Základy kryptologie

Kamil Malinka
malinka@fit.vutbr.cz
Fakulta informačních technologií

Detaily zkoušky

- Během semestru je možno získat maximální počet 100 bodů
  - projekty - 20b.
  - vnitrosemestrální písemka - 30b.
  - závěrečná písemka - 50b.

- Pro získání zápočtu je nutno splnit tyto podmínky:
  - Pro získání zápočtu potřebujete získat minimálně 50% možného bodového zisku z vnitrosemestrální písemky a projektu s podmínkou, že z každého projektu musíte získat minimálně 3 body (maximální zisk z jednoho projektu je 6 a 7b.). Bez získání zápočtu Vám nebude umožněno absolvovat zkoušku.

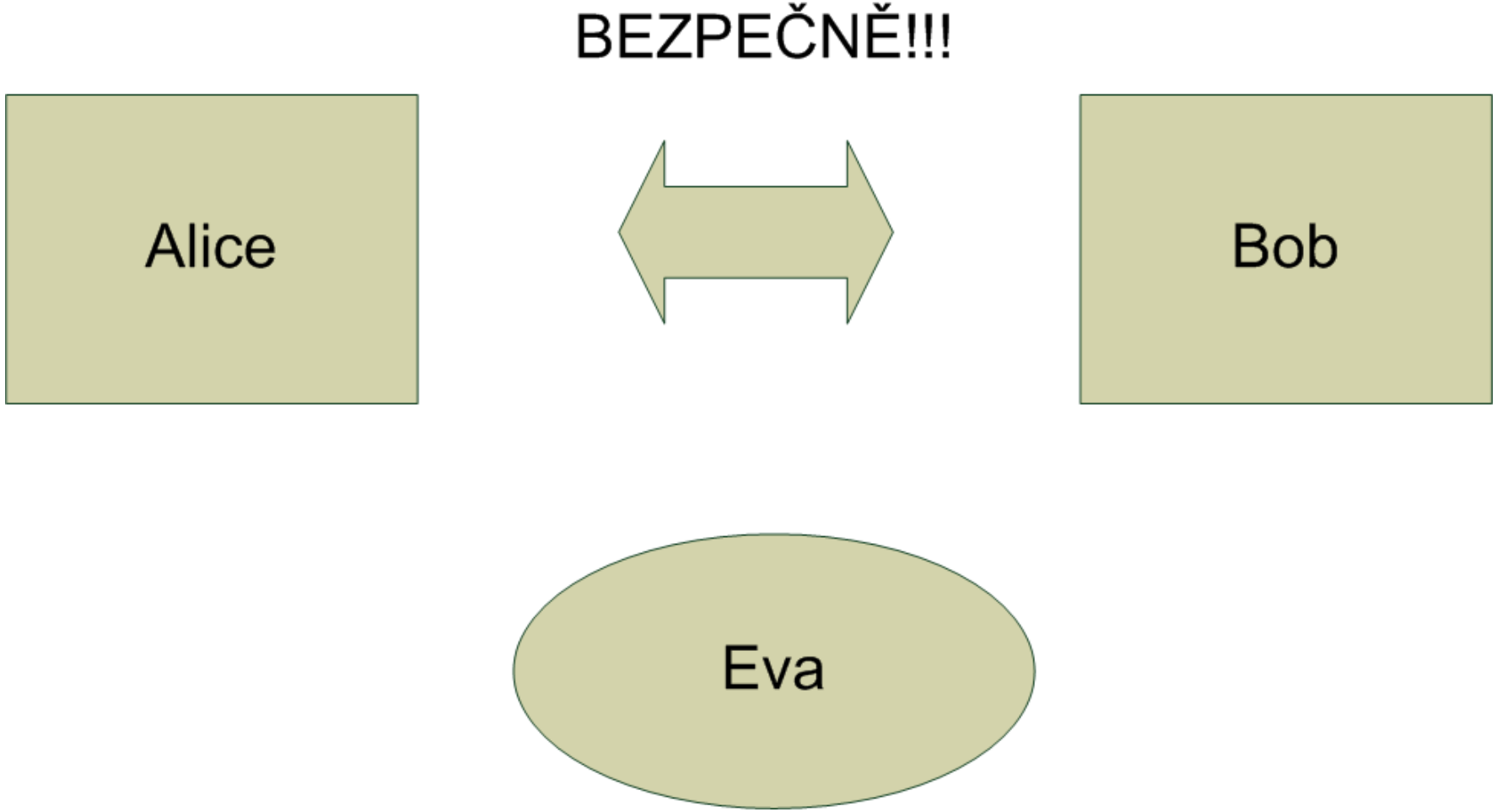
- Informace k předmětu
  - http://securityfit.cz/kib/
- Termíny cvičení viz aktuality

O čem je kryptografie?

