

TARTU ÜLIKOOL
Arvutiteaduse instituut
Infotehnoloogia mitteinformaatikutele õppekava

Indrek Purga

Võltsitud PDF dokumentide tuvastamise võimalused

Magistritöö (15 EAP)

Juhendaja: Kristjan Krips

Tartu 2019

Võltsitud PDF dokumentide tuvastamise võimalused

Lühikokkuvõte:

Magistritöö eesmärgiks oli luua tarkvaralahendus võltsitud PDF dokumentide tuvastamiseks. Rakendus on välja töötatud suurte failihulkade analüüsiks ning iga faili osas kuvatakse hinnang selle autentsuse kohta. Valminud rakendus võimaldab tuvastada PDF faili järk-järgulise muutmise ning kasutajale genereeritakse faili kõik eelmised vaheversioonid.

Rakenduses on välja töötatud meetodid fiktiivsete dokumentide tuvastamiseks ning PDF failidest piltide, teksti ja varjatud vahekihtide eraldamiseks.

Rakendus on kergelt laiendatav teistele failiformaatidele ning lisatud on algoritm võimalike manipuleeritud pildifailide tuvastamiseks.

Rakenduse loomiseks kasutati programmeerimiskeelt Java.

Võtmesõnad:

PDF, failiuuring, võltsitud dokument, fiktiivne dokument, pettus, Java

CERCS: P170 Arvutiteadus, arvutusmeetodid, süsteemid, juhtimine (automaatjuhtimisteooria)

Detection of forged PDF documents

Abstract:

The aim of the present thesis was to create a software solution for detecting forged PDF documents. The software was designed to handle a large number of files and to display an assessment of authenticity for each file. If the source code of the forged PDF file contains information about previous versions, then all previous versions of the file are made available.

The software includes methods for identifying fictitious documents and for extracting images, text and hidden layers from PDF documents.

The software is easily scalable for other file formats and an algorithm for detecting manipulated images was also added.

The software was developed using Java.

Keywords:

PDF, file forensics, forged document, fictive document, fraud, Java

CERCS: P170 Computer science, numerical analysis, systems, control

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Indrek Purga,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose

Võltsitud PDF dokumentide tuvastamise võimalused,

mille juhendaja on Kristjan Krips,

reprodutseerimiseks eesmärgiga seda ainult säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, alates 01.01.2030 kuni autoriõiguse kehtivuse lõppemiseni.

3. Olen teadlik, et punktis 1 ja 2 nimetatud õigused jäävad alles ka autorile.

4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Indrek Purga

14.08.2019