

OBRAZLOŽENJE PRIJEDLOGA TEME DOKTORSKE DISERTACIJE

Kandidat:

mr.sci Asmir Butković, dipl.el.ing

Radni naslov teme doktorske disertacije:

"Novi pristup u modeliranju geolociranja počinitelja sajber kriminala "
"A novel approach in modeling of geolocation cybercrime perpetrators "

1. Tip istraživanja

Jedna od definicija sajber kriminala je da se radio o *"nezakonitim kompjuterski-posredovanim aktivnostima koje se mogu obavljati putem globalne elektronske mreže"* [1]

Upravo otkrivanje i razjašnjavanje ovih krivičnih djela predstavlja trenutno najveći izazov policijskim i pravosudnim institucijama u zemljama širom svijeta. Ova doktorska disertacija će se bazirati na istraživanju procesa geografskog profiliranja sajber kriminala analizirajući metapodatake u zaglavlju e-maila. Područje istraživanja će obuhvatiti oblasti informacionih sistema, geografskih informacionih sistema i prepoznavanje oblika kriminala s aspekta integracije geopodataka, prepoznavanje trendova i obrazaca činenja krivičnih djela.

Priroda problem nužno uključuje potrebu holističkog pristupa u istraživanju kombinovanjem različitih metoda i tehnika geografskog profiliranja, IP geolociranja i e-mail forenzike.

Teorijska osnova će biti zasnovana na karakteristikama standarda RFC 2821 i RFC 2822 (*Simple Mail Transfer Protocol* i *Internet Message Format*) koji specificiraju kako bi trebalo da izgleda e-mail poruka i kako se e-mail poruke prenose između sistema, zatim Whois Based Geolocation (WBG) strategiji za geolociranje Internet hostova te upotrebi matematički i statistički metoda analize prostornog ponašanja kriminala.

U cilju validacija i verifikacije predloženog modela sve definisane metode i tehnike će biti implementirane unutar softverskog alata za procesiranje podataka upotrebom Microsoft .NET razvojnog okruženja i odgovarajućih API-ja. Na taj način će se izvršiti usporedba teorijski dobivenih rezultata s eksperimentalno određenim vrijednostima, u kontrolisanom (laboratorijskom) okruženju i na realnom sistemu što se može shvatiti kao praktični cilj ove disertacije.

2. Motivacija za istraživanje

Broj krivičnih djela počinjeni u sajber prostoru poput krađe podataka, Internet prevara, poslovne špijunaže, pornografije, seksualnog zlostavljanja, online seksualnog iskorištavanja djece, sajber terorizma u enormnom je porastu u posljednjih nekoliko desetljeća. To ima vrlo negativne efekte na korištenje Interneta za elektronsko poslovanje i sigurnost komunikacija. [2], [3], [4]

Sljedeća tabela pokazuje nekoliko najčešćih tipova sajber napada i njihov procenat zastupljenosti u ukupnom broju napada.

NAPAD	PRIJAVLJENIH SLUČAJEVA
Krađa podataka	33%
E-mail zloupotreba	22%
Neovlašteni pristup	19%
Izmjena podataka	15%
Virus napad	5%
DoS napad	3%
Drugi	3%

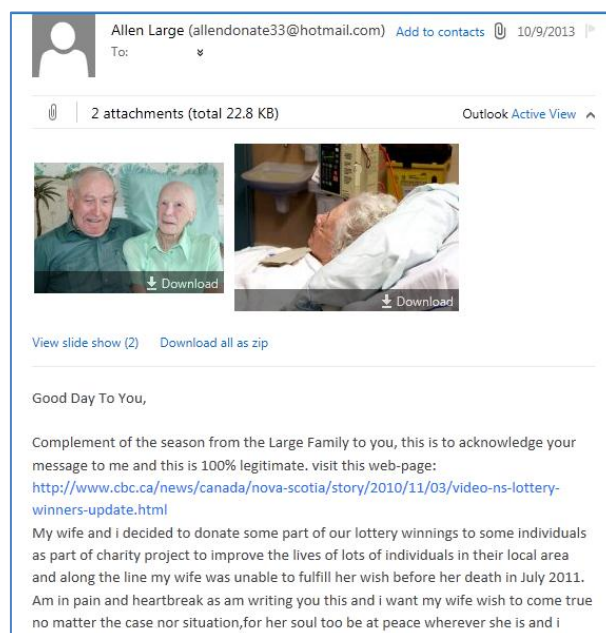
Tabela 1. Statistički podaci o različitim vrstama prijavljenih napada. [5]

U posljednjih dvadesetak godina e-mail je postao neizostavan dio načina na koji ljudi komuniciraju, privatno i javno. Nažalost, izvanredna praktičnost i učinkovitost ovog medija je prepoznata i iskorištena od osoba sklonih činjenju krivičnih djela koje su koristeći e-mail kao sredstvo za počinjenje krivičnih djela sajber kriminala pokazali njegovu tamnu stranu upotrebe. Pod kriminalnom povezanim sa upotrebom elektronske pošte (engl. e-mail related crime) ili kriminalom na bazi e-maila podrazumijeva se: [6], [7]

- Lažiranje e-mail poruka (engl. e-mail spoofing)
- Slanje malicioznog ili zlonamjernog code-a putem e-maila
- E-mail bombardovanja
- E-mail lotto prevare
- Slanje prijetećih e-mailova
- E-mail prevare
- Nigerijska prevara (Nigerijska šema 419)

