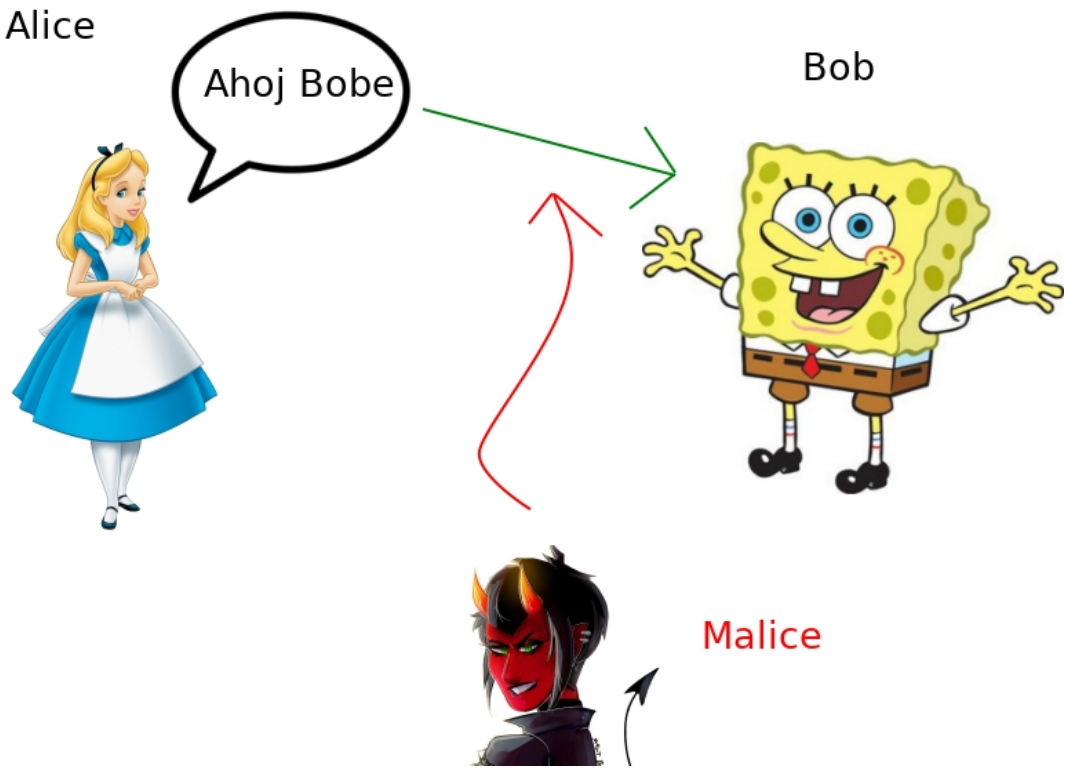
Asymetrické šifrování

Michal Hrbek

Matematický ústav Akademie věd ČR

Mírně upravená prezentace přednášky v rámci Týdne vědy a techniky AV ČR v roce 2019.

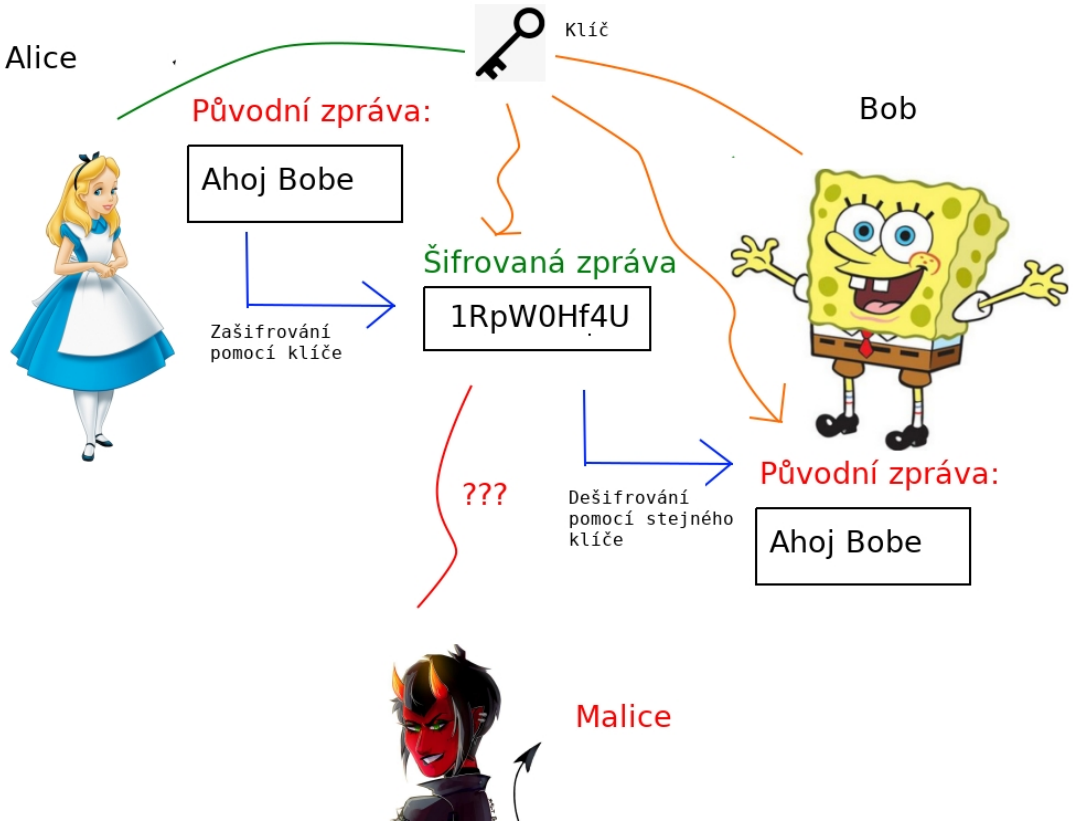
Kryptografie, neboli **šifrování**, je nauka o utajování zpráv tak, aby jejich obsahu mohl porozumět jen vybraný adresát.