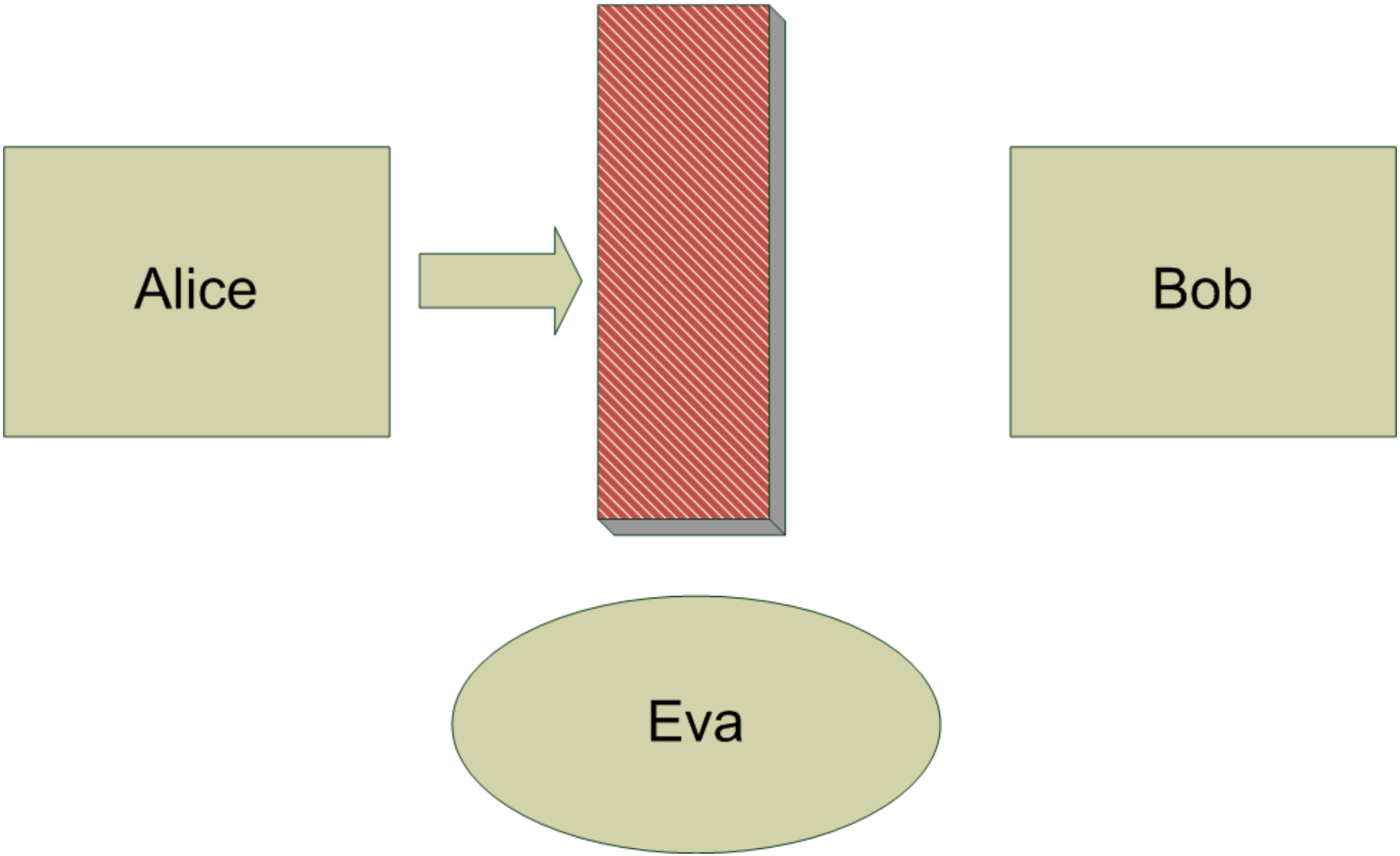
- mějme dvě komunikační strany – A (Alice), B (Bob)
- kryptografie umožňuje bezpečně komunikovat
- protokol
  - distribuovaný algoritmus
  - formální popis pravidel pro přenos dat…
  - cíle vs. hrozby – je nutno brát v potaz útočníka

- možnosti útočníka?