Takođe i sljedeće kriminalne aktivnosti bi se mogle podvesti pod kategoriju kriminala povezanog sa upotrebom elektronske pošte: [5], [8]

- Phishing ili mrežna krađa identiteta
- E-mail koji sadrži seksualno eksplicitni sadržaj ili pornografiju
- SPAM ili neželjena pošta



Slika 1. Primjer zlonamjernog e-maila

Povećan broj incidenata sajber kriminala putem anonimnih e-mailova, gdje su elektronske poruke direktno ili indirektno sredstvo počinjenja krivičnih djela nameće potrebu za efikasnim računarskim alatima i inteligentnim sistemima koji će automatizirati analizu i interpretaciju podataka o ovim krivičnim djelima. Isto tako prostorna analiza podataka o počinjenim krivičnim djelima putem mapiranja kriminaliteta poboljšava kvalitet analize podataka i može pružiti ključne informacije o određenim kriminalnim aktivnostima. [9], [10] Javlja se ideja o istraživanju mogućnosti primjene geografskog profiliranja kao tehnike za lociranje pošiljaoca zlonamjernih e-mail poruka, te iznalaženju model i algoritma za realizaciju datog procesa.

Prvi korak bilo koje e-mail analize je identifikacija izvora i načina nastanka e-maila, odnosno načina na koji se koriste e-mail klijent i server za generisanje e-maila.

Zatim tu je praćenje i lociranje geografskog položaja (geolociranje) pošiljaoca zlonamjernog e-maila što je i neophodan preduslov za krivično gonjenje od strane agencija za provođenje zakona. Mnoge agencije za provođenje zakona uspješno su integrisale mapiranje kriminala u svoje aktivnosti analize kriminala pomoću Geografskih informacionih sistema (GIS). Međutim ovo predstavlja najelementarniju upotrebu GIS-a budući da većini GIS sistema nedostaje jedan dodatni viši nivo funkcionalnosti, a to je podrška istražnim radnjama, kroz identifikaciju i prioritizaciju osumnjičenih i područja potrage. [11], [12]

Ovaj proces prati nekoliko problema koje treba prevladati a to su:

- Lažiranje podataka u zaglavlju e-mail poruke, takozvano "maskiranje" (engl. spoofing) e-maila.
- Nepostojanje direktne veza između IP adrese računara i njegovog geografskog položaja.
- Postojeće baze podataka koje nude informacije potrebne za IP geolociranje su nedovoljno precizne i često sadrže netačne i neažurne podatke.
- Nepostojanje odgovarajućih algoritama za geografsko profiliranje krivičnih djela sajber kriminala.

Istraživanja u oblasti analize elektronske pošte obično se fokusiraju na dva područja, analiza e-mail saobraćaja i analiza sadržaja e-maila, ali veoma malo u području vizuelne analitike (engl. visual analytics) e-mail poruka. [13], [14] Stoga se analiza prostornih informacija smatra veoma važnom fazom u okviru krivičnog istražnog postupka. Ovo se posebno odnosi na slučajeve serijskih krivičnih djela gdje kriminolozi i psiholozi već primjenjuju geografsko profiliranje kako bi modelirali geografsku distribuciju krivičnih djela i obrazaca ponašanja nasilničkog kriminaliteta (slučajevi serijskih ubistava i silovanja) s ciljem procjene najvjerovatnijeg boravišta počinioca.

U posljednjih nekoliko godina dostupnost naprednih računarskih matematičkih alata pruža mogućnost i uspostave nekih novih matematičkih modela koji mogu zamjene tradicionalne empirijske metode. [15], [16], [17]

Budući da se ovdje radi o relativno novom području istraživanja i primjene postoje određene nepoznanice i dvojbe o načinu kako, i u kojoj mjeri geografsko profiliranje može pomoći istragama sajber kriminala. Stoga je namjera ovog rada da na primjeru kriminala povezanog sa upotrebom elektronske pošte, uz određene adaptacije i prilagođavanja, pokažu mogućnosti ove tehnike za procesiranje kriminalnih aktivnosti te podstakne i promoviše upotreba geografskog profiliranja od strane agencija za provođenje zakona.

3. Stanje u polju primjene

E-mail pruža pogodan i efikasan način komunikacije u savremenom poslovnom okruženju. Međutim, problem upotrebe e-maila u zlonamjerne svrhe postaje sve izraženiji. E-mail forenzika, dio digitalne forenzike, predstavlja skup naučnih metoda i tehnika za proučavanje izvora i sadržaja e-maila, identifikaciju stvarnog pošiljaoca i primaoca e-maila, datuma/vremena slanja e-maila i sl. Većina zemalja priznaje e-mail kao legitimno dokazno sredstvo na sudu. E-mailovi se koriste kao značajan izvor dokaza u slučajevima ubistava, sajber uhođenja, uznemiravanja, krađe identitete i špijunaže. [18] Zajednica digitalne forenzike je zanemarila e-mail forenziku kao proces, uprkos činjenici da e-mail ostaje važan alat u izvršenju krivičnih djela. [19] U ovom trenutku postoji malo podrške za otkrivanje, preuzimanje i analizu e-maila, bez obzira na njegovu široku upotrebu. U mnogim slučajevima nemogućnost da se utvrdi porijeklo i autentičnost e-maila rezultira nedostatkom efikasnog dokaznog materijala za rješavanje sporova i procesiranje krivičnih djela sajber kriminala. Sa sve većom zloupotrebom e-maila istražitelji zahtijevaju efikasne automatizovane alate za brzu analizu e-maila. [24], [25], [26]