Alice by ráda Bobovi poslala zprávu "Ahoj Bobe" tak, aby si obsah zprávy mohl přečíst pouze Bob. Konverzaci se snaží odposlechnout zlomyslná Malice.

Symetrická kryptografie

V **symetrické šifře** se text **šifruje** i **dešifruje** pomocí jednoho stejného klíče, který zná odesílatel i příjemce.



Alice i Bob se předem dohodnou na tajném klíči. Pomocí tohoto klíče Alice svoji zprávu zašifruje a pošle Bobovi. Zašifrovaná zpráva je pro odposlouchávající Malici nesrozumitelná. Bob nakonec zašifrovanou zprávu dešifruje pomocí stejného tajného klíče.

Příklad symetrické šifry - Substituční šifra

Substituční šifra spočívá v nahrazení každého znaku zprávy jiným znakem podle nějakého pravidla.

- **Příklad: Caesarova šifra**
- Klíč: Číslo n od 1 do 25.
- Zašifrování: Každé písmeno v textu se posune o n pozic v abecedě dopředu.
- Dešifrování: Každé písmeno v textu se posune o n pozic v abecedě dozadu.
- Alice s Bobem se dohodnou v tajnosti třeba na klíči $n = 5$.
- Původní zpráva: **"AHOJ BOBE"**
- Šifrovaná zpráva: **"FMTO GTGJ"**
- Nepříliš bezpečná šifra - jen 25 možných klíčů.

- **Příklad: Substituční tabulka**
- Klíč: Tabulka následujícího tvaru

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

- Zašifrování: Každé písmeno v textu z horního řádku se přepíše na odpovídající písmeno z dolního řádku.
- Dešifrování: Přesně opačný postup.
- Původní zpráva: **"AHOJ BOBE"**
- Šifrovaná zpráva: **"QABX DBDL"**
- Spousta klíčů k dispozici (konkrétně 26!, tedy 26 faktoriál).
- **Kryptoanalýza** - umění dešifrování zpráv i přesto, že neznáme klíč.
- Substituční šifry jsou náchylné na tzv. **frekvenční analýzu**.

Příklad symetrické šifry - Permutační šifra

V permutační šifře nenahrazujeme písmena samotná, ale měníme jejich pořadí.

- Klíč: Tabulka tvaru

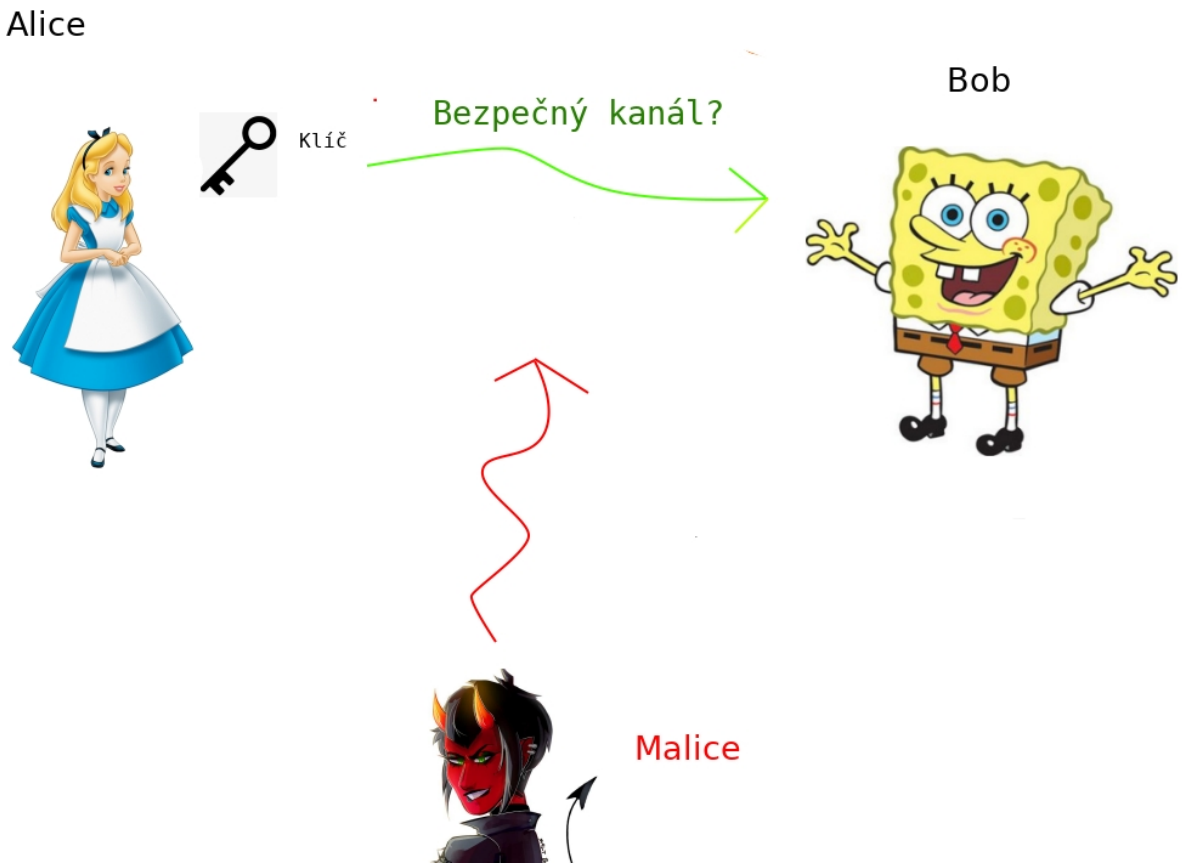
1	2	3	4
3	2	4	1

- Zašifrování: Pořadí každé následující čtveřice znaků zpřeházíme podle tabulky v klíči.
- Dešifrování: Přesně opačný postup.
- Původní zpráva: **"AHOJ BOBE"**
- Šifrovaná zpráva: **"JHAO EOBB"**

Moderní symetrické šifry a jejich výhody

- Jedna z nejpoužívanějších symetrických šifer - **Advanced Encryption Standard (AES)**. Jde o (velmi) chytrou kombinaci substitučních a permutačních šifer. Standardně se používají klíče délek 128, 192 a 256 bitů.
- Symetrické kryptosystémy, včetně zmíněného AES, jsou velmi rychlé. Je možno efektivně šifrovat široký tok dat za běhu nebo svižně zašifrovávat velké soubory, případně celý obsah disku.