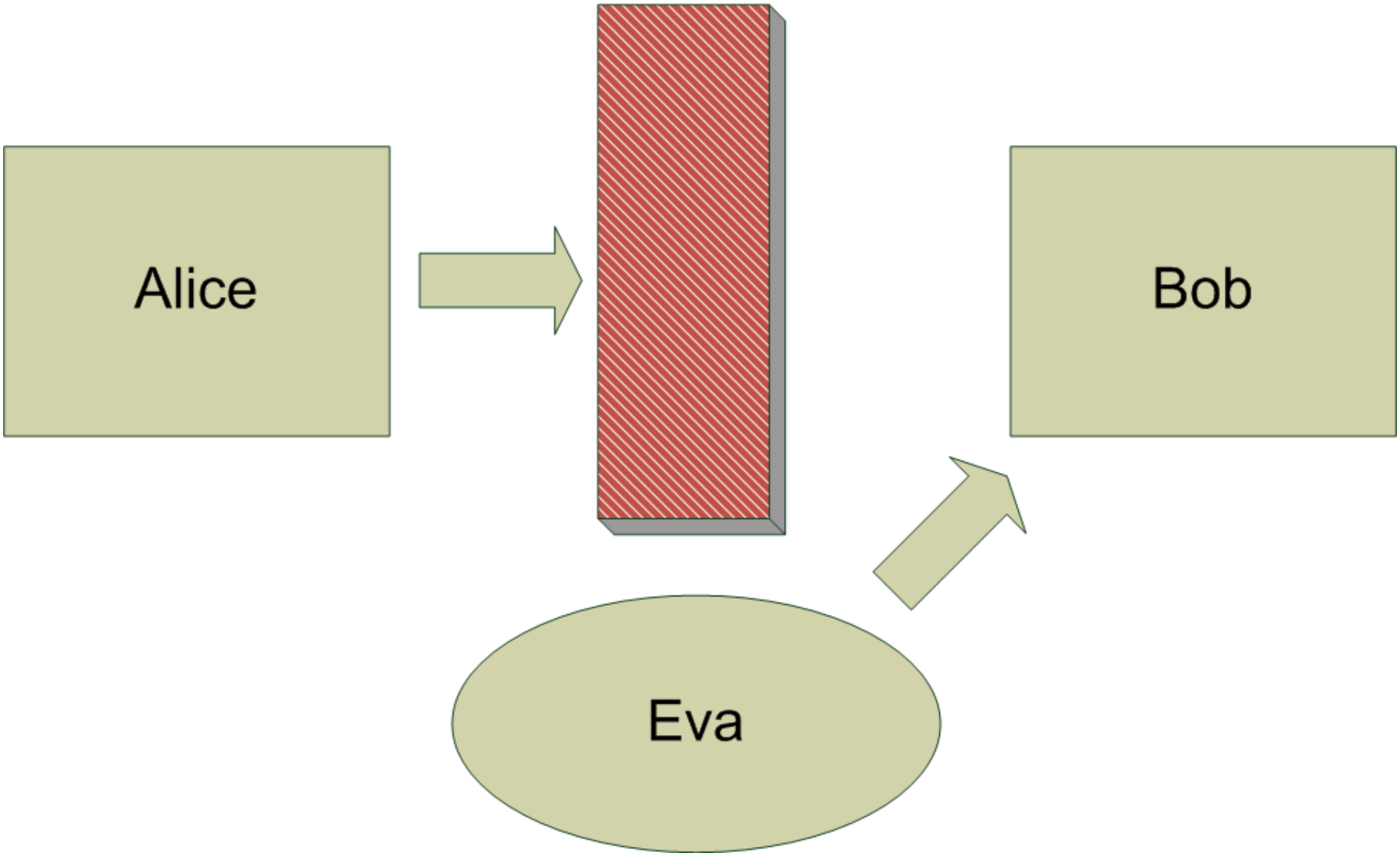
O čem je kryptografie?



