LIČNO APLIKATIVNO REŠENJE ZA IŠČITAVANJE EXIF TIPA METAPODATKA IZ JPEG FORMATA FOTOGRAFIJE, NJIHOVA ULOGA I PRIMENA U POLJU DIGITALNE FORENZIKE

Đorđe Todić*,
Petar Jakić,
Uroš Arnaut,
Aleksa Ćuk

Univerzitet Singidunum,
Beograd, Srbija

Rezime:

U ovom radu je opisano jedno rešenje alata za iščitavanje metapodataka iz fotografija formata JPEG. Aplikacija je razvijena u programskom jeziku JAVA kao GUI rešenje. Osnovna prednost u odnosu na ostala dostupna rešenja je u činjenici da je ovo desktop aplikacija, za razliku od većine koje su razvijane kao internet aplikacije. Dodatno, zbog svojih karakteristika, aplikacija omogućava brže dobijanje informacija kao i mogućnost poređenja exif tipa metapodatka između dve fotografije. Samim tim dobijene informacije imaju veliki značaj u polju digitalne forenzike pružajući mogućnost otkrivanja velikog broja relevantnih informacija o prethodno navedenom vidu digitalnog podatka. Jedan od glavnih pokretača za izradu ovog projekta jeste pokušaj da se suzbije vid verbalnog uznemiravanja koji je zastupljen putem raznih društvenih aplikacija, kao i to u kojoj meri bi se moglo dokazati izvorno poreklo digitalnog sadržaja. Aplikacija može naći primenu i u tehničkom smislu a to je analiza parametara fotografije. Analiza parametara pruža pomoć u usavršavanju tehnike fotografisanja i uviđenju grešaka prilikom setovanja uređaja.

Ključne reči:

Exif, Java, Jpeg, metapodaci, digitalna forenzika.

1. UVOD

Metapodaci predstavljaju “podatke o podacima”, u ovom slučaju dodatne informacije o fotografiji koja je sačuvana u jpeg formatu.[1]

U digitalnom smislu to su podaci koji dodatno opisuju, objašnjavaju i omogućavaju lakše upravljanje resursima. Neki od metapodaka su automatski generisani od uređaja koji je kreirao fotografiju. Takođe postoji način za njihovo ručno dodavanje ili menjanje uz pomoć adekvatnih softvera kao što su GIMP ili Adobe Photoshop. Zastupljeni su i primeri gde se metapodaci mogu dodatno ubaciti i putem samog uređaja, jedan od primera je digitalna kamera. Metapodaci takođe mogu pomoći u vidu zaštite intelektualne svojine. Mora postojati svest o tome da kao što postoji način da se nešto naknadno utisne kao dodatna informacija o fotografiji, isto tako postoji i način da se to ukloni ili izmeni. Metapodaci se mogu definisati u tri glavne kategorije. Tehničke metapodatke, opisne metapodatke i administrativne metapodatke. Ono čime se primarno bavi pomenuta aplikacija jeste analiza tehničkih metapodataka, kao

Odgovorno lice:
Đorđe Todić

e-pošta:
djole.toda@gmail.com

DOI: 10.15308/Sinteza-2019-256-261

tipova podataka koji su „fabrički“ utisnuti u fotografiju. Kao primer se navodi informacija o proizvođaču, modelu uređaja, vremenu i datumu kreiranja fotografije, u određenim slučajevima i GPS lokaciji. Iščitavanje metapodataka može biti otežano raznim tehnikama kompresovanja podataka koji su zastupljeni na društvenim mrežama. U predstojećem delu rada biće opisano nesto više o tagovima exif tipa metapodatka, u kojoj meri se kompresuje digitalni sadržaj putem društvenih mreža, u kojim situacijama možemo ili ne možemo iskoristiti exif tip metapodatka kako bi dokazali poreklo testiranog sadržaja, analiza sličnih aplikacija kao i sama struktura i funkcionisanje pomenute aplikacije.

2. GUBITAK METAPODATAKA SLANJEM DIGITALNOG SADRŽAJA PUTEM DRUŠTVENIH MREŽA

Facebook

Facebook koristi „zlib“ tip kompresovanja čiji je kod javno dostupan na „github“ sajtu. [2]

Na prvi pogled, po pitanju kvaliteta, ne mogu se uvideti razlike između originalne i fotografije poslate putem facebook društvene mreže.Kad su u pitanju metapodaci „zlib“ briše sve informacije o poslatoj fotografiji. Ne postoji razlika u tome da li je fotografija preuzeta(download) ili direktno sačuvana(Save As).