Hlavní problém symetrického šifrování

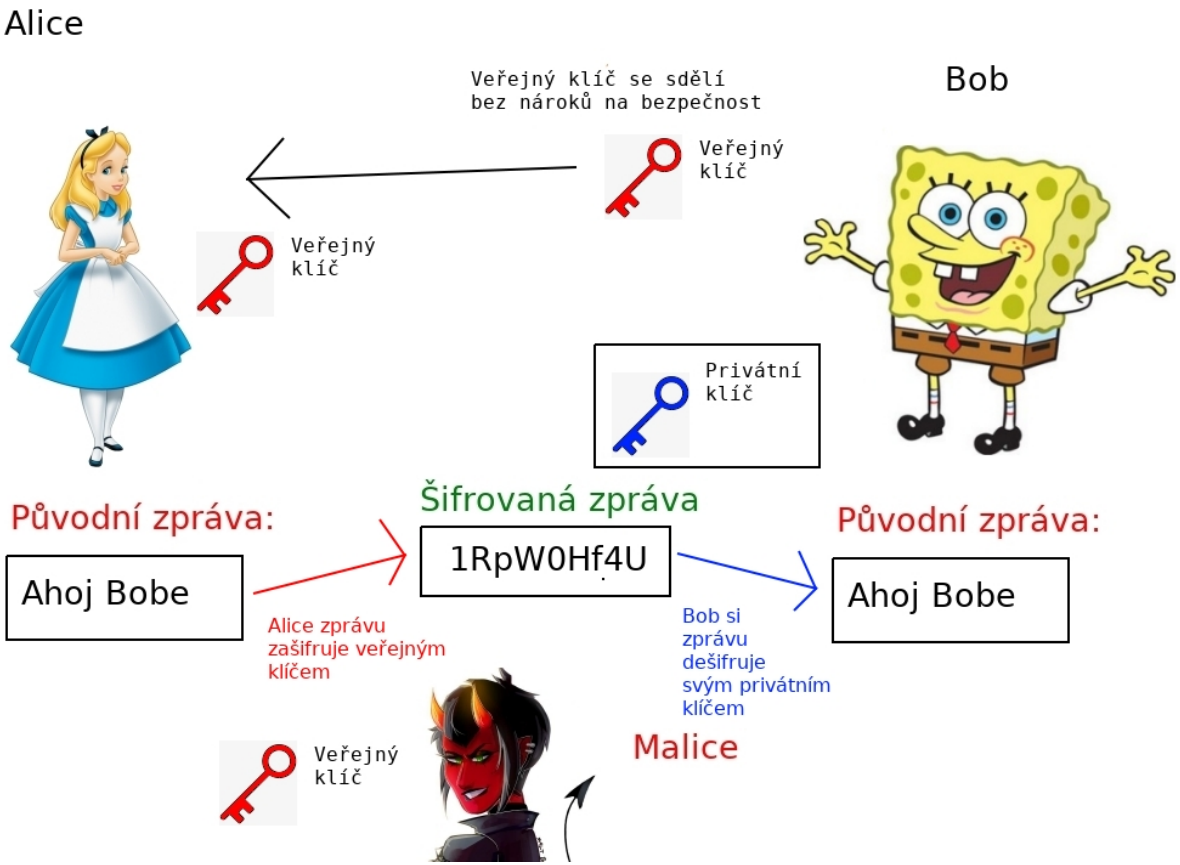


Alice potřebuje před navázáním zašifrované komunikace nějakým dokonale bezpečným způsobem sdělit Bobovi tajný klíč. **V praxi většinou takový bezpečný kanál neexistuje!**

Pokud by dokonale bezpečný kanál existoval (představujte si hlavně komunikaci po internetu), nemuseli bychom se šifrováním zabývat v první řadě.

