximací*** tohoto čísla; proto platí

$$\mathbf{I} = \frac{0,0377 \cdot n}{l \cdot (n-1)} + \frac{0,0385 \cdot n - 0,0762}{n-1}.$$

# Friedmanův test XIV

Zároveň víme, že ***index koincidence* I *je aproximací*** tohoto čísla; proto platí

$$\mathbf{I} = \frac{0,0377 \cdot n}{l \cdot (n-1)} + \frac{0,0385 \cdot n - 0,0762}{n-1}.$$

Vyjádříme-li si z výše uvedeného vztahu $l$, získáme důležitou Friedmanovu formuli pro délku klíčového slova:

$$\boxed{l \approx \frac{0,0377 \cdot n}{(n-1) \cdot \mathbf{I} - 0,0385 \cdot n + 0,0762}.}$$

# Friedmanův test XV

Použijme nyní tuto formuli na náš příklad. Najdeme-li všechna $n_i$, obdržíme

$$n = 368, \sum_{i=1}^{26} {n_i}^2 = 5924.$$

# Friedmanův test XV

Použijme nyní tuto formuli na náš příklad. Najdeme-li všechna $n_i$, obdržíme

$$n = 368, \sum_{i=1}^{26} {n_i}^2 = 5924.$$

Máme tedy

$$\mathbf{I} = \frac{5924}{135056} = 0,0439.$$

# Friedmanův test XV

Použijme nyní tuto formuli na náš příklad. Najdeme-li všechna $n_i$, obdržíme

$$n = 368, \sum_{i=1}^{26} {n_i}^2 = 5924.$$

Máme tedy

$$\mathbf{I} = \frac{5924}{135056} = 0,0439.$$

Jedná se tedy s velkou pravděpodobností o polyabecední šifrování. Spočtěme nyní délku klíčového slova $l$ :

$$l \approx 6,5.$$

# Friedmanův test XV

Použijme nyní tuto formuli na náš příklad. Najdeme-li všechna $n_i$, obdržíme

$$n = 368, \sum_{i=1}^{26} {n_i}^2 = 5924.$$

Máme tedy

$$\mathbf{I} = \frac{5924}{135056} = 0,0439.$$

Jedná se tedy s velkou pravděpodobností o polyabecední šifrování. Spočtěme nyní délku klíčového slova $l$ :

$$l \approx 6,5.$$

To ukazuje současně s výsledkem testu Kasiského na to, že délka klíčového slova ***je skutečně 5*** (a ne 10, 15 nebo 20).

# Určení klíčového slova I

Jakmile je zjištěna délka klíčového slova, jde o to poznat klíčové slovo samotné. Ale to už není tak těžké.

# Určení klíčového slova I

Jakmile je zjištěna délka klíčového slova, jde o to poznat klíčové slovo samotné. Ale to už není tak těžké.

Pokud kryptoanalytik Mr. X zná délku klíčového slova, ví, že písmena č. $1, l+1, 2l+1, \ldots$ příp. č. $2, l+2, 2l+2, \ldots$ atd. byla získána pomocí monoabecedního šifrování (dokonce pomocí posouvací šifry). Zpravidla tedy stačí nalézt ekvivalenty písmene **e**.

# Určení klíčového slova I

Jakmile je zjištěna délka klíčového slova, jde o to poznat klíčové slovo samotné. Ale to už není tak těžké.

Pokud kryptoanalytik Mr. X zná délku klíčového slova, ví, že písmena č. $1, l+1, 2l+1, \ldots$ příp. č. $2, l+2, 2l+2, \ldots$ atd. byla získána pomocí monoabecedního šifrování (dokonce pomocí posouvací šifry). Zpravidla tedy stačí nalézt ekvivalenty písmene **e**.

V našem příkladu je $l = 5$. Ze 74 písmen první "monoabecední" části je 14 rovno **V**. Proto odpovídá **e** písmenu **V**. Z Vigenerova čtverce pak obdržíme, že ***první písmeno klíčového slova*** je "***R***".