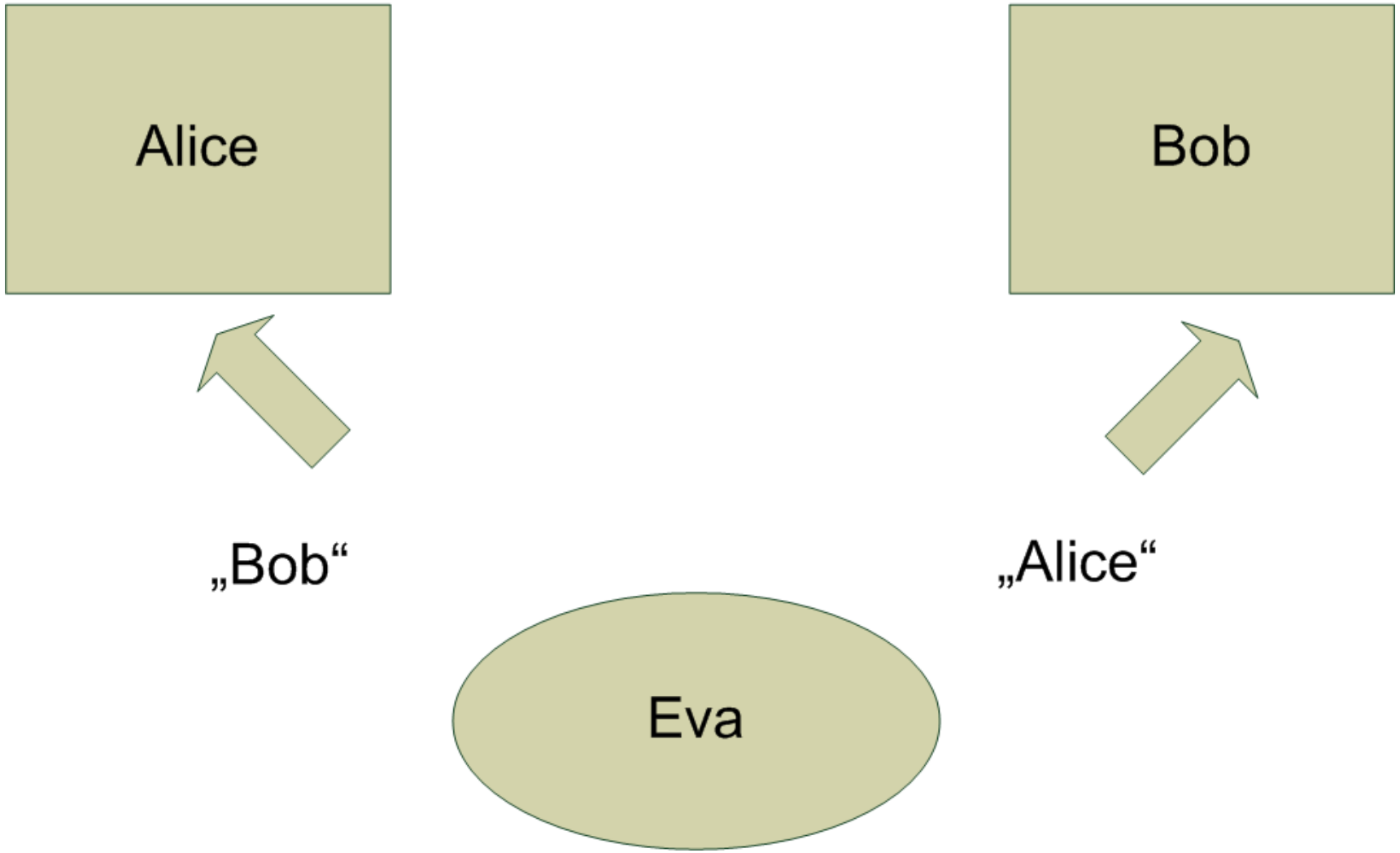
Zabránění doručení

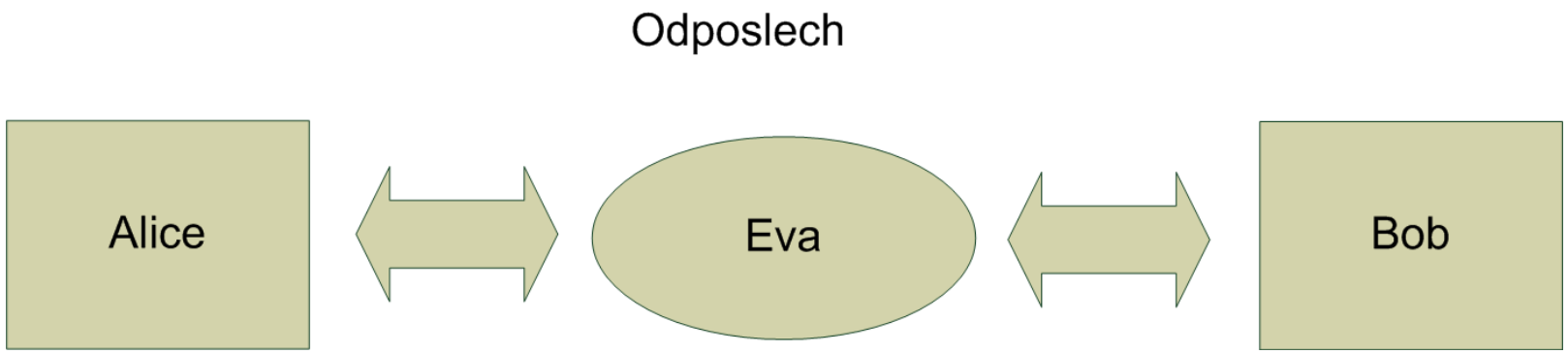


Vložení jiné zprávy



Předstírání jiné identity




Odposlech
Alice
Eva
Bob

Bezpečnost je tak silná jako nejslabší článek!

Bezpečnostní vlastnosti - cíle

- důvěrnost
  - utajení obsahu zprávy
- integrita
  - nemožnost změny zprávy
- autentizace
  - Eva se nemůže vydávat za Alici
- nepopiratelnost
  - lze ověřit vše co Alice a Bob udělali

- Otázka:
  - Jaký je rozdíl mezi autentizací a nepopiratelností?

Symetrická kryptografie

- Alice a Bob sdílejí klíč, který Eva nezná
- klíč – náhodný řetězec *k* bitů
  - jaký počet bitů je dostatečný?
    - počet lidí 5×10^9
    - hmotnost země 5.98×10^{27} grams
    - viditelný vesmír – 1.7×10^{77} atomů
  - délka se odvíjí od algoritmu a odolnosti proti útoku hrubou silou
  - v současnosti – 512 b. je považováno za bezpečné

Symetrická kryptografie

- Rozvoj ARPANETu -> Internet -> DES/Lucifer
- DES – moderní implementace původních principů
- DES -> 3-DES -> AES

