STEGANOGRAFIE S VYUŽITÍM NEURONOVÝCH SÍTÍ

Robert Jarušek
Katedra informatiky a počítačů, Ostravská univerzita v Ostravě, 30. dubna 22, 70103 Ostrava, email: robert.jarusek@osu.cz

Abstrakt

Pojem steganografie má svůj původ v řečtině, ve slovech stegos „skrytý“ a grafia „psaní“, čili „skryté psaní“ („cover writing“) [1]. Steganografie je metoda skrývání osobních nebo citlivých informací v něčem, co se na první pohled nejeví jako neobvyklé. Cílem steganografie je tedy ukrýt zprávu (informaci, data) tam, kde by ji nikdo nečekal, a zároveň tak, aby její přítomnost nebyla detekována. Zaměření tohoto příspěvku je prezentace výsledků aplikace vícevrstvých neuronových sítí v oblasti steganografie.

Klíčová slova: *steganografie, neuronové sítě, vodoznaky*

Úvod

Krycí médium *(cover medium)* je soubor, do kterého ukrýváme zprávu *(embedded message)*. Někdy se v procesu ukrývání používá i takzvaný *stego-klíč (stego-key)*, který může sloužit k vyššímu zabezpečení a jeho znalost je nutná pro proces extrakce souboru z krycího média. Výsledkem procesu ukrytí zprávy je *stego-médium*. V literatuře se lze často setkat i s jiným označením [2]:

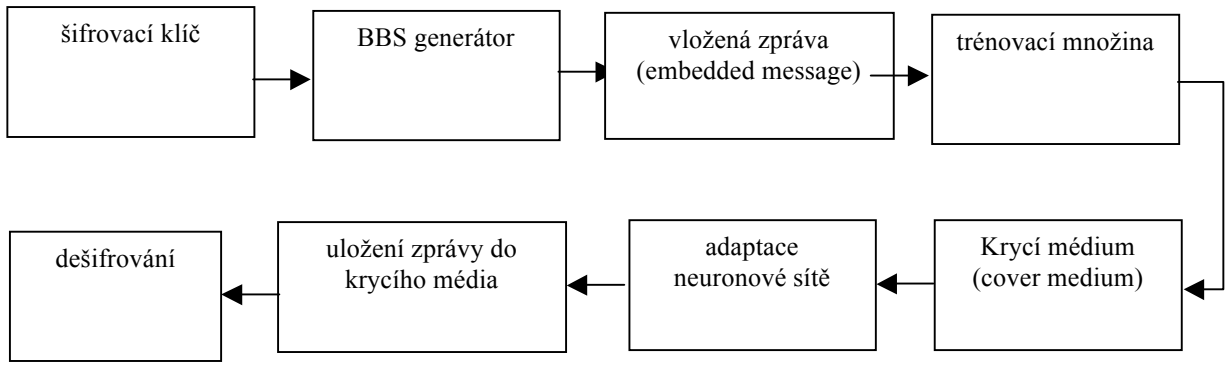
- cover medium (carrier medium, covertext)
- embedded message (hidden message, embedded data)

Obecnou rovnici stegano grafického procesu bychom mohli zapsat následujícím způsobem:

cover medium + embedded message [+stegokey] = stego-medium

Steganografický proces na bázi neuronových sítí

Návrh steganografického systému je uveden na obrázku 1.



Obrázek 1. Model steganografického systému

Práci navrhovaného steganografického systému si vysvětlíme při řešení konkrétního problému. Během naší experimentální studie jsme použili následující nastavení:

- šifrování klíč byl řetězec "123",
- "vložená zpráva" byla "Mendel",
- BBS [3] generátor pracuje s následujícími parametry: p = 1017 a q = 1907,
- Backpropagation pracuje s následujícími parametry: aktivační funkce byla binární sigmoida s parametrem strmosti $\lambda = 1$, koeficient učení $\alpha = 0{,}1$, parametr "momentum" byl nastaven na 0.

Krycí médium představuje digitální obrázek, který je uložen v režimu šedé škály. Dále určíme počet bitů, které budou representovat jednotlivé pixely, například v 8-bitovém barevném režimu, používá barevný monitor 8 bitů pro každý pixel, který umožňuje zobrazování 256 různých odstínů šedé barvy.