Na osnovu ovih rezultata, poreklo poslate fotografije je nemoguće dokazati.

Jedini uspešan način u ovom vidu testiranja jeste taj kada je slika, koja je poslata, prethodno arhivirana u neku od tipova arhiva(zip, rar, itd.). Tada „zlib“ ne pristupa dodatnom kompresovanju poslatog sadržaja i metapodaci se, nakon ekstraktovanja arhive, mogli iščitati.

Gmail

Prilikom prosleđivanja fotografije putem Gmail aplikacije dobija se veća količina upotrebljivih metapodataka. Jedna od glavnih jeste podatak o „proizvođaču“ uređaja preko koga je nastala fotografija, dok su detalji o modelu, lokaciji, vremenu i datumu izgubljeni. Samim tim ako je uznemirujući sadržaj prosleđen putem gmail-a i posedujemo uređaj optuženog, možemo uporediti da li odgovara proizvođaču uređaja koji on poseduje, što nam može pružiti pomak u daljoj istrazi.

Apple Imessage

Aplikacija koja podržava razmenu poruka putem interneta na svim apple uređajima od verzije „IOS 5“ i „OS X Mountain Lion“ verzije.

Prilikom istraživanja fotografije poslate sa jednog uređaja na drugi dolazi se do ohrabrujućih informacija.Gubitak metapodataka nije zabeležen.Samim tim se mogu posedovati dokazi o tačnom tipu i modelu uređaja, datumu i vremenu nastanka fotografije kao i podatke o tome gde je nastala fotografija, ukoliko je u tom trenutku bila uključena opcija lokaciranja (GPRS) na datom uređaju.

Viber i WhatsApp

Aplikacije funkcionišu po sličnom principu kao Facebook što prouzrokuje iste rezultate. Brišu se sve informacije o poslatoj fotografiji.

3. EXIF

Šta je Exif

Exif metapodaci su zapis koji prikazuje podatke digitalnih SLR(*Single-line reflex*) fotoaparata koje se koriste za snimanje određene fotografije. Ovi podaci se snimaju u stvarnu datoteku slike.

Svaka fotografija ima svoje jedinstvene podatke. Exif podaci prikazuju informacije o fotografiji, kao što su model kamere, ekspozicija, otvor blende, ISO, koji je režim kamere bio korišćen, da li je ili nije aktiviran blic, i dr.[3]

Exif specifikacija

Specifikacija Exif slikovne datoteke propisuje metod snimanja slikovnih podataka u datotekama i navodi sledeće stavke:

- Struktura datoteka sa slikovnim podacima,
- Oznake koje koristi ovaj standard i
- Definisanje i upravljanje verzijama formata.

Karakteristike specifikacije Exif slikovnih datoteka uključuju sledeće.

Format snimanja datoteka je zasnovan na postojećim formatima. kompresovane datoteke se snimaju kao

JPEG (ISO / IEC 10918-1iv) sa umetnutim segmentima markera aplikacije (APP1 i APP2). Nekompresovane datoteke se snimaju u TIFF Rev. 6.0v format. Upotreba postojećih formata znači da datoteke snimljene pomoću DSC(digital still camera) ili srodnog sistema mogu biti direktno pročitane od strane komercijalne aplikacije, i omogućava korišćenje funkcija za pregled i manipulisanje slikama.

Informacije o srodnim atributima za kompresovane i nekompresovane datoteke čuvaju se u formatu informacija o oznakama definisano u TIFF Rev. 6.0. Informacije koje su specifične za sistem kamere i nisu definisane u TIFF-u čuvaju se u privatnim oznakama registrovan za Exif. Specifikacija Exif fajla takođe određuje metod za snimanje sličica. Razlog za korišćenje TIFF Rev. 6.0 formata oznaka u kompresovanoj datoteci APP1 segment je da olakša razmenu atributa podataka između Exif kompresovanih i nekompresovanih datoteka.