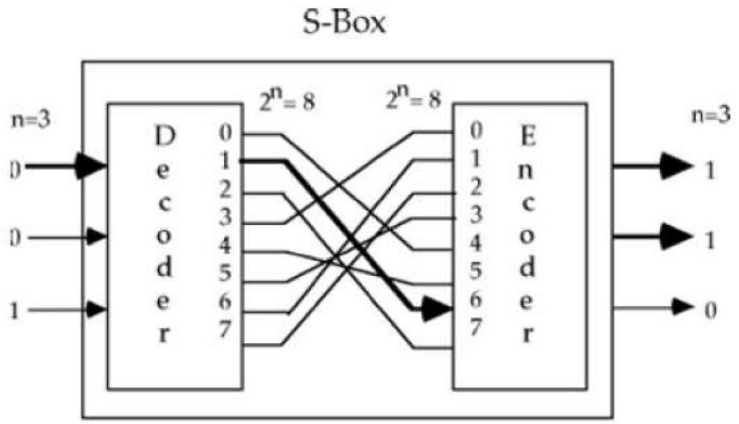
- problém s distribucí klíčů…

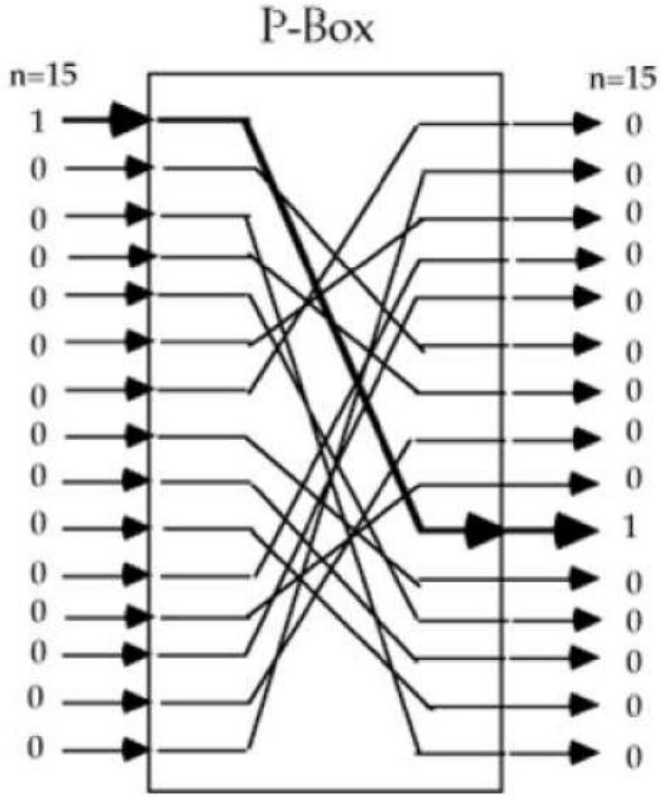
DES

- Základní vlastnosti algoritmu
  - Lavinovitost – změna jednoho vstupního bitu má vliv na přibližně polovinu bitů na výstupu
  - Úplnost – každý výstupní bit je funkcí všech vstupních bitů
  - Neexistence korelace – statisticky ověřeno, není vztah mezi otevřeným a šifrovaným textem, a dále mezi šif. textem a klíčem

Stavební bloky DES

- S-Box
- P-Box





Blokové schéma DES

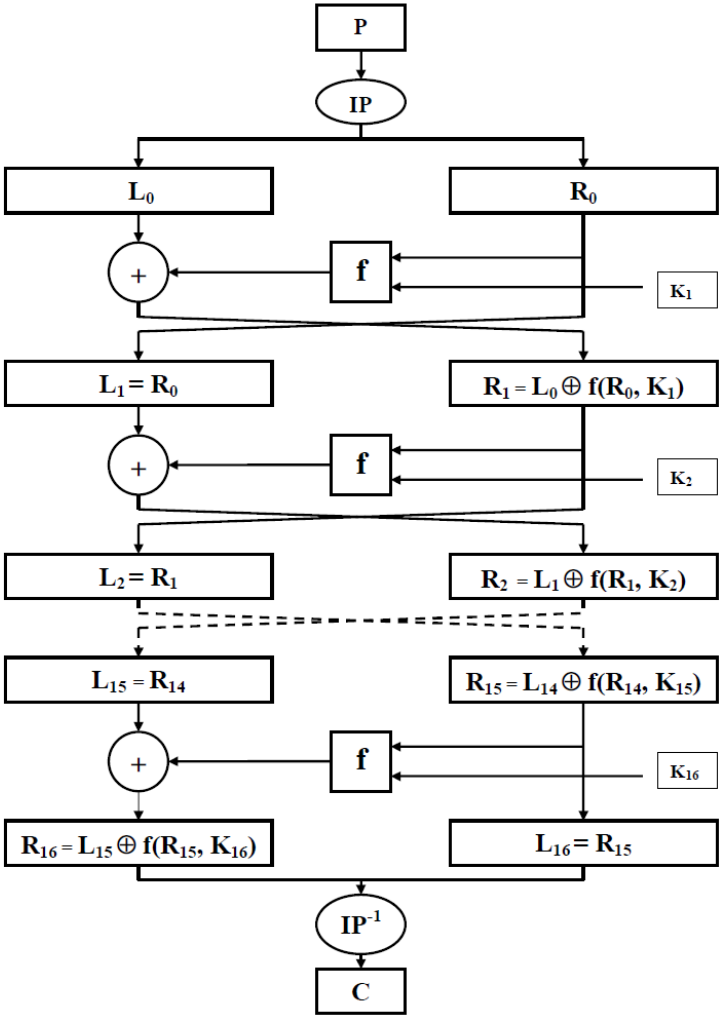
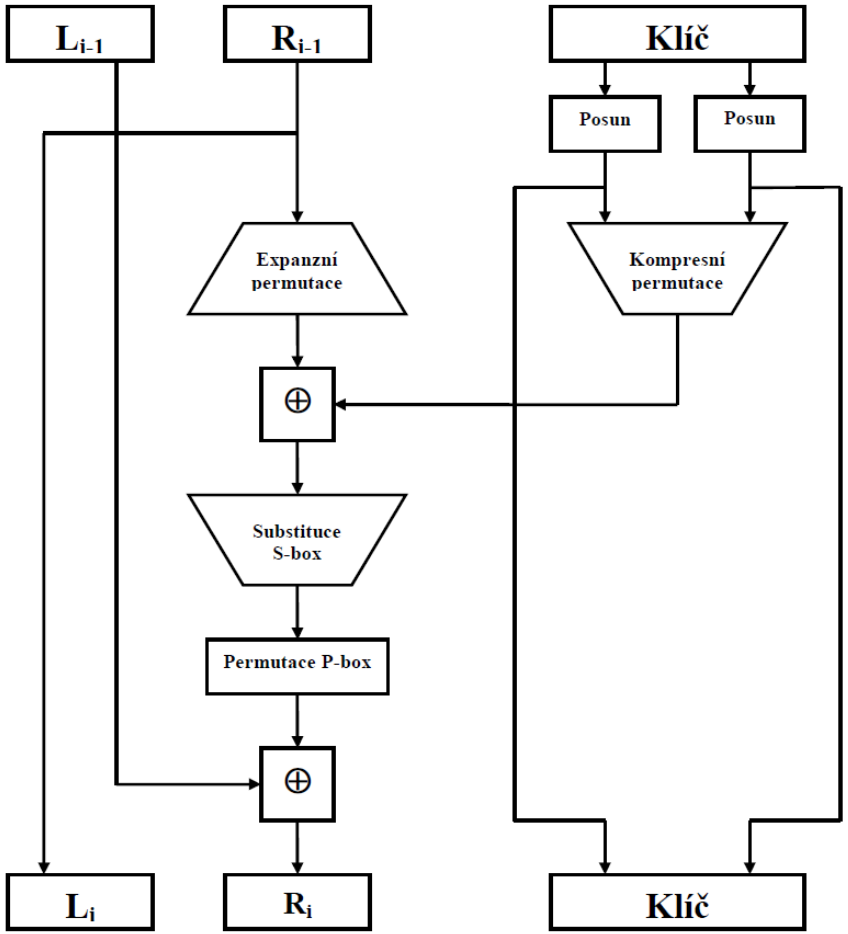