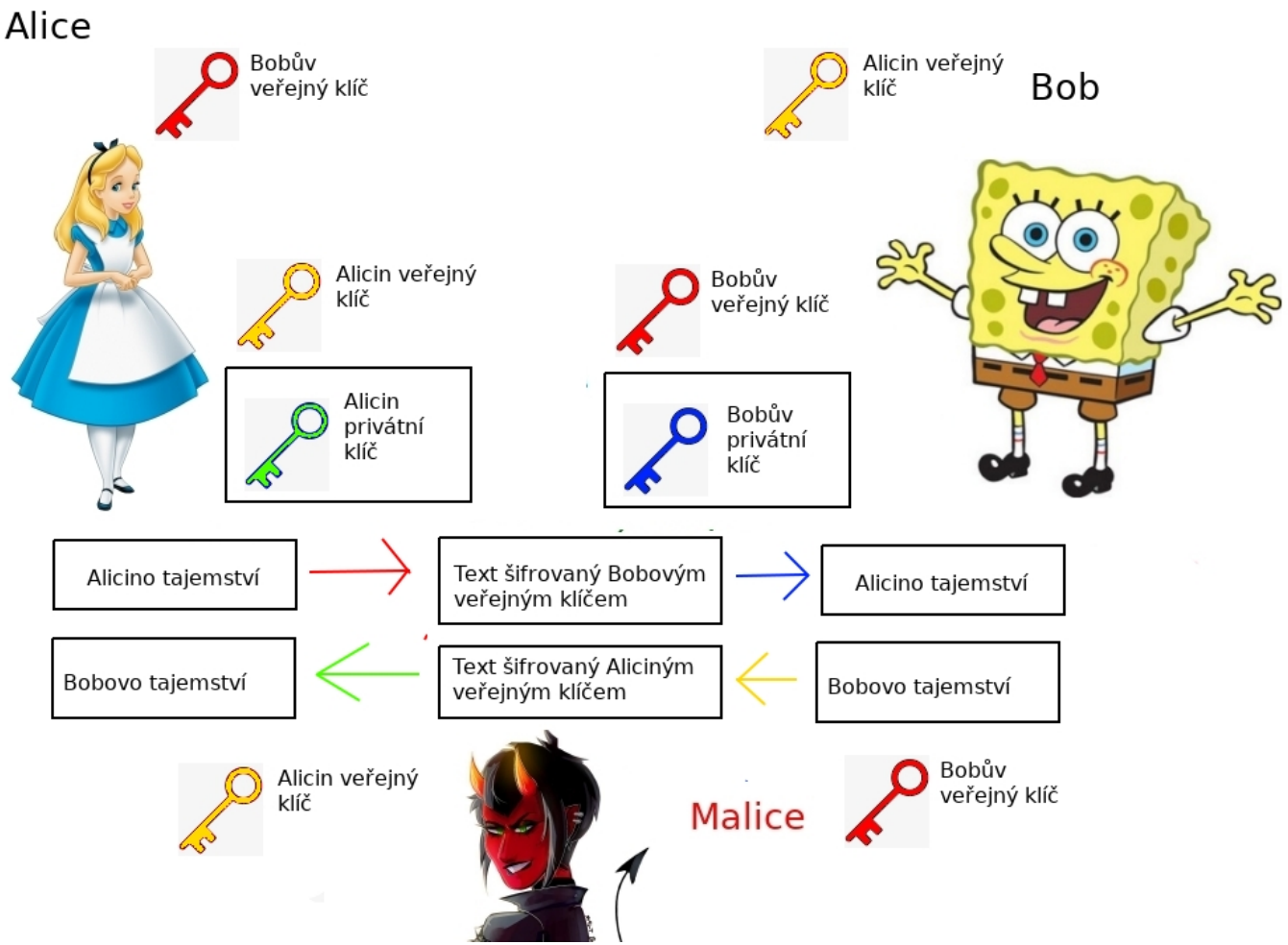
Řešení - Asymetrická kryptografie

V asymetrické kryptografii figurují klíče dva - jeden **privátní** a jeden **veřejný**. Pomocí veřejného klíče se zpráva zašifruje, ale jen držitel privátního klíče je schopen zprávu dešifrovat.



Veřejný klíč sdělí Bob Alici a v zásadě nevadí, že ho Malice může odposlechnout. Privátní (tajný) klíč zná pouze Bob a nesdělí ho vůbec nikomu. Alice šifruje veřejným klíčem, ale dešifrovat takto utajenou zprávu může **jen** Bob.

Diffieho-Hellmanovo schéma

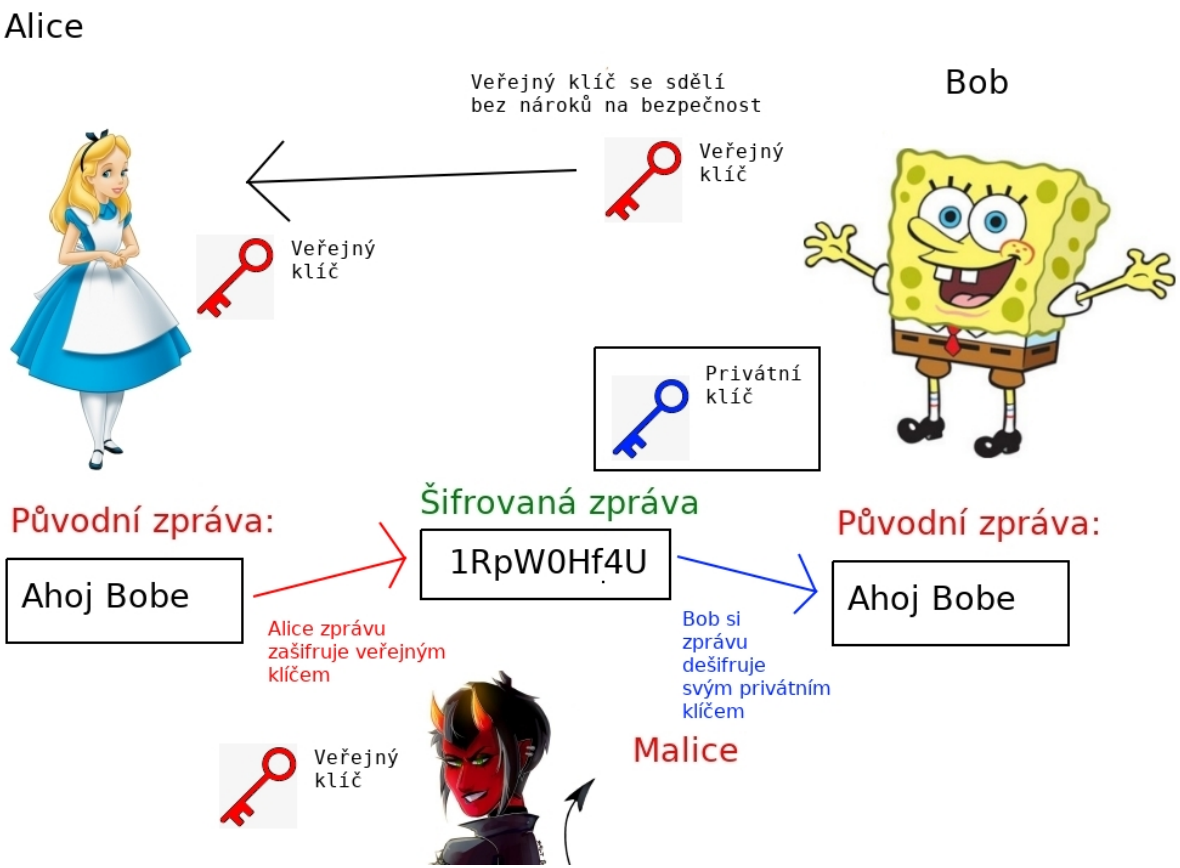


Diffieho-Hellmanovo schéma je vlastně jen zdvojený systém asymetrické šifry z předchozího obrázku. Aby spolu mohli Alice a Bob komunikovat obousměrně, každý z nich si vytvoří svůj pár veřejného a privátního klíče. Oba dva pak své veřejné klíče pošlou tomu druhému, čímž je potenciálně zveřejní i Malici. Své privátní klíče si nechají pro sebe a používají je na dešifrování zpráv od svého protějšku.