Kriminalističko-istražni postupak uključuje dvije faze: (1) pronalaženje počinioca i (2) dokazivanje krivice (Rossmo 2006). Geografsko profiliranje može pomoći u prvoj fazi ovog procesa kroz ulogu osumnjičenog ili određivanju prioritnog područja čime pomaže u zadatku upravljanja informacijama. Metodologija nije dizajnirana za rješavanje krivičnih djela, to se može uraditi samo kroz svjedoke, priznanja, ili fizičke dokaze. [27], [28] Geografski kodirane informacije iz policijskih evidencija se mogu koristiti za otkrivanje trendova kriminaliteta, prepoznavanje obrazaca činenja krivičnih djela kao i potvrdu prisustva počinioca u okviru geografskog područja. U nastavku je dat detaljan pregled stanja po pojedinim oblastima istraživanja koje su od ključnog značaja za postizanje opšteg cilja disertacije.

3.1 Analiza zaglavlja e-mail poruka

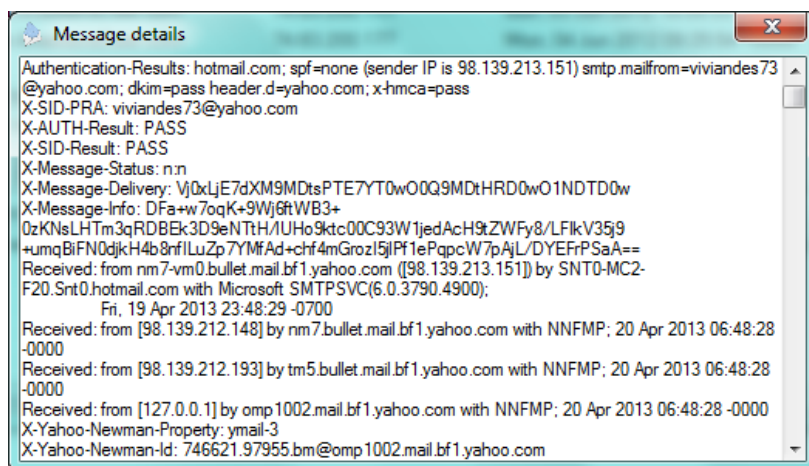
Elektronska pošta se sastoji od dva dijela, zaglavlja i tijela. Dio zaglavlje, nosi informacije koje su potrebne za e-mail usmjeravanje, predmet poruke i vremensku oznaku, dok tijelo sadrži stvarnu poruku/podatke e-maila. [29]

Informacije sadržane u zaglavlju mogu pomoći istražiteljima u praćenju pošiljaoca e-mail poruke. Potpuna istraga e-mail zaglavlja treba da sadrži:

- Ispitivanje e-mail adrese pošiljaoca
- Ispitivanje inicijalnog protokola u kojemu je nastala poruka (HTTP naspram SMTP)
- Ispitivanje Message ID
- Ispitivanje IP adrese pošiljaoca

E-mail praćenje je proces ispitivanja informacija sadržanih u zaglavlju e-mail poruke kako bi se utvrdi njen izvor.

Tipično e-mail zaglavlje izgleda ovako:



Slika 2. Zaglavlje e-mail poruke

Postoji više različitih metoda kojim se može generisati e-mail. Najčešće korišteni metodi su:

- Korištenje HTTP baziranih e-mail servisa kao što su Hotmail, Gmail, Yahoo, i drugi.
- Korištenje SMTP klijentskih aplikacija kao što su Outlook, Expedia, Pine i druge.
- Ručno kreiranje e-maila. Obično se koristi za lažne e-maile koristeći Telnet za povezivanje na port 25 (SMTP) ili korištenje drugih aplikacija ili alata koji se povezuju na e-mail servere ili primopredajne servere. [30], [31], [32]

Istraživanje e-maila kako bi se utvrdilo njegovo porijeklo se može razlikovati zavisno o njegovom metodu iniciranja.

Web bazirani e-mail je jedan od najčešćih načina slanja zlonamjernih e-mailova zbog niza faktora, među kojima su:

- Pogodnosti slanja i primanja e-mail preko Web stranice
- Besplatni Web bazirani e-mail računi se mogu dobiti vrlo lako i postavljanje traje svega nekoliko minuta
- Lažni podaci mogu biti dati prilikom postavljanja računa
- E-mail može biti poslat sa bilo kojeg mjesta, uključujući Internet kafe i javna mjesta
- "Anonimnost" - web bazirani e-mail servisi nude dodatan nivo sigurnosti za pošiljaoce zlonamjernih e-mailova

Većina e-mail sistema koji šalju poštu preko Interneta koriste SMTP za slanje poruka sa jednog servera na drugi, a poruke se onda mogu preuzeti sa e-mail klijentom koristeći POP3 ili IMAP. SMTP protokol propisuje da svaki SMTP relej koji učestvuje u slanje e-maila mora dodati na početku liste zaglavlja poruke liniju "received" koja sadrži podatke o SMTP serveru koji je primio poruku, od kog servera je primio poruku i vremensku oznaku kada je dodana zaglavlju.