Kompresovane datoteke mogu snimati proširene podatke koji prelaze 64 kilobajta tako što ga dele više APP2 segmenta. APP2 segment se koristi prilikom snimanja FlashPix[vi] ekstenzija.

Značajka Exif slikovnih datoteka je njihova kompatibilnost sa standardnim formatima u širokoj upotrebi danas, omogućavajući im da se koriste na personalnim računarima i drugim informacionim sistemima. Namera je da se promoviše široka upotreba digitalne fotografije kamere.[4]

Tagovi – IFD structure

IFD (image file directory) koji se koristi u ovom standardu sastoji se od 2-bajtnog broja, 12-bajtnih polja kompatibilnosti, i 4-bajtni pomeraj na sledeći IFD, u skladu sa TIFF Rev. 6.0.

Svaka od 12-bajtnih kompatibilnosti polja se sastoji od sledećih četiri elementa.

- Bytes 0-1 Tag
- Bites 2-3 Tip
- Bites 4-7 Brojač(*count*)
- Bites 8-11 Value Offset

Tag

Svakoj oznaci se dodeljuje jedinstveni 2-bajtni broj za identifikaciju polja. Oznaka brojeva u Exif nultog IFD i 1. IFD isti su kao brojevi TIFF oznaka.

Tip

Sledeći tipovi se koriste u Exif-u:

- 1 = BITE je 8-bitni nepotpisani ceo broj,
- 2 = ASCII je 8-bitni bajt koji sadrži jedan 7-bitni ASCII kod. Završni bajt se završava sa NULL,
- 3 = SHORT je 16-bitni (2-bajtni) nepotpisani ceo broj,
- 4 = LONG je 32-bitni (4-bajt) nepotpisani ceo broj,
- 5 = RATIONAL - Prvi LONG je numerator, a drugi LONG izražava imenilac,
- 7 = UNDEFINED je 8-bitni bajt koji može uzeti bilo koju vrednost u zavisnosti od definicije polja,
- 9 = SLONG je 32-bitni (4-bajtni) potpisani celobrojni (dvokomplementna notacija),
- 10 = SRATIONAL 2 SLONGs. Prvi SLONG je numerator, a drugi SLONG je imenilac.

Vrednost

Broj vrednosti. Treba pažljivo napomenuti da brojanje nije suma bajtova. U slučaju jedne vrednosti od SHORT (16 bita), na primer, broj je ‘1’ iako je dva bajta.[4]

Vrednost pomaka

Ovaj tag beleži pomak od početka TIFF zaglavlja do pozicije u kojoj je snimljena sama vrednost. U slučajevima kada se vrednost uklapa u 4 bajta, sama vrednost se beleži. Ako je vrednost manja od 4 bajta, vrednost je u 4-bajtnoj oblasti počevši sa leve strane, tj, sa donjeg kraja zone odstupanja bajta. Na primer, u velikom endian format, ako je tip KRATKO, a vrednost 1, on se snima kao 00010000.H.

Obratite pažnju da kompatibilnost polja mora biti zabeležena u nizu počevši od najmanjeg broja oznake. Nema odredba u vezi sa redosledom ili pozicijom snimanja vrednosti oznake.[4]

3. POREĐENJE POSTOJEĆIH APLIKACIJA I IMPLEMENTACIJA SOPSTVENOG REŠENJA

Postoje različita rešenja za iščitavanje exif tipa metapodatka iz JPEG formata slika i upoređivanja istog sa

dve ili više različitih slika, što je svrha rada aplikacije opisane u radu. Neke od postojećih softvera su Internet rešenja u vidu online aplikacija, druga rešenja su gotove desktop aplikacije otvorenog i zatvorenog koda. Svaka od njih ima svoje prednosti i mane.

Internet aplikacije

Prednosti Internet rešenja ogledaju se u dostupnosti i brzini dobijanja povratnih informacija. Za verifikaciju slike je dovoljno da posedujete na Vašem računaru one slike koje želite da proverite, da posedujete Internet pretraživač i konekciju na Internet.

Internet aplikacije mogu da budu rešene u vidu vebsajtova, korišćenjem različitih tehnologija kao što su PHP, .NET, JavaScript, itd. Najčešća rešenja su korišćenjem JavaScript jezika u kombinaciji sa HTML-om za prikaz stranica. Takve stranice prikazuju formu u kojoj imate mogućnost da izaberete fotografiju iz Vašeg računara i pročitate njen sadržaj metapodataka i dobijete tabelarni prikaz istih.