Použitím šifrovacího klíče, který je reprezentován řetězcem „123", je vygenerována série náhodných čísel x_i podle vzorce (1) za použití BBS se specifickou inicializaci proměnných p=1017 a q=1907.

$$\begin{aligned} x_{n+1} &= x_n^2 \bmod (M) \\ x_0 &= 123 \\ M &= 1017 * 1907 \end{aligned} \qquad (1)$$

S použitím vzorce (1) bylo dále každých 8 po sobě jdoucích hodnot normalizováno na intervalu $\langle 0,1 \rangle$. Tato série pak reprezentuje vstupní vektor trénovací množiny pro jeden znak vložené zprávy a výstupní vektor pak tvoří reprezentace tohoto znaku podle ASCII. Proto trénovací množina každé neuronové sítě zahrnuje pouze jeden vzor, tj. každá neuronová síť se učí pouze na jeden znak vloženého zprávy. V naší experimentální studii jsme použili šest vícevrstvých neuronových sítí s topologií 8 - 8 - 8. Každý jednotlivý řádek v tabulce 1 představuje jeden vzor trénovací množiny.

Adaptace stegosystému

Adaptace prezentovaného stegosystému probíhá pomocí modifikovaného adaptačního pravidla backpropagation [4]. Váhové hodnoty každé použité neuronové sítě jsou extrahovány z jasové informace pixelu v černobílém obrázku, které jsou v tomto případě vybírány sekvenčně (jiné přístupy budou předmětem dalšího výzkumu). Pro každou váhovou hodnotu je použito n pixelů. Tento počet je určen před začátkem výpočtu. U každého pixelu budeme měnit pouze poslední 2 významné bity (tj. nejméně významné bity). Pokud bychom pro váhu použili jen jeden pixel, pak by nabývala maximálně 4 hodnoty (00, 10, 01, 11), což by pro naučení neuronové sítě bylo nedostatečné.

V rámci této experimentální studie jsme nastavili n = 100, tj. každá váha byla reprezentována 100 pixely (řetězení po řádcích použitého krycího média – obrázku). Příslušná váhová hodnota w_{ij} pak byla vypočítána jako průměrná hodnota z jasové informace všech 100 příslušných pixelů x_i dle vztahu (2). Pro každou neuronovou síť s topologií 8 – 8 – 8 tak použijeme 1440 pixelů obrázku. Pořadí pixelů příslušející jednotlivým váhovým hodnotám je pevně stanoveno před začátkem výpočtu a v jeho průběhu se nemění.

$$w_{ij} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{2}$$

Každá váhová hodnota $\bar{w}_{ij}$ je následně normalizována samostatně do intervalu $\langle -1,1 \rangle$ podle vzorce (3).

$$\bar{w}_{ij} = 2 \cdot \left(\frac{w_{ij}}{255} \right) - 1 \tag{3}$$

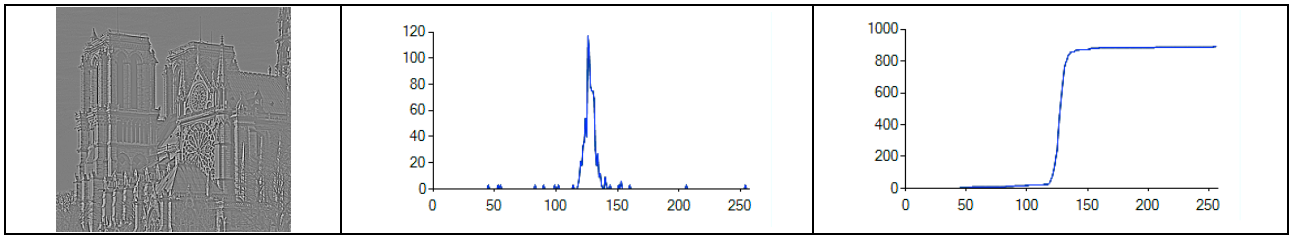
V průběhu adaptace metodou backpropagation jsou přírůstky vah omezeny skokovou funkcí odpovídající bitovému skoku ± 1 u každého pixelu. Tyto přírůstky jsou rozpočítány na jednotlivé pixely, tak aby se vždy změnily nanejvýš 2 jejich poslední nejméně významné bity. Proto je nutné určit samostatně pro každou váhovou hodnotu přípustný rozsah jejích hodnot pro adaptaci.