Schéma jedné rundy DES



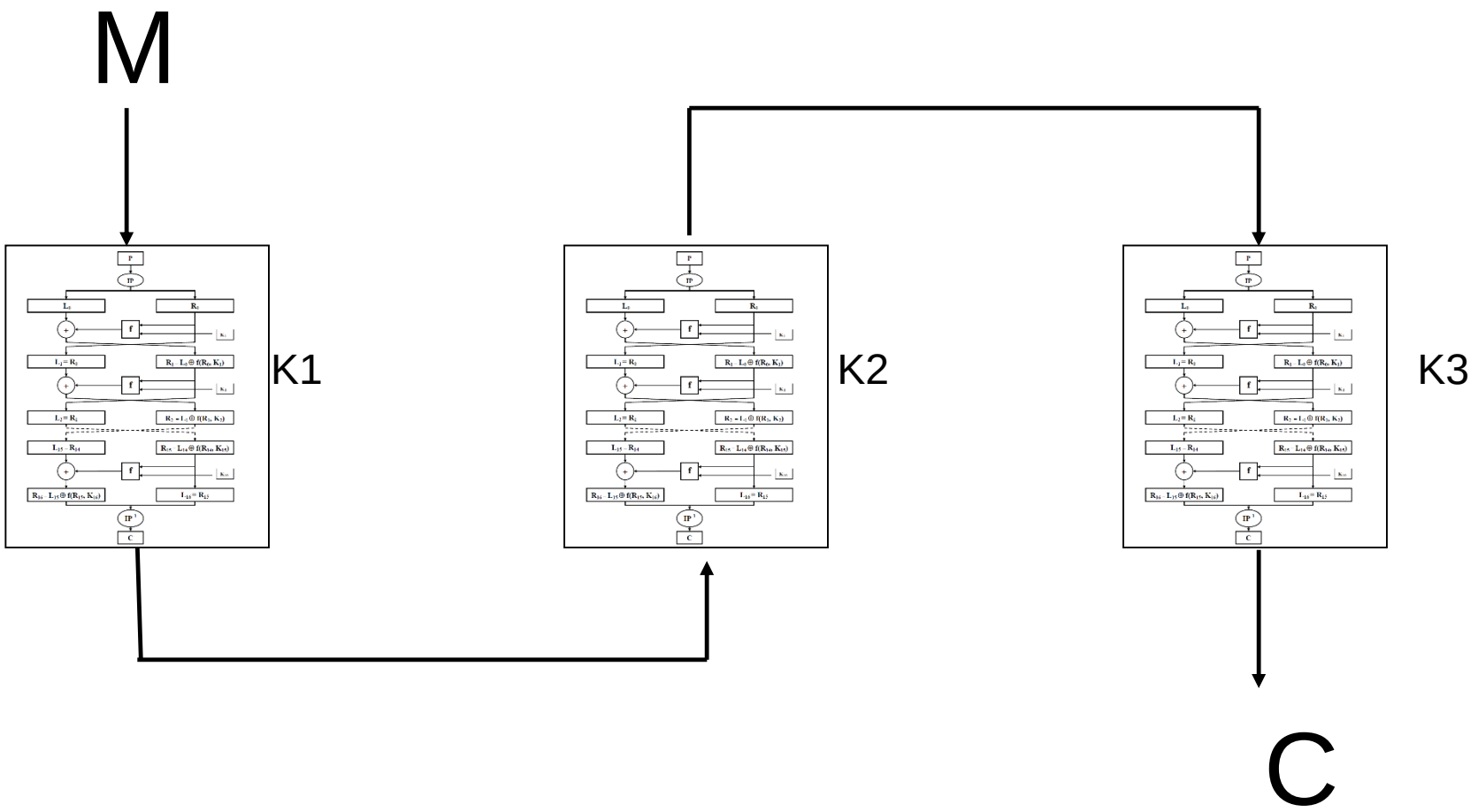
Triple DES

- Varianty EEE, EDE …

- C=E(K3,D(K2,E(K1,M)))
- M=D(K1,E(K2,D(K3,C)))