Tipično e-mail prolazi kroz najmanje četiri releja prije nego stigne u inbox primaoca. Tokom SMTP transakcije ime hosta i IP adrese releja koji učestvuju u prijenosu e-maila evidentiraju se u "Received: from " liniji u zaglavlju maila, tako da sadrži IP adrese svih servera uključeni u usmjeravanje e-maila s jednog mjesta na drugo. [33], [34]

Ova linija, kada se čita od dna do vrha, pruža uvid u putanju kroz koju prolazi e-mail od pošiljaoca do primaoca. Sa ovom adresom korisnik može pratiti porijeklo e-maila i eventualno identitet pošiljaoca.

Praćenje SMTP baziranog e-maila je nešto drugačiji od praćenja HTTP baziranog e-maila. Neke karakteristike su zajedničke, kao što su statičke naspram dinamičkih IP adresa, važnost vremena, i tuneliranje. Vrijeme ima vrlo značajnu vrijednost na sudu. Ako informacija o vremenu nije sinhronizovana i ne može se potvrditi, biti će vrlo teško dokazati taj slučaj na sudu. Sinhronizacija vremena je bitna za sve servere koji su uključeni u proces praćenja e-maila.

Neka pitanja kao što su otvoreni releji (primopredajni serveri) i lažna zaglavlja su češća pojava kod SMTP e-mailova. [35], [36]

Jasno je da je nemoguće da svi policijski službenici postanu eksperti za Sistem elektronske pošte na Internetu. Postoji potreba da se razviju više intuitivne i prilagođene metode kako bi se pomoglo inspektorima u procesiranju zlonamjernih e-mailova, bez potrebe da u potpunosti razumiju detalji Sistema elektronske pošte na Internetu.

E-mail analiza se može smatrati kao rudarenje podataka ugrađenih u zaglavlju ili tijelo poruke elektronske pošte. Razne tehnike analize teksta koje izvlače nepoznate i korisne informacije iz nekog skupa e-mailova mogu se koristiti za provođenje e-mail analize. Najčešće korištene tehnike se mogu svrstati u sljedeće kategorije: automatizovano prepoznavanja autora teksta (engl. authorship attribution), analiza sadržaja, profiliranje phishing e-mailova i filtriranje neželjene pošte (spam-a).

Prepoznavanja autora e-maila znači identificirati najvjerovatnijeg autora anonimnog e-maila iz grupe potencijalnih osumnjičenih. Za prepoznavanje autora koriste se različite tehnike proučavane od strane raznih istraživača. Razne teme o kojima su napisani radovi se odnose na pol, jezik, različit stil pisanja i slično, potencijalnog autora. [37], [38]

Analiza sadržaja se definiše kao naučna metoda analize poruka kvantitativnim i kvalitativnim tehnikama koristeći naučne metode (pazeći na objektivnost, pouzdanost, valjanost, mogućnost generalizacije, zamjenjivost i testiranje hipoteze) koja nije ograničena na tipove varijabli koji bi mogli biti izmjereni ili na kontekst u kojem su poruke kreirane ili predstavljene.

Phishing se može definisati kao prevara u kojoj se e-mail korisnici navode da predaju privatne informacije koje će se koristiti za krađu identiteta. Ovi napadi koriste socijalni inženjering i tehničke mahinacije za krađu ličnih podataka i šifri bankovnog računa. To je jedna od najbrže rastućih prevara na Internetu. Ekskluzivna motivacija prevarantima je finansijska dobit.

John Yearwood koristi strukturne karakteristike primljenih e-mailova i informacije dobivene na osnovu hiperlinkova iz "Whois" baze podataka za profilisanje phishing e-mailova. [39], [40]

Takođe postoje istraživanja čije se metode klasifikacije zasnivaju na informacijama iz zaglavlja e-maila (a ne na sadržaju e-maila). Svoje klasifikacije baziraju na tri vrste analize zaglavlja: DNS bazirana analiza zaglavlja, analiza društvenih mreža i Wantedness analiza. [41], [42]

Spam, također poznat kao neželjena elektronska pošta, nanosi sve više štete e-mail saobraćaju. Filtriranje je jednostavan i učinkovit način borbe protiv spama. Klasifikacijski algoritmi sa mogućnošću mašinskog učenja pokazuju odlične performanse kod filtriranja spama. [43], [44]

Postoji dokazi da je spam pokretačka snaga botneta: Najčešća strategija za eksplatisanje botneta je slanje spam elektronske pošte, gdje spam uključuje tradicionalne e-mail oglase, phishing e-mailove, e-mail poruke sa virusima i druge neželjene e-mail poruke.

