Mgr. Jan Butora, Ph.D.
CNRS
43 Av. le Corbusier
59800 Lille, Francie
Tel.: +33 662534002
E-mail: jan.butora@univ-lille.fr

# Posudek oponenta diplomové práce Bc. Dominika Sepaka

# "Hledání praktických robustních klasifikátorů"

Tato diplomova prace se zabyva problemem Cover-source Mismatch (CSM) in digital image
steganalysis a hledanim robustnich klasifikatoru, ktere jsou schopny tento problem odstranit ci
limitovat. V prvni kapitole student predstavuje zakladni myslenku obecnejsiho problemu robustnosti
klasifikatoru. Druha kapitola se venuje Supervised Machine Learning. Definovany jsou zakladni pojmy
jako Risk minimization a Empirical minimization a prakticke rozdily mezi nimi. Uvedene pojmy jsou
ilustrovany na dvou prikladech 1) Linearni regrese pomoci least squares 2) Neural network a gradient
descent. Dale se definuje Vapnik-Chervonenkis (VC) dimenzi. Student poukazuje, ze pokud je VC
dimenze daneho prediktoru prilis velka (napr. vetsi nez kolik mame datapoints), pak hrozi riziko
overfitting. Toto je take diskutovano za Vetou 2.1, ktera dava vztah mezi risk minimization loss a
empirical loss minimization. Zbytek kapitoly se diskutuje nesrovnalosti mezi training a testing sets.
Treti kapitola predstavuje koncept of Environments a predstavuje nekolik moznosti jak pomoci nich
merit robustnost klasifikatoru.
Ve ctrvte kapitole se zamerujeme na cil teto prace, coz je Image steganography, konkretne pak cover-
source mismatch (CSM). CSM nastava v pripade, kdy steganalyst (ktery se snazi detekovat
steganografii) trenuje svuj detektor v jednom cover source, i kdyz steganographer vytvari images z
jineho cover source. V predchozi literature se ukazuje, ze v takovem pripade je vetisnou takovy detek-
tor k nicemu. Hlavni prinos teto prace je pak experimentalni cast na konci ctvrte kapitoly, kdy student
natrenoval nekolik detektoru (jeden detector pro jeden cover source) a implementoval nekolik ruznych
robustnich pristupu, ktere vysvetlil ve treti kapitole. Pro mereni robustnosti potom

pouzil nekolik fo-
toaparatu A-F pro trenink a dva jine G-H nechal pro testy robustnosti.
I kdyz dane vysledky nejsou robustni skrze ruzne JPEG Quality Factors (QF), nevidim to vubec jako
problem, protoze JPEG kvalita se da casto zhruba odhadnout z quantization table ktera je vzdy v
JPEGu pritomna. Naopak, pokud se omezime na jeden QF, tak vysledky naznacuji slibny smer vyzku-
mu, jelikoz CSM se da do jiste miry limitovat metodami, ktere byly v teto praci predstaveny.
Prace se velmi dobre cte a je od zacatku jasne, kam smerujeme. Chtel bych podotknout, ze dokoncit
vsechny experimenty muselo zabrat slusnou CPU/GPU spotrebu, zejmena hledani hyperparametru pro
jednotlive detektory, kterych je mnoho.
Komentare

- stavebni bloky jako section, figure, table, theorem atd. by meli zacinat velkym pismenem pokud
se odkazuji do textu.
- Na nekolika mistech ‘steganograph’ -> ‘steganographer’
- Str. 14: ‘the contents of the cover image’ -> ‘the content of the cover image’
- Str. 16: ‘one of the most general lost functions it the is the’ -> ‘one of the most general lost
functions is the’
- Definice stochastic gradient descent na strane 20 by spise mela definovat batch gradient
descent, jelikoz stochastic se nazyva gradient descent kde se gradient pocita pouze z jednoho
samplu, ale v uvedene definici se gradient pocita pres subset, tedy nekolik samplu.
- Definice 2.5: ‘VC dimension in \infty’ -> ‘VC dimension is \infty’
- Velke O je definovano za Vetou 2.1 (str. 22), ale bylo pouzito uz drive na strane 20
- Str. 24: ‘used by the steganograph.’ -> ‘used by the steganographer.’
- Str. 26: ‘steganograph may select the cover source to with minimize the accuracy’
->’steganographer may select the cover source to minimize the accuracy’
- Str. 27: ‘in the experimental section of this thesis 4.3’ -> ‘in the experimental Section 4.3 of this
thesis’
- Str. 28: ‘The clairvoyant method does not recognize between the environments’ -> ‘The
clairvoyant method does not distinguish between the environments’
- Bylo by dobre na konci Sekce 3.2.1. zminit, proc se najednou namisto pojmu clairvoyant
pouzivame holistic
- Co znamena superscript ‘regtr’ v rovnici (3.5)? Zrejme jde o preklep, jelikoz v (3.4) se pro
stejny vyraz pouziva ‘reg’
- Str. 31: ‘prediction 3.3’ -> ‘prediction (3.3)’
- Rovnice (3.6) neni spravne definovana, koeficienty \lambda by meli utvaret konvexni