Výhody a nevýhody asymetrického šifrování

+ Zásadní výhodou je, že pro navázání zabezpečené komunikace není třeba předání tajné informace nezašifrovanou zprávou.
– Malice má přístup k veřejnému klíči. Může tak šifrovat svoje zprávy, a pak se je snažit podvrhnout jako zprávy od Alice/Boba.
+ (Některé) asymetrické algoritmy se dají využit k robustnímu systému digitálních podpisů, což efektivně řeší předchozí problém.
– Oproti symetrickým šifrám jsou ty asymetrické komplikovanější. Co je podstatnější, jsou také výrazně náročnější na výpočetní výkon.
+ Asymetrické šifrování se dá výborně kombinovat se symetrickým. Pomocí kanálu navázaného asymetrickou šifrou je možno si bezpečně předat klíč k jiné symetrické šifře, použité pro následnou rychlejší zabezpečenou komunikaci.

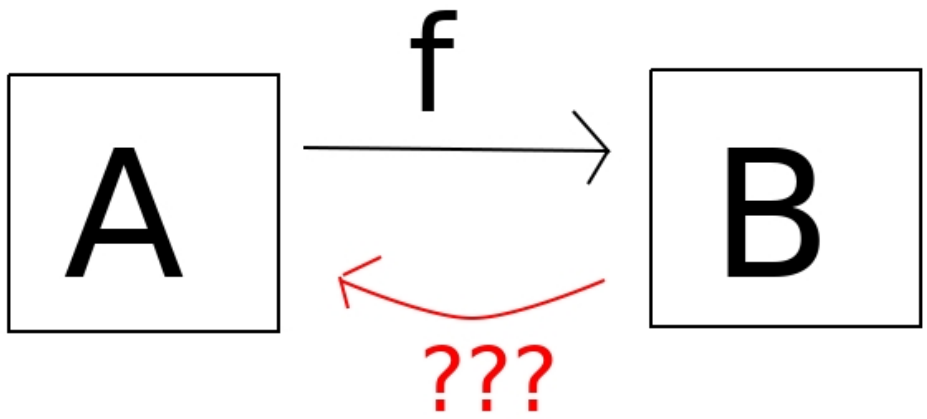
Jak ale asymetrická šifra funguje?



Jak Bob vytvoří privátní a veřejný klíč tak, aby zprávy zašifrované veřejným klíčem mohl přečíst jen on?

Jednosměrné funkce

Za funkcí asymetrické šifry stojí teoretický koncept z matematiky - tzv. **jednosměrná funkce**.



Jednosměrná funkce je přiřazení f z množiny A do množiny B, pro které je **"snadné"** spočítat hodnotu $f(a)$ pro každý prvek a, ale je **"extrémně těžké"** nalézt pro hodnotu b nějaký vzor a, tak aby $f(a) = b$.
Jak se dají použít v asymetrickém šifrování:

- Bob zvolí a. Hodnota $b = f(a)$ - **veřejný klíč**.
- Původní hodnota a - **privátní klíč**, který je těžké odhalit.

Kandidát jednosměrné funkce

Jednosměrné funkce se dají definovat matematicky **"rigorózně"**, my si zde však vystačíme s touto nepřesnou intuicí. Je jedním z velkých problémů současné matematiky rozhodnout, zda takové přesně definované jednosměrné funkce vůbec existují.
Většina matematiků však věří, že jednosměrné funkce existují.
Někteří "nadějní kandidáti" se dokonce používají v denní praxi!

Kandidát jednosměrné funkce

Definujeme funkci f, která dvěma **prvočíslům** p, q přiřadí jejich součin: $f(p, q) = p \cdot q$.

- Prvočísla p a q - **privátní tajemství**.
- Součin $N = p \cdot q$ - **zveřejněná informace**.