Botnet ima višestruke štetne posljedice: otpočinjanje DDoS napada, krađa korisničkih lozinki i identiteta, generisanje prevara klikom (engl. Click fraud), i slanje spam e-mailova. [45], [46]

Isto tako predlaže se selekcija atributa uz klasifikacijske tehnike za detekciju zlonamjernih e-mailova, posebno terorističkih e-mailova. Fokus je na algoritme mašinskog učenja kao što su Stabla odlučivanja, Logistička regresija, Naivni Bayes (engl. Naïve Bayes), i Mašine potpornih vektora (engl. Support Vector Machine) za detekciju e-mailova koji sadrže sumnjiv sadržaj. [47], [48] Ovaj dio analize u doktorskom radu će se prvenstveno koristiti za identifikovanje tehnike koju je moguće adaptirati i integrisati u sistem i koja daje najbolje rezultate kod utvrđivanja stvarne IP adrese uređaja sa koga je poslata zlonamjerna e-mail poruka.

3.2 IP geolociranje

IP geolociranje je proces pronalaženja geografske lokacije Internet Protokol adrese.

Internet tehnologija geolociranja (IP geolokacija) ima za cilj utvrditi fizičku (geografsku) lokaciju korisnika Interneta i uređaja. Igra ključnu ulogu u mrežnim uslugama zasnovanim na lokaciji (engl. location-aware), kao što su geolokacijski marketing, lokalizacija sadržaja, ograničavanje prodaje digitalnih sadržaja na korisnike zahtijevane dobi i jurisdikcije. Takođe za autorizaciju transakcija, tako da se izvode samo na unaprijed utvrđenim mjestima, lociranje osumnjičenih za sajber kriminal i osiguranje forenzičkih dokaza u agencijama za provođenje zakona. Važna primjena IP geolociranje je i za lociranje hitnih poziva iniciranih preko VoIP-a (Voice over IP). [49], [50] IP adresa je numerička adresa koja identifikuje čvor na mreži. Dodjela IP adresa nije proizvoljna, jer postoji organizacija koja je odgovorna za distribuciju adresnog prostora. IANA (Internet Assigned Numbers Authority) je odgovorna za globalnu koordinaciju sistema internetskog adresiranja, kao i AS (Autonomous System) brojeva koji se koriste za rutiranje Internet saobraćaja.

ISP-ovi alociraju IP adrese iz lokalnog Internet registra (LIR) ili Nacionalnog Internet Registra (NIR), ili iz njihovog odgovarajućeg Regionalnog Internet Registra (RIR) : AfriNIC - Afrički region, APNIC - Azijsko-pacifički region, ARIN - Sjeverna Amerika, LACNIC - Latinska Amerika i Karibi, RIPE NCC - Europa, Bliski Istok i Centralna Azija. [51]

Utvrdjivanje lokacija Internet hostova iz njihovih IP adresa je izazovan problem, jer ne postoji direktna veza između IP adrese hosta i njegovog geografskog položaja. [52] Rješenje se može izraziti različitim stepenom granulacije; za većinu aplikacija rezultat bi trebalo biti dovoljno precizan da odredi grad u kojemu se IP nalazi, ili se vraća ime grada ili longituda i latituda u kojoj se nalazi meta. [53]

Geografsko lociranje IP adresa je važno za akademsko istraživanje, komercijalne aplikacije i aplikacije koje mogu biti prijetnja nacionalnoj sigurnosti. Stoga su, obje komercijalne i akademske baze podataka i alati dostupni za mapiranje IP adresa u geografske lokacije. Procjena tačnosti ovih servisa za mapiranje je vrlo složena.

Ponuda servisa za geolociranje se kreće od besplatnih servisa, preko servisa koji koštaju nekoliko stotina dolara do servisa koji koštaju desetine hiljada dolara godišnje. [54]

Trenutne tehnike geolokacije se mogu svrstati u dvije glavne kategorije:

1. "statičke tehnike" koje koriste pasivni pristup geolociranju IP adrese, i
2. "tehnike zasnovane na mjerenju" koje rade aktivna mrežna mjerenja.

Statičke tehnike koriste bazu podataka ili Domain Name Service (DNS) imena susjednih rutera za geolociranje IP adrese. Jednostavan pasivni metod određivanja geografske pozicije IP adresa je korištenje javnih whois baza podataka, koje pružaju informacije o registrantu ili vlasniku IP adresnog bloka. Međutim, informacije whois baze podataka mogu biti nepotpune, zastarjele ili netačne.

Dalje, ako je veliki blok IP adresa dodjeljen jednom pravnom subjektu, onda whois baza podataka ne daje informacije o geografskom položaju individualnih IP adresa u okviru tog bloka. (npr. veliki ISP) [55], [56]

Ove baze podataka mogu biti vlasničke ili javne. Javne baze podataka obuhvaćaju one upravljane od strane Regionalnih Internet Registara (npr., ARIN, RIPE) .

Iako tačan metod izgradnje ovih baza podataka nije javno dostupan, one su ponekad zasnovane na kombinaciji whois servisa, DNS LOC zapisa i AS brojeva. [53]

Najpoznatiji geolokacijski servisi sa vlasničkim bazama podataka za mapiranje IP adrese u geografsku lokaciju, srednjeg opsega cijena su Maxmind GeoIP, IPLigence, and IP2Location. Maxmind spada među najistaknutije on-line servise koji pružaju informacije o lokaciji IP adrese, i široko se koristi u Internet industriji. Maxmind tvrdi da ostvaruje "83% preciznosti za gradove u SAD-u unutar radijusa od 40 km." To može biti dobro za njegov glavni cilj korištenja, a to lociranje korisničkih uređaja do tačnosti poštanskog broja ili ZIP koda za pomoć u marketingu, upravljanju digitalnim pravima i elektronsku trgovinu.