# Určení klíčového slova II

Analogicky, ze 74 písmen druhé, třetí, čtvté a páté "monoabecední" částí je 11 rovno **E**, 8 rovno **H**, 21 rovno **M** a 13 rovno **S**.

# Určení klíčového slova II

Analogicky, ze 74 písmen druhé, třetí, čtvté a páté "monoabecední" částí je 11 rovno **E**, 8 rovno **H**, 21 rovno **M** a 13 rovno **S**.

Proto odpovídá **e** písmenu **E**, **H**, **M** a **S**.

# Určení klíčového slova II

Analogicky, ze 74 písmen druhé, třetí, čtvté a páté "monoabecední" částí je 11 rovno **E**, 8 rovno **H**, 21 rovno **M** a 13 rovno **S**.

Proto odpovídá **e** písmenu **E**, **H**, **M** a **S**.

Opětovným nahlédnutím do Vigenerova čtverce pak obdržíme, že další písmena klíčového slova jsou po řadě "**A**", "**D**", "**I**" a "**O**".

# Určení klíčového slova II

Analogicky, ze 74 písmen druhé, třetí, čtvté a páté "monoabecední" částí je 11 rovno **E**, 8 rovno **H**, 21 rovno **M** a 13 rovno **S**.

Proto odpovídá **e** písmenu **E**, **H**, **M** a **S**.

Opětovným nahlédnutím do Vigenerova čtverce pak obdržíme, že další písmena klíčového slova jsou po řadě "**A**", "**D**", "**I**" a "**O**".

Rozšifrování textu použitím klíčového slova "**RADIO**" je již standardní záležitostí.

# Určení klíčového slova III

Snadným porovnáním získáme

**Zpráva**

**denhoechs t eno rganisa**
**t ionsstander f uhrdiek**
**ryptologi e inv enedigw**
**osieinforme i n erstaat**
**l ichenbue r o t a et igke i**

**uka tenimmonat bekamen**
**eswurded a fuer geso r g t**
**dasssiew a ehre ndih r e r**
**arbeitn i chtge stoe r tw**
**urdensi e dur f t enih r eb**

**tausgeueb t wu r deesgab**
**schluesse l sek retaere**
**dieihrbue r o im dogenpa**
**lasthatten und fuer ihr**
**etaet igke i t r u ndzehnd**

**ue rosabe r auch nich t ve**
**r lassenb e vors ieei neg**
**estellte aufga bege l oe**
**sthatten**

# Závěrečné poznámky I

Viděli jsme, že každé Vigenerovo šifrování s dostatečně krátkým klíčem (aby se mohla ke slovu dostat pravděpodobnost ve sloupcích) lze jednoduchým způsobem rozšifrovat.

# Závěrečné poznámky I

Viděli jsme, že každé Vigenerovo šifrování s dostatečně krátkým klíčem (aby se mohla ke slovu dostat pravděpodobnost ve sloupcích) lze jednoduchým způsobem rozšifrovat.

Uvažujme nyní Vigenerovo šifrování ***s dlouhým klíčovým slovem***. Budeme předpokládat, že klíčové slovo je dlouhé právě tak, jak je délka zprávy. Ukážeme dva triky, které Mr. X znemožní účinně využít výše uvedených testů.

# Závěrečné poznámky I

Viděli jsme, že každé Vigenerovo šifrování s dostatečně krátkým klíčem (aby se mohla ke slovu dostat pravděpodobnost ve sloupcích) lze jednoduchým způsobem rozšifrovat.

Uvažujme nyní Vigenerovo šifrování ***s dlouhým klíčovým slovem***. Budeme předpokládat, že klíčové slovo je dlouhé právě tak, jak je délka zprávy. Ukážeme dva triky, které Mr. X znemožní účinně využít výše uvedených testů.