- Varianty použití klíčů:
  - Klíče jsou nezávislé
  - K1 a K2 jsou nezávislé, K1 = K3
  - K1 = K2 = K3

Varianta EDE



AES

- Viz DEMO Rijndael

Problém distribuce klíčů

- telefon? kurýr? – logistické problémy…
- v historii řešeno silou peněz

- první náznak řešení distribuce klíčů – Diffie - Hellman
- založen na modulární matematice (těleso zbytkových tříd)

Diffie - Hellman

- protokol pro ustanovení klíčů
- prvočíslo p, generátor α tvoří multiplikativní grupu Zp
- Alice zvolí tajemství *x*, Bob zvolí tajemství *y*
- $A \rightarrow B: \alpha^x \bmod p$
- $A \leftarrow B: \alpha^y \bmod p$

- Alice a Bob sdílí $K=\alpha^{xy} \bmod p$

Asymetrická kryptografie

- Alice a Bob mají své vlastní veřejné a soukromé klíče
- Vyměňují se pouze veřejné klíče – Eva je může znát
- velikosti klíčů – 2048 b. – 4096 b.
- algebraická záležitost
- problém:
  - jak spojit klíče s uživateli?
  - vždy zajištění integrity – o kolik je to jednodušší než zajistit nepopiratelnost?

Symetrická vs. asymetrická kryptografie

- problém s distribucí klíčů
- výpočetní složitost
- revokace klíčů

- v současnosti oblíbený princip:
  - asymetrická kryptografie pro ustanovení klíče, další šifrovaní pomocí symetrické (rychlost)

Protokol Needham-Schroeder

- Protokol pro obousměrnou autentizaci

  $A \rightarrow B$: A, B, $(N_A, A)PK_B$

  $B \rightarrow A$: B, A, $(N_A, N_B)PK_A$

  $A \rightarrow B$: A, B, $(N_B)PK_B$

- útok vypadá následovně:

  $A \rightarrow I$: A, I, $(N_A, A)PK_I$

  $I(A) \rightarrow B$: A, B, $(N_A, A)PK_B$

  $B \rightarrow I(A)$: B, A, $(N_A, N_B)PK_A$

  $I \rightarrow A$: I, A, $(N_A, N_B)PK_A$

  $A \rightarrow I$: A, I, $(N_B)PK_I$

  $I(A) \rightarrow B$: A, B, $(N_B)PK_B$

- problém – 2 a 3 krok neobsahují identifikátory

Teorie pravděpodobnosti

- Narozeninový problém
  - v nádobě je *m* míčků očíslovaných 1 až m, n míčků je taženo s opakováním
  - pravděpodobnost aspoň jedné koincidence (stejná kulička vybrána dvakrát)

    $P(m, n)=1-m!/n!*m^n$

    - if $n=\sqrt{m}$, $m\rightarrow\infty$, $P(m,n) \approx 1-\exp(-n^2/(2m))$
    - $P(m, n) = 1-\exp(-m/(2m)) = \sqrt{((\pi m)/2)} = 1.25\ \sqrt{m}$

Narozeninový problém

- jiný příklad:
  - v jedné místnosti x lidí, kolik lidí má narozeniny ve stejný den?
  - reálná čísla: 40 lidí v místnosti vs. 365 dní
  - očekávání: cca 11%, 100% if 365 lidí
  - realita: pnost. cca 90%, 60 lidí – 99%
  - v reálném světě výsledky ovlivňují populační exploze, oblíbené dny, výpadky elektřiny apod.
- Problém v kryptografii:
  - např. využívání pro vyhledávání kolizí v hashích
  - tzv. Birthday attack – snížení velikosti prostoru klíčů
    - př. 64 bitový hash, aprox. 1,8 x 10^{19} výstupů, best case: 5,1 x 10^{9} pokusů k nalezení kolize
    - nutnost volby klíčů tak, aby byly 2x delší než pro obyčejný brute-force attack

Entropie

- střední hodnota informace jednoho kódového znaku (přirozený jazyk má špatnou entropii – viz. frekvenční tabulka)
- pravděpodobnost výskytu určité proměnné
  - muž/žena 50%
- $X = \{x_1, x_2, \ldots, x_n\}$ je náhodná proměnná a $P(X=x_i)=p_i$, $\sum p_i=1$
- Entropie je definována jako $H(X)= -\sum p_i \lg p_i =$ $= \sum p_i \log (1/p_i)$, dle konvence ($p_i \log (1/p_i)=0$ iff $p_i=0$)
- Vlastnosti
  - $H(X) \in \langle 0, \lg n \rangle$
  - $H(X)=0 \Leftrightarrow$ existuje p_i takové, že $p_i=1$
  - $H(X)=\lg n \Leftrightarrow p_i=1/n$
- nejvyšší entropie – rovnoměrné rozložení pravděpodobností -> důležitost náhodných generátorů (je málo zdrojů vysoké entropie)
- Při šifrování chceme zavést co nejvyšší náhodnost