U takvoj aplikaciji u samo nekoliko klikova dolazite do željenih rezultata. Problem kod ovakvih aplikacija se javlja kod poverljivosti. Ne možete da verujete drugoj strani da neće kompromitovati Vaše podatke ili ih sačuvati za sebe radi svojih daljih istraživanja. Postoji mogućnost da će sve što postavite (eng. Upload) na takvom vebsajtu biti iskorišćeno u neke druge svrhe.

Desktop aplikacije

Bolja rešenja su desktop aplikacije, kod kojih lokalno možete uporediti slike i iščitati podatke bez konekcije na Internet. Međutim, idalje postoji verovatnoća da je osoba ili kompanija, koja Vam je dostavila svoje rešenje, ostavila sebi mogućnost da dobije podatke iz aplikacije sa Vašeg računara (eng. Backdoor). To može da bude prilikom prve sledeće konekcije na Internet iako aplikacija ne radi.

Desktop aplikacije (otvorenog i zatvorenog koda) nude iste ili proširene mogućnosti u odnosu na Internet rešenja. Nude veću sigurnost, u odnosu na prethodno opisane, iz razloga što veće ili manje korporacije nude svoje usluge i stoje iza njih. Takva rešenja su uglavnom u Java i C# programskim jezicima. Ti jezici nude pregršt biblioteka koje nude različite funkcionalnosti.

Kod gore navedenih rešenja takođe postoji mogućnost neprikazivanja određenih delova metapodataka koji su Vama neophodni za upoređivanje, to se ogleda kroz neefikasnost i nepouzdanost. Dodatni nedostaci se mogu ispoljiti u nemogućnosti ostvarivanja konekcije na Internet, što predstavlja nepouzdanost, jer se ne možete osloniti na takve informacione sisteme. Šta više, jedna od najvećih opasnosti je prikupljanje Vaših podataka, i njihovim skladištenjem u bazama podataka.

Nezvanično, korporacije mogu da plate programerima da ostave ulaz u Vašu aplikaciju kako bi prikupljale podatke i kasnije ih koristile protiv Vas ili u svrhu istraživanja za sopstvene potrebe.

Implementacija i pojašnjenje sopstvenog rešenja

Iz prethodno navedenih razloga, možda je najbolje rešenje razvijanje sopstvenog informacionog sistema koji će obavljati sve potrebne funkcionalnosti. Prednost ovakvog rešenja je što u svakom trenutku poznato koji podaci će biti pročitani, na koji način, koje biblioteke će se koristiti, koji tipovi fotografija su podržani, troškovi su izrade i održavanja smanjeni, itd.

Kao što je prethodno opisano u radu, napravili smo aplikaciju koja će učitavati dve slike sa računara i raditi njihovo poredjenje na osnovu metapodataka. Kada aplikacija uporedi slike, možemo da zaključimo da li su nastale na istom uredjaju, lokaciji, itd.

Uzmimo za primer da imate neke slike za koje se ne sećate gde su nastale i da li su iz istog perioda ili iz istog mesta. U softveru ćete izabrati slike za koje želite da dobijete informacije, kada ste to uradili, videćete da postoji opcija za otpakivanje metapodataka. To je prvi korak koji je potrebno uraditi. Sledeći korak je da izvršimo poređenje ovih podataka, koje je glavni cilj ovog rešenja.

Da bi aplikacija mogla da čita meta podatke korišćena su biblioteke :

- com.drew.imaging.ImageMetadataReader;
- com.drew.metadata.Directory;
- com.drew.metadata.Metadata;
- com.drew.metadata.Tag;

Jedan od jednostavniji načina da se čitaju meta podaci je preko „ImageMetadataReader“, koji koristi „FileTypeDetector“ kako bi odredio koji metod dekodiranja će biti najbolji da se iskoristi. Ova biblioteka dobija metapodatke iz svih podržanih formata datoteka, uključujući JPEG, RAW (NEF / CRv / CR2) i TIFF.