**Trik č.1** Mohli bychom se pokusit použít jako klíč text knihy. Takový klíč má zcela určitě tu výhodu, že ho lze přenést bez velkých problémů.

# Závěrečné poznámky II

Např. stačí podat příjemci informaci ***Eugen Eichhorn: Felix Hausdorff - Paul Mongré***, aby mohl začít dešifrovat kryptogram pomocí následujícího slova:

> *As you have already heard, Hausdorff was born in the Silesian metropolis Breslau, today called Wroclaw. In the last days of the Second World War, the German Wehrmacht declared Breslau a fortress; the result was its complete destruction. That happened . . .*

# Závěrečné poznámky II

Např. stačí podat příjemci informaci ***Eugen Eichhorn: Felix Hausdorff - Paul Mongré***, aby mohl začít dešifrovat kryptogram pomocí následujícího slova:

> *As you have already heard, Hausdorff was born in the Silesian metropolis Breslau, today called Wroclaw. In the last days of the Second World War, the German Wehrmacht declared Breslau a fortress; the result was its complete destruction. That happened . . .*

V případě použití takovéhoto klíče se všechny metody na určení délky klíče minou účinkem. Protože však klíč tvoří souvislý text nějaké řeči (angličtina, němčina, apod.), působí na kryptogram statisticky signifikantní data, takže nemůžeme takovouto šifru označit za zcela bezpečnou.

# Závěrečné poznámky III

První, kdo odhalil tuto slabinu, byl opět Friedman. Proto půjdeme ještě o kus dál.

# Závěrečné poznámky III

První, kdo odhalil tuto slabinu, byl opět Friedman. Proto půjdeme ještě o kus dál.

**Trik č.2** V případě triku č.1 mohl Mr. X ještě použít nějakou statistiku v důsledku tvaru klíčového slova.

# Závěrečné poznámky III

První, kdo odhalil tuto slabinu, byl opět Friedman. Proto půjdeme ještě o kus dál.

**Trik č.2** V případě triku č.1 mohl Mr. X ještě použít nějakou statistiku v důsledku tvaru klíčového slova.

Proto nyní zvolíme za klíčové slovo prakticky nekonečnou, náhodnou posloupnost písmen, na kterou si se statistickými testy nepřijdeme. Např. lze za ni zvolit výsledky opakování vrhu ideální 26-hranou kostkou.

# Závěrečné poznámky III

První, kdo odhalil tuto slabinu, byl opět Friedman. Proto půjdeme ještě o kus dál.

**Trik č.2** V případě triku č.1 mohl Mr. X ještě použít nějakou statistiku v důsledku tvaru klíčového slova.

Proto nyní zvolíme za klíčové slovo prakticky nekonečnou, náhodnou posloupnost písmen, na kterou si se statistickými testy nepřijdeme. Např. lze za ni zvolit výsledky opakování vrhu ideální 26-hranou kostkou.

Lze pak ukázat, že takovýto způsob šifrování je **dokonce teoreticky bezpečný!** Jinak řečeno: nabízí nám **perfektní bezpečnost**.

# Závěrečné poznámky III

První, kdo odhalil tuto slabinu, byl opět Friedman. Proto půjdeme ještě o kus dál.

**Trik č.2** V případě triku č.1 mohl Mr. X ještě použít nějakou statistiku v důsledku tvaru klíčového slova.

Proto nyní zvolíme za klíčové slovo prakticky nekonečnou, náhodnou posloupnost písmen, na kterou si se statistickými testy nepřijdeme. Např. lze za ni zvolit výsledky opakování vrhu ideální 26-hranou kostkou.

Lze pak ukázat, že takovýto způsob šifrování je **dokonce teoreticky bezpečný!** Jinak řečeno: nabízí nám **perfektní bezpečnost**.

Takovýmito perfektními systémy se budeme zabývat v následující kapitole.