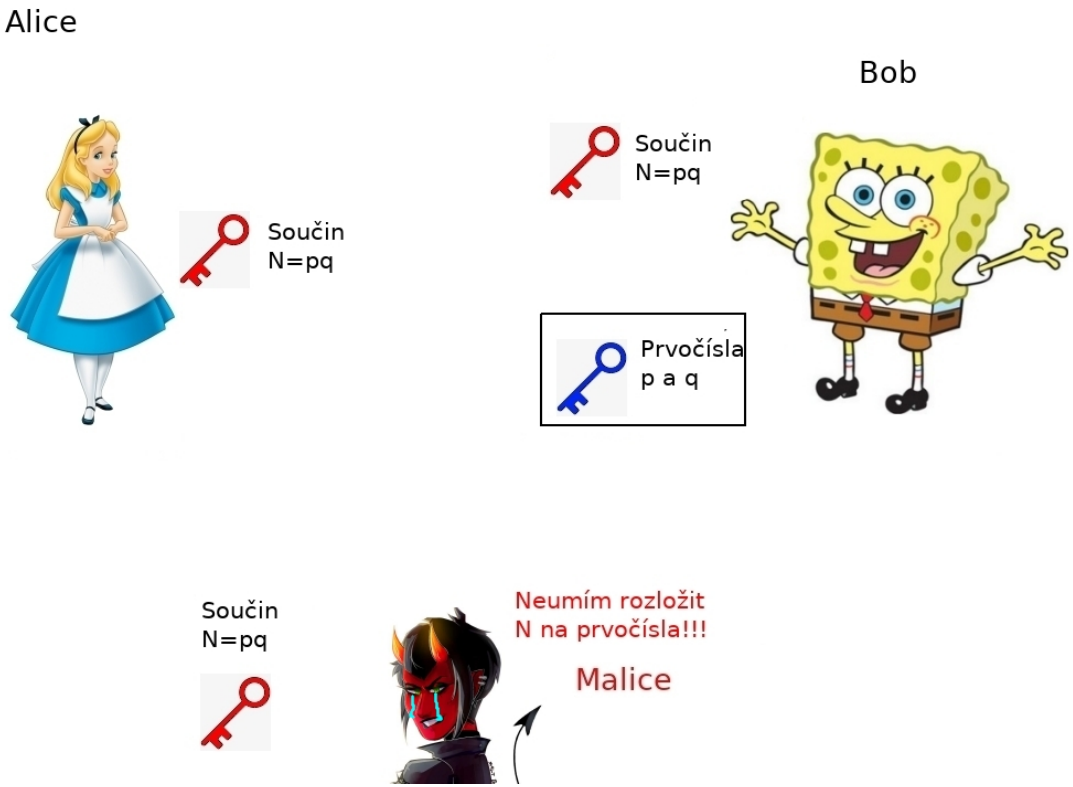
Obtížnost faktorizace na prvočísla

Zkuste si: Rozložte číslo 1517 na součin prvočísel.

Pravděpodobně použijete metodu podobnou Eratosthenově sítu - budete postupně zkoušet dělitelnost čísly 2 až $\sqrt{1517}$ a po cestě zahazovat násobky již odzkoušených čísel. Pro velká čísla taková metoda může zabrat dlouho!

Pokud zvolíme dostatečně **náhodným** způsobem dostatečně **velká** prvočísla p a q, ani supersilný počítač nedokáže v rozumném čase rozložit součin $N = p \cdot q$ zpět na prvočísla p a q.

- Pro velká čísla (~100 cifer a více) neexistuje efektivnější algoritmus faktorizace než jisté, poměrně mírné, vylepšení starodávného Eratosthenova síta.
- Největší rozložené číslo (zvolené RSA-768) mělo 232 dekadických cifer a zabralo obrovské síti počítačů 2 roky (Lenstra et al., 2009).
- Doporučená délka klíče pro RSA je dnes 2048 bitů.
- Běžnému stolnímu počítači by faktorizace takového čísla zabrala několik statisíců násobků stáří celého vesmíru.



Nápad, jak zvolit privátní a veřejný klíč, bychom měli. Stále ale nevíme, jak šifrovat zprávy. Naštěstí na to již před námi přišli tři chytří pánové - Ron Rivest, Adi Shamir a Leonard Adleman.

Kryptosystém RSA

Obrázek: Shamir, Rivest, Adleman

Modulární aritmetika

V srdci šifry RSA je matematika, přesněji řečeno, teorie čísel.

- Místo běžného počítání s celými čísly se nám bude hodit počítat **"modulo"** nějaké zvolené číslo N.
- Zvolme si nějaké kladné číslo N, zvané **modulus**.
- Operace **modulo** N vrací zbytek po celočíselném dělení modulem N. Tedy například pro $N = 13$:

$$50 \equiv 11 \pmod{13}$$
$$26 \equiv 0 \pmod{13}$$
$$120 \equiv 3 \pmod{13}$$

Čti postupně: "50 dává zbytek 11 po dělění 13", "26 dává zbytek 0 po dělění 13", "120 dává zbytek 3 po dělění 13".

- **Uhádnete?** Kolik vychází $43275320587 \equiv \ ? \pmod{10}$?
Nápověda: Nemusíte nic počítat, stačí si uvědomit, co to znamená počítat zbytek po dělení desítkou.

Rovnice v modulární aritmetice

Následuje trocha ilustrace toho, jak se dá počítat v modulární aritmetice.

- Vyřešme rovnici: $5x \equiv 1 \pmod{11}$

x	0	1	2	3	4	5	6	7	8	9	10
$5x$	0	5	10	15	20	25	30	35	40	45	50
$5x \pmod{11}$	0	5	10	4	9	3	8	2	7	1	6

  Nalezli jsme řešení $x = 9$. Všechna řešení rovnice jsou tvaru $x = 9 + 11k$ pro libovolné k.

- Vyřešme rovnici: $6x \equiv 1 \pmod{9}$

x	0	1	2	3	4	5	6	7	8
$6x$	0	6	12	18	24	30	36	42	48
$6x \pmod{9}$	0	6	3	0	6	3	0	6	3

  Rovnice nemá řešení.

- Povšimněte si: čísla 5 a 11 jsou **nesoudělná**, zatímco čísla 6 a 9 mají největší společný dělitel roven 3.
- Řešení modulárních rovnic úzce souvisí s **dělitelností** a tvoří důležitou součást oboru zvaného **teorie čísel**.

Eulerova funkce φ

Leonhard Euler, 1707 - 1783.

Eulerova funkce φ

Eulerova funkce a Eulerova věta tvoří základ matematické myšlenky šifry RSA.

- Pro kladné číslo N definujeme hodnotu **Eulerovy funkce** $\varphi(N)$ jako:

  $\varphi(N) =$ počet čísel menších než N, která jsou s N **nesoudělná**.

- $\varphi(10) = 4$, protože $1, 3, 7, 9$ jsou všechna čísla mezi 1 a 10 nesoudělná s 10.
- $\varphi(18) = 6$, protože $1, 5, 7, 11, 13, 17$ jsou všechna čísla mezi 1 a 18 nesoudělná s 18.
- Pokud je p **prvočíslo**, pak vždy platí $\varphi(p) = p - 1$.