kombinaci jednotlivych robust objectives. Jedna moznost jak to napravit je indexovat danou
sumu skrze i=1,...q, a mit e_1, ..., e_q vsechny training environments. Pak staci nahradit kazde
e za e_i a pridat lambda_i jako weight daneho robust objective. To se take odrazi ve (3.8) a
(3.9).

- Za rovnici (3.11) se definuje \eps_S a \eps_T stejnym zpusobem. Bylo by dobre zminit, ze se
jedna o source a target error, jelikoz nelmusi byt ihned jasne co S a T reprezentuji.
- Nektere pojmy ve Vete 3.2 by zaslouzili vysvetleni. Konkretne distance d_F a mnozina F. I
kdyz F je definovana v (3.15), z dane definice se zda ze F = {-1,0,1}, coz zrejme neni spravne
- V Sekci 4.2.1. by bylo vhodne zminit, ze se v praci pro jednoduchost a rychly feedback
nepouzivaji state-of-the-art metody ani pro steganography ani pro steganalysis
- Bylo by rovnez vhodne zminit, kolik images je v pouzivanem Dataset
- Experiment 5b) v sekci 4.2.2.: ‘The risk and regret are computed on the training dataset’ ->
‘The risk and regret are computed on the validation dataset’
- Na konci str. 49 student vysvetluje, proc je steganography detected mnohem lepe pro nizsi QF.
Avsak s jeho vysvetlenim uplne nesouhlasim, nebot napr. jak zminuji v moji prvni otazce,
number of embedding changes at QF 100 should be in fact higher. Spise bych argumentoval tim
jak funguji JRM features a nsF5 algoritmus, ale vysvetlovani proc je QF 100 mene detectable
neni cilem teto prace.

Ke studentovi mam nasledujici otazky:

1. Na konci Sekce 4.2.1. se zminuje prumerny pocet embedding changes pri embedding messages with
nsF5, 0.04 bpnzAC, pricemz bpnzAC znamena bits per non-zero AC DCT coefficient. Tvrdi se potom,
ze pro JPEG Quality Factor (QF) 100 je tento pocet embedding changes zhruba 1163 a pro QF 75 je to
2176. Avsak pro QF 100 mame mnohem vice non-zero AC DCT coefficients, tedy absolutni delka of
the secrete message bude vetsi nez pro QF 75, tedy musime delat vice embedding changes pro QF 100.
Proc je to tedy naopak? Myslel jsem si, ze jde o preklep, ale toto je zminovano znovu na konci strany
49.
2. Table A.2 je zajimava. Omezme se pouze na QF 75, protoze maji obecne mnohlem mensi PE. Jde
videt, ze pokud trenujeme na kamere s danym ISO setting a pak testujeme na

kamere s vyssim ISO,
detekce neni az tak spatna (kamery s nejvyssim ISO jsou D a G s ISO 1600 a 3200). K tomu bych mel
dve otazky
a) Proc je kamera B outlier, neboli proc B s ISO 100 vubec nedetekuje D a G? To ovsem neni
jednoducha otazka a odpoved nemusi byt vubec zrejma, ale muze vezt k zajimavym poznatkum.
b) Proc to nefunguje i naopak: training na vysokem ISO a detekce na nizkem? To jde videt ve sloupcich
D a G kdy PE je ‘male’ jenom pro kamery D a G, ale jde si vsimnout, ze G->D (train on G and test on
D) je mnohem horsi nez D->G, nebot G ma vetsi ISO.
Z vyse uvedeneho vyplyva, ze student splnil zadani Diplomove Prace. Prace je zpracovana na
vysoke odborne i jazykove urovni. Vzhledem k vyse uvedenemu jednoznacne navrhuji hodnotit tuto
diplomovou praci znamkou A - vyborne a doporucuji ji k obhajobe inzenyrskeho titulu.
V Lille, dne 22. května 2022
Mgr. Jan Butora, Ph.D.