Ako se zna tačan tip određenog fajla, moguće je koristiti specifičan „MetadataReader“ koji će biti namenjen takvom tipu podatka (JpegMetadataReader za JPEG datoteke, ili

TiffMetadataReader za TIFF i RAW datoteke), s tim da takvi namenski čitači ne nude mnogo bolje performanse.

Objekat meta podataka sadrži nula ili više objekata direktorijuma (eng. Directory) i sadrže nula ili više „Tag“ objekata. Oznake sadrže vrednosti koje predstavljaju metapodatke za izvornu sliku.

Da bi fajl mogao da se otvori u aplikaciji, korišćene su biblioteke:

- java.io.File;
- javax.swing.JFileChooser;

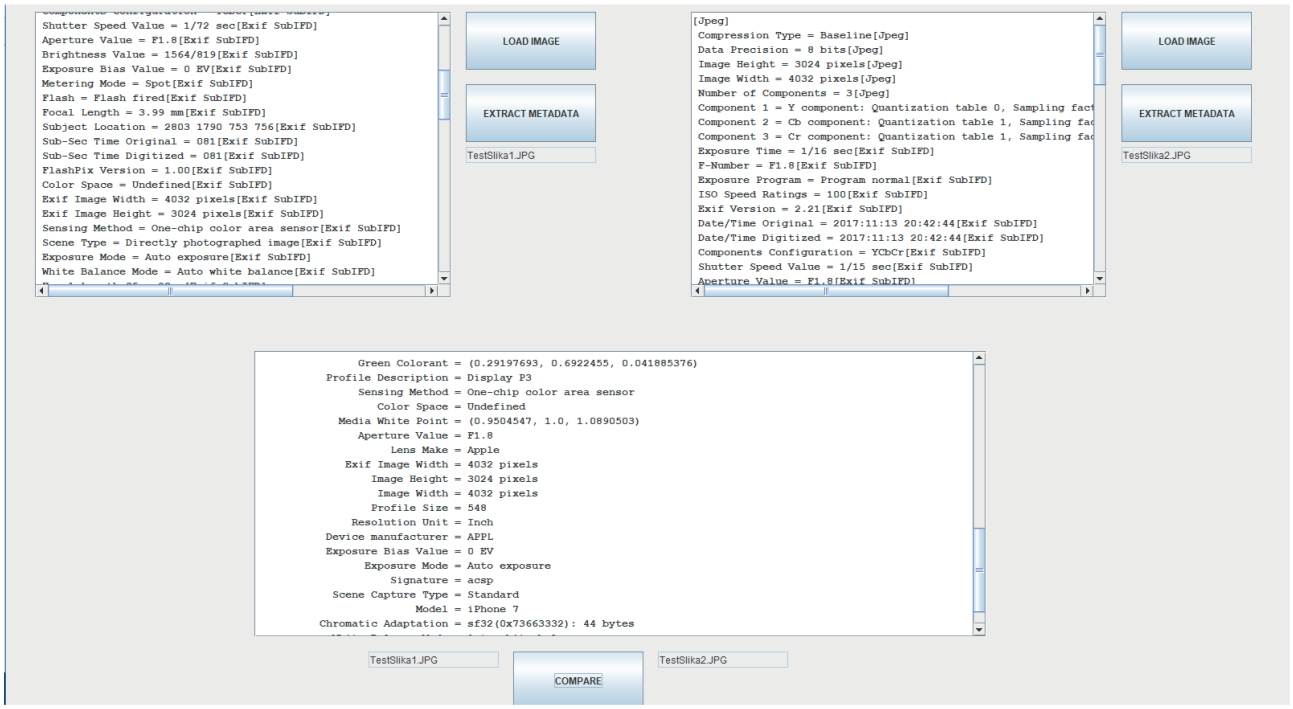
Java klasa javax.swing.JOptionPane ima mogućnosti za kreiranje prozora za dijalog, koji se može pojaviti na radnoj površini računara (na ekranu) da bi se od korisnika zatražio unos poruke ili za prikazivanje poruke korisniku. Na slici 1 se može primetiti da je kreiran korisnički interfejs koji ima za cilj jednostavniju interakciju između klijenta i računara. [5, 6]

Grafički interfejs od klijenta zahteva unos slike, čije ime ispisuje u labeli namenjenoj za čuvanje imena slike. Dugme „Extract metadata“ je jedna od dve namene ovog softvera i nudi prikaz meta podataka slike unutar okvira namenjenog za ispis teksta, taj okvir ima predefinisane dimenzije sa mogućnošću pomeranja sadržaja po x i y osi. Druga namena aplikacije je upoređivanje sadržaja dve slike, upoređeni sadržaj se prikazuje kao tekst na dnu grafičkog interfejsa sa mogućnošću pomeranja sadržaja po x i y osi.