Eulerova věta

- **O něco těžší:** Pokud jsou p a q **různá prvočísla**, pak vždy platí $\varphi(pq) = (p-1)(q-1)$.

$$\varphi(10) = \varphi(2 \cdot 5) = 1 \cdot 4 = 4 \text{ ...alespoň tady to sedí!}$$

Eulerova věta (1763)

Mějme dvě nesoudělná kladná čísla a, N. Pak platí následující formulka:

$$a^{\varphi(N)} \equiv 1 \pmod{N}.$$

Příklad: Zvolme $N = 10$ a $a = 3$. Pak máme:

$$a^{\varphi(N)} = 3^4 = 9 \cdot 9 = 81 \equiv 1 \pmod{10}.$$

Zpátky k šifrování - kryptosystém RSA

Jak Bob vyrobí privátní a veřejný klíč:

1. Bob zvolí dostatečně náhodným způsobem dvě veliká různá prvočísla p a q. **Příklad:** $p = 3, q = 11$
2. Bob spočítá jejich součin $N = p \cdot q$. **Příklad:** $N = 3 \cdot 11 = 33$
3. Bob spočítá hodnotu Eulerovy funkce $\varphi(N) = (p-1)(q-1)$. **Příklad:** $\varphi(N) = 2 \cdot 10 = 20$
4. Bob si zvolí číslo e v rozmezí $1 < e < \varphi(N)$ takové, aby čísla e a $\varphi(N)$ byla nesoudělná. **Příklad:** $e = 3$
5. Bob nalezne číslo d jako řešení modulární rovnice podobné, jako jsme viděli pár slidů zpátky:

$$e \cdot d \equiv 1 \pmod{\varphi(N)}$$

Příklad: Řešením rovnice $3d \equiv 1 \pmod{20}$ je 7, neboť $3 \cdot 7 = 21 \equiv 1 \pmod{20}$.

Zpátky k šifrování - kryptosystém RSA

Jak Bob vyrobí privátní a veřejný klíč:

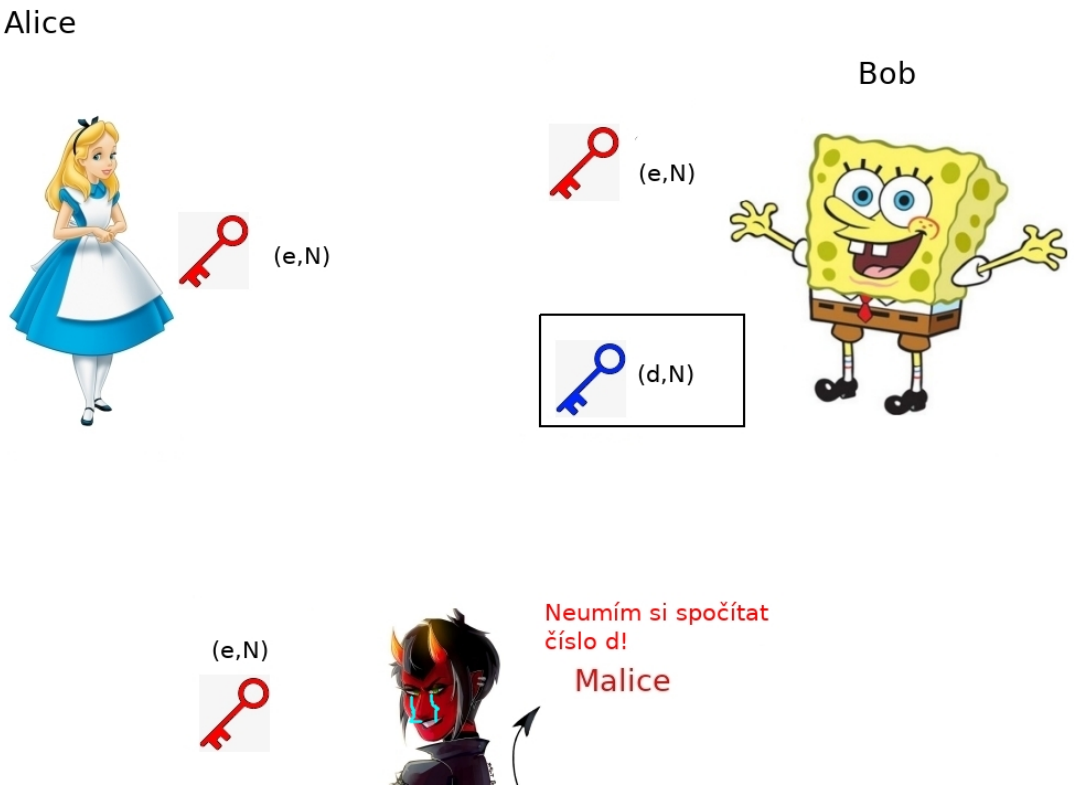
1. Bob zvolí dostatečně náhodným způsobem dvě veliká různá prvočísla p a q. **Příklad:** $p = 3, q = 11$
2. Bob spočítá jejich součin $N = p \cdot q$. **Příklad:** $N = 3 \cdot 11 = 33$
3. Bob spočítá hodnotu Eulerovy funkce $\varphi(N) = (p-1)(q-1)$. **Příklad:** $\varphi(N) = 2 \cdot 10 = 20$
4. Bob si zvolí číslo e v rozmezí $1 < e < \varphi(N)$ takové, aby čísla e a $\varphi(N)$ byla nesoudělná. **Příklad:** $e = 3$
5. Bob nalezne číslo d jako řešení modulární rovnice podobné, jako jsme viděli pár slidů zpátky:

   $$e \cdot d \equiv 1 \pmod{\varphi(N)}$$

   Příklad: Řešením rovnice $3d \equiv 1 \pmod{20}$ je 7, neboť $3 \cdot 7 = 21 \equiv 1 \pmod{20}$.
6. Nakonec Bob zapomene vše kromě čísel N,e,d.

Kryptosystém RSA

Bob si uloží dvojici čísel (d, N) jako svůj **privátní klíč** a zveřejní dvojici čísel (e, N) jako **veřejný klíč**. Zbytek čísel z výpočtu může s klidem zapomenout.



V našem příkladu $(d, N) = (7, 33)$ a $(e, N) = (3, 33)$.