Može se takođe uočiti da Maxmind i InfoDB imaju istu distribuciju udaljenosti od referentnih vrijednosti. To je zbog činjenice da se InfoDB temelji na besplatnoj verziji Maxmind baze podataka. Manje od 20% je procijenjena tačnost za Maxmind i InfoDB unutar desetak kilometara od referentnih tačaka. [57]

IPLigence je geolokacijski servis provajder, koji postoji od 2006 godine. Njegov vodeći proizvod IPLigence Max sadrži geografske podatke, kao što su zemlja, region i grad, longituda i latituda, i dodatne opšte informacije kao što su vlasnik i vremensku zonu.

Hexasoft održava IP2Location geolokacijsku bazu podataka sa širokim spektrom geolokacijskih informacija, od konverzije IP adresa u kod odgovarajuće zemlje, do osiguravanja informacija kao što su propusnost i vrijeme. [54]

BAZA PODATAKA	BLOKOVI	(LAT; LONG)	ZEMLJE	GRADOVI
HostIP	8,892,291	33,680	238	23,700
IP2Location	6,709,973	17,183	240	13,690
InfoDB	3,539,029	169,209	237	98,143
Maxmind	3,562,204	203,255	244	175,035
Software77	99,134	227	225	0

Tabela 2. Osnovne karakteristike nekih geolokacijskih baza podataka [58]

Postoji problem s koordinatama, jer politika raspodjele IP adresa rezultira blokovima koji često imaju zajedničku upravu (centralnu administraciju) što uključuje istu fizičku lokaciju. Iz Tabele 2 se može uočiti velika razlika između broja IP blokova i broja jedinstvenih geografskih (parova longituda i latituda) lokacija. [58], [59] Alternativni pristup geolociranju se zasniva na izdvajanju geografskih informacija iz DNS imena krajnjeg hosta ili susjednog rutera.

Mrežni operateri često dodjeljuju domenska imena mrežnim ruterima sa ugrađenim geografskim kodovima. Izdvajanje i identifikovanje ovih geografskih kodova sa mrežnog rutera u blizini mete može pružiti korisnu procjenu njihovog geografskog položaja. Međutim, ovaj pristup nije pouzdan jer nemaju svi ruteri deskriptivna imena. Štaviše, pošto ne postoji standard za imenovanje rutera, identifikovanje ove informacije može biti težak zadatak.

Lociranje IP adrese mjerenjem kašnjenja proučava upotrebu mrežnog kašnjenja ili topologije mjerenja za procjenu prostorne (geografske) lokacije Internet čvora. Geolokacija zasnovana na mjerenju uključuje aktivna mjerenja RTT-ova od računara na poznatoj lokaciji do ciljane IP adrese. Izazov geolokacije zasnovane na mjerenju je da se pronađe odgovarajući model za predstavljanje veze između mjerenja mrežnog kašnjenja i geografske udaljenosti. Round-trip time (RTT) između mrežnih čvorova je vrijeme potrebno da se pošalje paket na odredište i vrati ponovo nazad i često se koristi kao mjera mrežnog kašnjenja. Ono se sastoji od propagacionog kašnjenja (engl. propagation delay), transmisionog kašnjenja (engl. transmission delay), kašnjenja zbog procesiranja (engl. processing delay) i kašnjenja zbog čekanja u redovima (engl. queuing delay). [60], [61]

Algoritmi geolociranja zasnovani na mjerenju koriste skup geografski distribuiranih orijentir čvorova sa poznatim lokacijama za lociranje ciljane IP adrese. Ovi orijentiri mjere različite mrežne karakteristike, kao što su kašnjenje, i saobraćaj između njih i cilja.

Algoritmi geolociranja se uglavnom oslanjaju na ping i traceroute mjerenja. Ping mjeri vrijeme povratnog puta (RTT) kašnjenja između dva čvora na Internetu, a traceroute otkriva i mjeri RTT do rutera duž putanje prema datom odredištu. [39]

Internet Control Message Protocol (ICMP) echo zahtjevi (pingovi) se koriste za prikupljanje vrijednosti RTT između par računara. Ove tehnike pretpostavljaju da su RTT-ovi i udaljenosti između računara u korelaciji.

Tehnike geolociranja na bazi kašnjenje koriste dva seta čvorova:

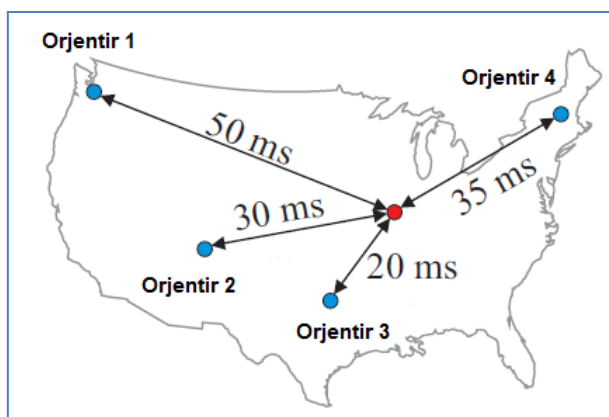
1. čvorovi ispitivači, koji pokreću pingove do drugih čvorova
2. orijentir čvorovi, koji reaguju na pingove poslate od čvorova ispitivača.