Biblioteka java.util.HashMap je takođe korišćena u aplikaciji. Heš mapa je deo Java biblioteka od verzije 1.2. Ona obezbeđuje osnovnu implementaciju „Map“ interfejsa Jave. Ova biblioteka interno koristi listu veza da čuva podatke u parovima koji su poznatiji kao ključ-vrednost, što znači da biste pristupili vrednostima, morate znati ključ. HashMap koristi tehniku pretvaranja velikog Stringa u mali String koji predstavlja isti niz. Kraća vrednost pomaže u indeksiranju i bržem pretraživanju.[7, 8]

Za rad softvera su korišćene i druge biblioteke koje doprinose pouzdanijem i bržem radu, ali neće biti opisane u ovom radu. Ovakvo rešenje problema iščitavanja exif tipa metapodatka iz slika doprinosi bržem, efikasnijem i pouzdanijem radu. Prednosti su mogućnost rada u „offline“ režimu, odnosno da aplikacija ne mora da radi isključivo zahtevajući Internet konekciju. To doprinosi zaštiti podataka u smislu da ćemo biti sigurni da neće doći do kompromitovanja istih od strane trećeg lica. Dodatni vid zaštite bi bio da se aplikacija koristi na zasebnom računaru koji ne bi imao pristup internetu i služio bi isključivo za namene ovog softvera.

Još jedna od prednosti ogleda se u brzini rada. Jednostavnost grafičkog interfejsa omogućava korisniku da u samo nekoliko klikova iščita podatke i uporedi ih.



Slika 1. Slika prozora aplikacije za učitavanje i poređenje metapodataka

4. ZAKLJUČAK

U poslednjih par godina došlo je do velike ekspanzije društvenih mreža kao primarnog sredstva za komuniciranje među ljudima. Sa druge strane, sa sve većom upotrebom javlja se i pretnja za verbalnim nasiljem putem istih. Najveći problem je usklađivanje prava na privatnost i prava na dostupnost informacijama, pošto postoji potreba za kontrolom i praćenjem toka svih informacija, ali i pravo pojedinca na privatnost. Nivo anonimnosti koji počinioci mogu imati je visok jer oni mogu da izvrše napad, obrišu nalog i negiraju krivicu. Razmatrana aplikacija se može usavršavati u smeru otkrivanja eventualnih promena u samoj fotografiji, prikazivanju stepena tih promena, kao i pokušaju dovođenja svih metapodataka na početno stanje.Integritet digitalnog dokaza mora biti sačuvan ukoliko treba biti prezentovan na sudu, samim tim se radi na daljem razvoju deskop aplikacije i licenciranju iste.

LITERATURA

[1] M. Milosavljević, G. Grubor, Digitalna forenzika računarskog sistema, Univerzitet Singidunum, 2009

[2] Facebook, Zstandard - Fast real-time compression algorithm, github.com/facebook/zstd, poslednji pristup 28. 03. 2019.

[3] T. Punti, What Is Exif Metadata - SLR Photography Guide, slrphotographyguide.com/what-is-exif-metadata/, poslednji pristup, 29. 03. 2019.

[4] Digital Still Camera Image File Format Standard (Version 2.1), Japan Electronic Industry Development Association, Japan, pp 288, -1998.

[5] JavaTpoint, Java JOptionPane, www.tutorialspoint.com/java/util/java_util_hashmap.htm, 2.4.2019.

[6] Oracle, JOptionPane (Java Platform SE 7), docs.oracle.com/javase/7/docs/api/javax/swing/JOptionPane.html, 2.4.2019.

[7] Beginners book, HashMap in Java with Example, beginnersbook.com/2013/12/hashmap-in-java-with-example, 1.4.2019.

[8] Oracle, HashMap (Java Platform SE 7), docs.oracle.com/javase/8/docs/api/java/util/HashMap.html, 2.4.2019.