Steganografie

Intro 1

Osnova

Co je to steganografie? 1

Formátování digitálního obrazu. 1

Akvizice digitálního obrazu. 1

Útoky na LSB embedding v různých formátech. 3

Pokročilejší metody (±1 embeddig, statistická restorace, F5) 1

Maticové vkládání 2

Efektivita maticového vkládání 1

Wet paper codes (psaní na mokrý papír) 1

Vkládání pomocí Viterbiho algoritmu 1

Historie

Steganos = přikrytý kryptos = schovaný, tajný
grafia = psaní

Hérodotos - příběh o otrokovi se zprávou vytetovanou na hlavě
- použití dřevěné desky pokryté voskem - zpráva napsaná na dřevě pod voskem.

Ukrývání zpráv v textu - např. první písmena každé věty.
např. Giovanni Boccaccio takto zakódoval celé tři sonety do počátečních písmen v básních
- složitější schéma: písmena skrývající zprávu jsou určena nějakým klíčem.
- formátování textu - kurzíva, vzdálenost mezi řádky
posun řádku o $\frac{1}{300}$ in $\doteq \frac{1}{10}$ mm není pozorovatelný okem a přitom se zachová při kopírování.

Neviditelný inkoust: - roztok cukru, octu nebo mléka a zahřátí nad svíčkou
- inkoust viditelný pod ultrafialovým světlem.

Mrkání: 1966 Americký zajatec poskytl rozhovor na kamenu morseovou mrkáním sdělil, že ho vietnamci mučí.

Bankovky: chráněny proti kopírování pomocí EURion (spíš vodoznak)
Barevné laserové tiskárny ukládají žluté tečky s informací o tiskárně (Machine Identification Code)

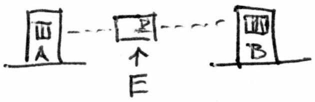
Mikrotečky – mikroskopické fotografie textu o ~~rozměr~~ průměru ~1 mm.

- Používali je Němci během obou světových válek. Mikrotečka se doložila mezi vrstvy papíru v pohlednici.
- Identifikace kradených automobilů – mikrotečky v laku. (spíše vodoznak)

Skutečný rozmach oboru nastal v posledních 15 letech. 1998 – (díky Internetu)

Co je steganografie a co to není

Problém dvou vězňů



169/1999 Sb. § 17 odst. 2
Vězeňská služba je oprávněna provádět kontrolu korespondence, přitom je oprávněna seznámit se s obsahem zasílaných písemností. Zakládá podezření, že je připravován nebo páchán tr. čin ⇒ zadržení

Vězni A a B, strážce E.

- Vězni jsou v oddělených celách. Chtějí plánovat společný útěk.
- Smějí komunikovat, ale komunikace je monitorovaná.
- ~~[illegible] odhalí komunikaci~~
- Když strážce odhalí utajovanou komunikaci, kanál přeruší a vězně potrestá.
- Předpokládá se, že strážce zná steganografický algoritmus, ale nezná klíč. (Např. výběr písmen v textu, ale neví, která písmena jsou vybírána.)

- Pasivní strážce: monitoruje komunikaci a analyzuje zprávy ← (cílem je odhalit, že dochází ke skr. komunikaci.)
- Aktivní strážce: modifikuje komunikaci, aby zničil potenciální ukrytý obsah zpráv.
  - např. změna velikosti obrázku, ořezání okrajů, ztrátová komprese.
  - Změnit pořadí slov nebo nahradit synonymem. Američtí telegrafisté za 1. sv. v.
- Zákeřný strážce: Vkládá zprávy do utajené komunikace. Vydává se za protistranu.

! Stegosystém je považován za prolomený pokud lze odhalit, že dochází ke skryté komunikaci. Zjištění obsahu komunikace je druhořadé.

K čemu je steganografie dobrá? Neonacisti si posílají dopisi z vězení.

- Ochrana před represivními režimy, které zakazují šifrování. (Bitové a hesla)
- Obecně nechceme-li upoutat pozornost.
- Máme-li v cestě aktivní firewall, který blokuje určité protokoly, pak můžeme zakázaný protokol tunelovat přes povolený protokol pomocí stego.