Samotná adaptace každé sítě proběhne podle modifikovaného algoritmu backpropagation, který pracuje následovně. Adaptace probíhá ve dvou fázích. Po každém adaptačním kroku následuje kontrolní fáze. V jejím průběhu se zjišťuje, zda mají výstupní neurony hodnoty odpovídající požadovanému výstupu. Pokud je skutečná hodnota každého výstupního neuronu větší (resp. menší) než stanovená prahová hodnota $b = 0.5$, pak danou hodnotu příslušného výstupního neuronu považujeme za 1 (resp. 0). Tato částečná adaptace probíhá tak dlouho, dokud není splněna tato podmínka pro všechny neurony ve výstupní vrstvě. Pak je učení ukončeno.

V průběhu adaptace musíme zabránit, aby váhová hodnota nepřekročila povolený rozsah. Pokud některá váha dosáhne povolené hraniční hodnoty, pak se již v této iteraci nemění. Takto naučenou konfiguraci sítě zpětně rozpočítáme do nejméně významných pixelů použitého obrázku a obrázek aktualizujeme.

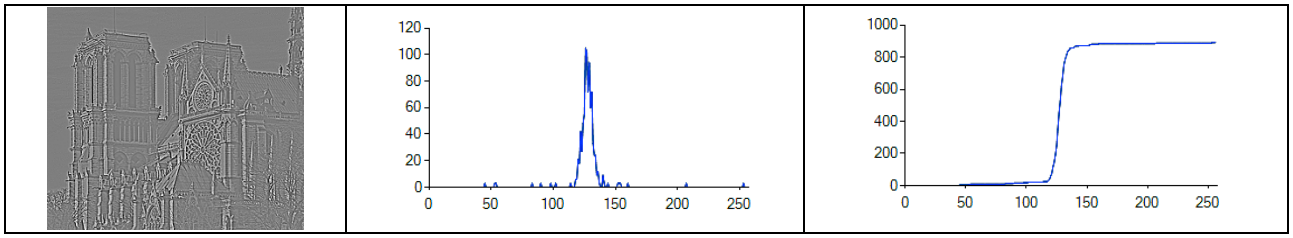
Experimentální ověření

Obrázky 2 a 3 zobrazují stejný snímek s vloženou zprávou a bez vložené zprávy včetně histogramů. Je zřejmé, že rozdíly v odpovídajících příslušných histogramech jsou minimální.



(a) (b) (c)

Obrázek 2. Snímek bez vložené zprávy (a), histogram (b) a kumulativní histogram (c)



(a) (b) (c)

Obrázek 3. Snímek s vloženou zprávou (a), histogram (b) a kumulativní histogram (c)

Závěr

Dle výsledků experimentální studie lze konstatovat, že na rozdíl od standardních steganografických metod, které se snaží ukrýt informaci do nejméně významných bitů v obrázku (tj. aniž by braly v potaz datovou reprezentaci), je využití neuronových sítí pro ukrytí informace perspektivní zejména proto, že pro adaptaci neuronové sítě cíleně využíváme vizuálních markantů obrázku (krycího média).

Literatura

[1.] MORKEL, T.; ELOFF, J. H. P.; OLIVIER, M. S.: An Overview of Image Steganography, Proceedings of the Fifth Annual Information Security South Africa Conference (ISSA2005), Sandton, South Africa, June/July 2005. Dostupné na internetu: http://mo.co.za/open/stegoverview.pdf [cit. 2015-1-25]

[2.] NAGESH H. R.; ADARSH R. K.; CHANDRA S. K.: Digital steganography: seeing the unseen, Manipal Institute of Technology, Manipal, India, datum nenalezeno. Dostupné na internetu: http://studentprogress.com:8080/uploads/cogrec/impnts/stegnocamera.pdf [cit. 20158-1-25].

[3.] BLUM, L., BLUM, M., AND SHUB, M. A.: Simple Unpredictable Pseudo-Random Number Generator. In SIAM Journal on Computing, volume 15 (2): 364-383. (1986)

[4.] FAUSETT, L. V.: Fundamentals of Neural Networks. Prentice-Hall, Inc., Englewood Cliffs, New Jersey 1994.

Abstract

The concept of steganography has its origins in Greek, in the words stegos "hidden" and grafia "writing" or "hidden writing" ("writing cover") [1]. Steganography is a method of concealing personal or sensitive information in a medium, what seems to be normal at first sight. The aim of steganography is to hide messages (information, data), where nobody expects them as well as its presence has not been detected. The focus of the paper is to present results of the application of multilayer neural networks in the field of steganography.