Čvorovi ispitivači i orijentir čvorovi imaju poznate lokacije.

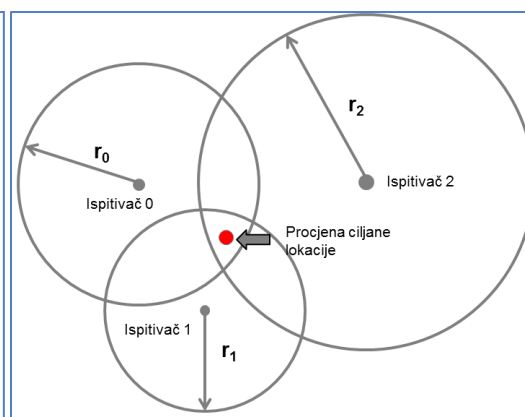
GeoPing uspoređuje ciljani vektor kašnjenja za sve orijentire pomoću Euklidske udaljenosti, a orijentir koji daje najmanju udaljenost je proračunata ciljane lokacija. Ova metoda koristi ograničen broj mjesta i na taj način daje diskretni izlaz. Rezolucija ove tehnike je reda 10^2 kilometara. [62], [63]

Druga tehnika je Constraint Based Geolocation (CBG). CBG poboljšava GeoPing koristeći mjerena kašnjenja kao ograničenja. Umjesto mapiranja mete prema jednom od orijentira, CBG koristi multilateraciju da kombinuje vrijednosti kašnjenja više čvorova kako bi dobio region u kome se nalazi ciljane lokacija. [64]

Za procjenu udaljenost do cilja iz RTT vrijednosti, svaki čvor ispitivač pinguje host da se dobije "mapa latencije" od (udaljenost, RTT) parova. (primjer Slika 4)



Slika 3. GeoPing



Slika 4. CBG u WAN

Ova tehnika ima pouzdanosti utvrđivanja lokacije hosta od 50% unutar radijusa od 80 km. CBG modifikacija, posebno Topology Based Geolocation i Octant, koristi dodatna ograničenja za ponovnu estimaciju ciljane lokacije i dobiva rezultat sa 50% pouzdanosti unutar 35 km. [37], [66]

3.3 IP bazirano geokodiranje

Postoji nekoliko načina za pronalaženje lokacije korisnika. Najbolje i najpreciznije je ako uređaj koristi aplikaciju koja ima GPS. Ali ako ne, postoje načini za lociranje, mada ne sa istom preciznošću. Desktop preglednici obično nemaju pristup GPS, ali računari priključeni na mrežu imaju IP adresu, koja može otkriti približnu lokaciju korisnika. Jedan od načina za pronalaženje lokacije korisnika je preko IP baziranog geokodiranja. [66]

Prostorna analiza kriminala ne fokusira se samo na ispravno geokodiranje, nego i na postizanju visokog stepena tačnosti geokodiranja. Geokodiranje je proces pretvaranja lokacija, poput adrese žrtve ili napadača, u mrežu koordinata i ovaj zadatak redovno obavljaju mnogi kriminalistički analitičari.

Kriminal je inherentno prostorna pojava i mapiranje kriminaliteta ima tendenciju da postane posebno područje proučavanja. Dok su neki zločini teži za mapiranje (Internet prijevare, utaje poreza i slično), većina kriminalnih aktivnosti se može lako prostorno analizirati. [67], [68]

Za datu IP adresu, servis za mapiranje treba da vrati geografski položaj hosta kome je ta IP adresa dodijeljena. To je izazovan problem jer IP adrese u sebi ne sadrži nikakav pokazatelj geografske lokacije. [69], [70], [71]

Pokretanje Google Maps 2005 godine je napravilo revoluciju u servisima za online mapiranje aplikacija na Internetu. Baziran na Asynchronous JavaScript and XML (AJAX), Google Maps uvodi novu vrstu interakcije server/klijent koji održava konstantnu vezu sa serverom za neposrednije preuzimanje dodatnih kartografskih podataka. Osim toga, Google također pruža programerima slobodan pristup njegovom kodu u obliku Application Programming Interface (API). [72], [73], [75]

Postoji nekoliko rješenja za mapiranje uključujući Yahoo! Maps, Microsoft Bing Maps, Nokia Ovi Maps, ESRI ArcGIS Maps, ali se Google Maps pokazao kao najpopularniji do sada. Google Maps API, prema Programmableweb.com je najpopularniji API na Internetu i prema statistici iz januara 2014, 39% svih mashup-ova koriste ovaj API.

Aplikacije i web sajtovi koji koriste podatke i funkcionalnost iz više različitih izvora, se obično nazivaju mashup-ovi. Ove aplikacije postaju sve popularniji i napravile su revoluciju u načinu na koji se informacije koriste i vizualizuju. [66]

Google Maps API V3 pruža vrlo efikasan mehanizam za isporuku digitalnih kartografskih informacija Internet korisnicima sa brzim vremenom odziva i interakcijom prilagođenom korisniku (engl. user friendly). Korištenjem standardne Google Maps "Map Type" kontrole, korisnik je u mogućnosti da odabere jedan od četiri tipa prikaza mapa: Roadmap, Satelite, Terrain i Hybrid.