Složitost

- Algoritmus je dobře definovaná výpočetní procedura
  - má vstupní proměnné a vždy se zastaví a vrátí výstup
- Velikost dat (vstup, výstup) je minimální počet bitů potřebných pro reprezentaci hodnot příslušných dat
- složitost ***časová*** vs. ***prostorová***
- je možno spočítat průměrnou hodnotu a nejhorší možný případ
- důvody: asymptotické ohraničení dob běhu, hranice pro nekonečné množství hodnot

Složitost – asymptotické hranice

asymptotická horní hranice – $f(n) = O(g(n))$
pokud existuje pozitivní konstanta c a přirozené číslo n_0 takové, že
$0 \leq f(n) \leq c\, g(n)$, pro každé $n \geq n_0$

asymptotická spodní hranice – $f(n) = \Omega(g(n))$ …
$0 \leq c\, g(n) \leq f(n)$, pro každé $n \geq n_0$

asymptotic tight bound – $f(n) = \Theta(g(n))$ constants c_1, c_2, … $c_1\, g(n) \leq f(n) \leq c_2\, g(n)$, for all $n \geq n_0$

o-notace – *f(n) = o(g(n)) if for **any** positive constant c …*
$0 \leq f(n) < c\, g(n)$, for all $n \geq n_0$

Třídy složitosti

- algoritmus s polynomiální časovou složitostí -worst-case running time $O(n^k)$, n – velikost vstupu
- algorimus, který nelze ohraničit $O(n^k)$ je nazýván algoritmem s exponenciální časovou složitostí
- $O(n^{\ln n \ln n})$ vs. $O(n^{100})$?
- **třída složitosti P** – polynomiální časová složitost
- **třída složitosti NP** – nedeterministický polynomiální problém
- **NP úplné problémy**

- problém obchodního cestujícího

Faktorizace celých čísel

- jedná se o NP problém
- $n=p_1^{e1}p_2^{e2}\ldots p_k^{ek}$, p_i jsou prvočísla
- jak najít tato prvočísla?

RSA problem

- dána pozitivní přirozená čísla $n = p.q$, e takové, že g$cd(e, (p-1)(q-1)) = 1$
- zlomení šifry znamená, že pro celé číslo c najdeme m takové, že $m^e \equiv c \ (mod\ n)$
- rozklad 129 místného čísla => 64x65 místné…

RSA příklad

- p = 61 (první prvočíslo)
- q = 53 (druhé prvočíslo)
- $n = pq$ = 3233 (modul, veřejný)
- e = 17 (veřejný, šifrovací exponent)
- d = 2753 (soukromý, dešifrovací exponent)

- Pro zašifrování zprávy 123 probíhá výpočet:
  - šifruj(123) = 123^{17} mod 3233 = 855 Pro dešifrování pak:
  - dešifruj(855) = 855^{2753} mod 3233 = 123

Klasická kryptografie

- substituce
  - monoalfabetická vs. polyalfabetická
- permutace

Moderní kryptografie

- používáním počítačů jsme schopni provádět obrovské množství operací
- šifry stále používají dva principy
  - S-boxy a permutace (pevně definované v návrhu šify)
- **Kerckhoffův princip**
  - kryptosystém by měl být bezpečný i v případě, kdy všechno ohledně konstrukce a fungování systému je veřejně známé – **kromě klíče**
  - porušen např. během 2.světové války
  - security through obscurity

One-time pad

- šifra s bezpečností dokázanou na úrovni teorie informace
- Alice a Bob se dohodnou na dlouhé náhodné sekvenci *a*
- každá zpráva je xorována s nikdy nepoužitou částí sekvence
- důkaz je založen na násl. předpokladech:
  - *a* ani žádná její část nebude nikdy použita dvakrát
  - *a* je opravdu náhodná (porušeno během 2.svět.války)
- zisk částečné informace M1 xor M2
  - přirozený jazyk obsahuje velkou redundanci

Generátory náhodných čísel

- vždy potřebujeme nějaké náhodný bity
- hardware vs. software RNG
- generátory pseudonáhodných čísel
- libovolná konečná sekvence je náhodná, jestliže neexistuje žádná kratší sekvence, která by ji plně popisovala v nějakém jednoznačném matematickém zápisu
- analýza RNG – hodnotíme bezpečnost jejich návrhu a principy jejich fungování
- statistické testování
- jak vytvořit z řídkého zdroje dobrou náhodnou sekvenci?
  - xor
  - van Neumannova metoda (middle-square method)
    - 1111 -> 01234321 -> 05489649 …

Kryptografická primitiva

- všechny algoritmy je možno vytvořit z několika málo základních funkcí
- generátory náhodných čísel
- jednocestné funkce
- náhodné funkce se zadními vrátky

- proudové šifry
- blokové šifry
- schémata veřejného klíče

- nutno různé přístupy pro různé scénáře
  - multimedia vs. bloky..

Kryptografické útoky

- pasivní
  - užívány pro schémata veřejného klíče,útočník zná veřejný klíč
  - útoky nezávislé na klíči – výběr textu nezávislého na klíči
  - útoky závislé na klíči
- aktivní
  - útočník je aktivním účastníkem protokolu
  - chosen plaintext
  - chosen ciphertext