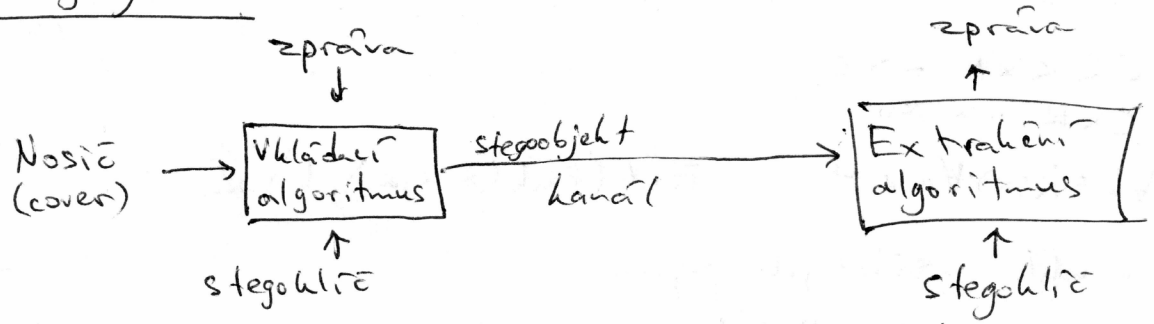
Vodoznak

Př: Internetoví prodejci hudby vkládají do MP3 vodoznak, kterým lze identifikovat kupce. V případě nelegálního šíření pak lze odhalit původce. Vodoznak tedy musí být těžko odstranitelný.

Rozdíly:	Steganografie	Vodoznak
Velikost zprávy	Velká (dopis)	malá (stačí id. číslo, nebo samotná přítomnost vodoznaku = 1bit)
Robustnost	nevyžadujeme u pasivního strážce	vyžadujeme
Nedetekovatelnost	vyžadujeme	nevyžadujeme.
Nosič	slouží k odlákání pozornosti	je významný sám o sobě a vodoznak ho pouze doplňuje.

- Robustnost (neodstranitelnost) a nedetekovatelnost jsou zpravidla protichůdné požadavky.

Stegosystém



cílem je nedetekovatelnost tj. neschopnost odlišit stegoobjekt od nosiče (uvažujeme pasivního strážce)

1) Steganografie volbou nosiče
- Máme databázi obrázků a každý obrázek má svůj tajný význam (stegoklíč).
  Např obrázek psa = začátek útoku
  obrázek kočky = ústup.
  = sada interpretačních pravidel.
- Varianta: vybereme obrázek, ~~který tedy se zhashuje~~
  jehož hash má požadovanou hodnotu čili hash = zpráva. (ještě lépe MAC)
  výhoda: obrázek je zcela přirozený
  nevýhoda: exponenciální složitost v délce zprávy.
  ? problém: při opakovaném použití (se stejným klíčem resp. heslem) může posloupnost odesílaných obrázků vykazovat (teoreticky) detekovatelnou vlastnost.

2) Steganografie tvorbou nosiče

- Např. naaranžujeme nějakou scénu, kterou vyfotografujeme
  mísa s ovocem – rozmístění ovoce může skrývat zprávu.
- Texty: vygenerujeme zprávu, která vypadá jako spam. Ten zpravidla
  používá zvláštní slova a značené výrazy, kterými lze
  kódovat zprávu.
  www.spammimic.com

3) Steganografie modifikací nosiče

C množina nosičů
$x \in C$ nosič
$K(x)$ množina klíčů pro daný nosič
$M(x)$ množina zpráv pro daný nosič
(Různé nosiče z C mohou mít různou kapacitu, proto je množina zpráv závislá na volbě nosiče.)

$Emb: C \times K \times M \to C$
$Ext: C \times K \to M$
$\forall x \in C \; \forall k \in K(x) \; \forall m \in M(x) \quad Ext(Emb(x,k,m),k) = m.$

Zavedeme několik užitečných pojmů:

- Kapacita nosiče: $\log_2 |M(x)|$, $x \in C$. (měříme v bitech)
- Relativní kapacita nosiče: $\frac{\log_2 |M(x)|}{n}$, $x \in C$, kde n je počet prvků x.
  v případě rastrového obrázku např. počet pixelů.
  (měříme v bpp bitech na pixel).
- Distorze $d: C \times C \to [0, \infty)$ (x nosič, y stegoobjekt.)
  $d(\bar{x}, \bar{y}) = \sum_{i=1}^{n} (x_i - y_i)^2$
- Efektivita vkládání $e = \frac{\underset{x}{E}\left[\log_2 |M(x)|\right]}{\underset{x,k,m}{E}\left[d(x,y)\right]}$ ← průměrná kapacita, kde $y = Emb(x,k,m)$ ← průměrná distorze

Je to tedy počet vložených
bitů na jednotku distorze – snažíme se maximalizovat.