Šifrování a dešifrování v RSA

Zašifrování:

1. Alice nejprve text zprávy přepíše předem dohodlou metodou jako sérii čísel. **Příklad:** **(v praxi se dělá trochu jinak)**
   "AHOJ BOBE"$\Rightarrow$ "$1, 8, 15, 10, 0, 2, 15, 2, 5$"
2. Alice postupně každé číslo m ze zprávy zašifruje pomocí veřejného klíče (e, N) následovně: $m \mapsto m^e \pmod N$.
3. $c = m^e \pmod N$ je pak (jedním číslem) zašifrované zprávy.
   Příklad: Za použití veřejného klíče $(e, N) = (3, 33)$ počítáme:

m	1	8	15	10	0	2	15	2	5
$c = m^3 \pmod{33}$	1	17	9	10	0	8	9	8	26

Šifrování a dešifrování v RSA

Dešifrování:

1. Bob přijme zašifrovanou zprávu sestávající se z čísel c.
   V příkladu: "$1, 17, 9, 10, 0, 8, 9, 8, 26$"
2. Bob postupně každé číslo c ze zprávy dešifruje pomocí privátního klíče (d, N) následovně: $c \mapsto c^d \pmod{N}$.

Proč dostane Bob po dešifrování správný výsledek? Počítejme:

$$c^d = (m^e)^d = m^{ed}$$

Zajímá nás výsledek modulo N.

Šifrování a dešifrování v RSA

$$c^d = (m^e)^d = m^{ed}$$

Zajímá nás výsledek modulo N.

- Vzpomeňme si, že $ed \equiv 1 \pmod{\varphi(N)}$, a tedy existuje číslo $k \geq 0$, že
$$ed = k \cdot \varphi(N) + 1$$

- Dosaďme:
$$m^{ed} = m^{k \cdot \varphi(N)+1} = (m^{\varphi(N)})^k \cdot m.$$

Šifrování a dešifrování v RSA

$$c^d = m^{k\cdot\varphi(N)+1} = \left(m^{\varphi(N)}\right)^k \cdot m.$$

- Podle Eulerovy věty platí: $m^{\varphi(N)} \equiv 1 \pmod{N}$.
- Tedy můžeme zjednodušit:

$$\left(m^{\varphi(N)}\right)^k \cdot m \equiv 1^k \cdot m \equiv m \pmod{N}.$$

Dohromady máme $c^d \equiv m \pmod{N}$, a tedy dešifrováním zašifrovaného textu skutečně získáme původní zprávu.

Šifrování a dešifrování v RSA

- **Velmi pozorný čtenář si mohl všimnout PODVODU!**
- Na předchozím slidu bylo: Podle Eulerovy věty platí: $m^{\varphi(N)} \equiv 1 \pmod{N}$.

Eulerova věta však požaduje, aby čísla m a N byla nesoudělná, vzpomeňme si:

Eulerova věta (1763)

Mějme dvě **nesoudělná** kladná čísla a, N. Pak platí následující formulka:

$$a^{\varphi(N)} \equiv 1 \pmod{N}.$$

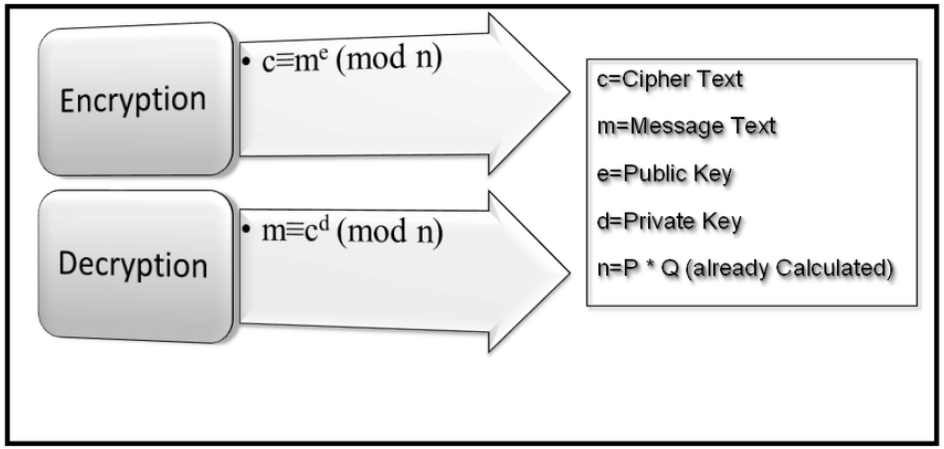
- Můžete si sami nalézt dvojici soudělných čísel a a N, pro které znění Eulerovy věty skutečně neplatí!
- Algoritmus RSA je i přesto v pořádku - případ, kdy m a N mají společného dělitele nezpůsobuje problém z jiného, méně "koncepčního" důvodu, který zde zameteme pod koberec.

Historie asymetrické kryptografie

- Vznik šifry RSA se datuje do roku 1977, avšak nejde o historicky první návrh asymetrické šifry. Krátce před nimi s trochu jiným konceptem přišli Whitfield Diffie a Martin Hellman.
- Dokonce ještě o pár let dříve podobný návrh vynalezl James Ellis. Pracoval však pro britskou tajnou službu a informace o jeho objevu byla zveřejněna až v 90. letech.
- RSA je stále nejpoužívanější asymetrickou šifrou, mimo jiné i proto, že se hodí i pro jiné aplikace, než šifrovaní - např. zmíněné schéma digitálních podpisů.

Bude šifra RSA bezpečná i v budoucnu?

- V současné době neexistuje rigorózní důkaz matematické bezpečnosti žádné asymetrické šifry. Tento problém souvisí s jedním ze sedmi z tzv. Problémů Milénia známým pod zkratkou **P versus NP**.
- Šifra RSA je však ohrožena i méně abstraktním způsobem. Ví se, že dostatečně funkční a výkonný **kvantový počítač** dokáže šifru RSA prolomit. Přesněji řečeno, existuje efektivní **kvantový algoritmus** na faktorizaci čísel na prvočíselný rozklad, tzv. **Shorův algoritmus**. Dostatečně výkonné a spolehlivé kvantové počítače jsou však stále hudbou budoucnosti.
- I tak již existují modernější asymetrické šifry, které by měly být na **kvantové útoky** odolné.



Děkuji za Váš zájem!