Geokodiranje je integralni dio Google Maps API-ja. Za poslatu adresu vraća natrag koordinate te adrese. Google-ov servis za geokodiranje je javno dostupan, ali ima neka ograničenja, jer izvodi procesorski prilično zahtjevne operacije. To je trenutno ograničenje na 2.500 zahtjeva za geokodiranje sa jedne IP adrese unutar 24 sata. To nije beznačajan broj zahtjeva, tako da u većini slučajeva, to je više nego dovoljno. [76]

Sve funkcionalnost za obavljanje geokodiranja se nalazi u Geocoder objektu.

U cilju interakcije sa Google geocoder-om, svaki zahtjev se pravi kroz HTTP zahtjev prema Google serverima. Ovi serveri prevode ove parametre u URL zahtjev i vraćaju izlaz onako kako je specificirao korisnik. Parametri u zahtjevu su:

- q - Formatirana adresa koja će se geokodirati
- output - Željeni izlazni format (xml, kml, csv, or json)
- key - Google Maps API ključ

API funkcije se mogu koristiti s raznim programskim jezicima. Ove funkcije određuju izgled mape, uključujući razmjer, položaj, i još neke dodatne informacija u obliku tačaka, linija, ili područja.

Google Maps API za .NET je brza i jednostavna klijentska biblioteka za Google Maps API, koja pruža sve funkcionalnosti dostupne u Google Maps API-u. Razvijena je u C# programskom jeziku za .NET Framework 3.5 i na osnovu dosadašnjeg istraživanja pokazuje se najprikladnijom za zadatke geokodiranja koji će se obavljati u ovom projektu.

Mapiranje kriminaliteta je oživjelo interes i promijenilo svijest mnogih kriminologa o važnosti geografije kriminaliteta kao odrednice kriminala koja može biti jednako važna kao i kriminalna motivacija. [77], [78]

3.4 Geografsko profiliranje

Mjesto izvršenja krivičnog djela je jedan od prvih i najvažnijih dokaza na raspolaganju policiji u njihovim istragama, i može sadržavati važne detalje o počiniocu krivičnog djela (Rossmo, 1997). [79]

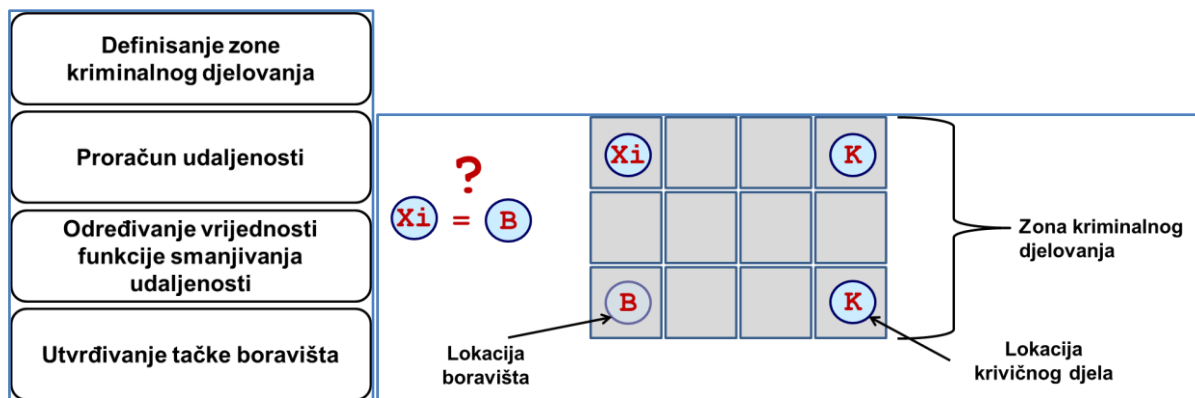
Geografsko profiliranje je kriminalistička istražna tehnika koja nastoji pružiti informacije o najvjerojatnijoj "bazi operacija" počinioca krivičnog djela na osnovu prostornih i vremenskih karakteristika niza povezanih krivičnih djela. Baza operacija može biti objekat stanovanja osumnjičenika, mjesto zaposlenja, rekreacije ili neka druga frekventna lokacija. Predviđanja su zasnovana na lokacijama ovih zločina i drugim geografskim informacijama o slučaju i osumnjičenom, kao i određene pretpostavke o putovanju osumnjičenika radi počinjenja krivičnog djela. (udaljenost između neke fiksne lokacije počinioca i lokacije na kojoj je izvršeno krivično djelo) [84] Vjerovatnoća prostornog ponašanja počinioca može se odrediti na osnovu informacija sadržanih na lokacijama mjesta zločina, njihove geografske povezanosti i demografskih karakteristika okolnih naselja. Geografsko profiliranje omogućava policijskim upravama da usmjere istražne radnje na prioritetna područja i osumnjičene, te koncentrišu njihove jedinice u one zone u kojima počinioci najčešće borave.

Paul i Patricia Brantingham 80-tih godina dvadesetog stoljeća proučavali su prostorno ponašanje počinitelja kako bi opisali lokacije na kojima se najčešće dešavaju zločini. Geografski profiliranje predstavlja na neki način pokušaj da se obrne ovaj model, koristeći lokacije počinjenja krivičnih djela kao osnovu za predviđanje najvjerojatnijeg područja boravišta počinitelja. [80], [81]

Istraživanje provedeno na Univerzitetu Sajmon Frejzer i odjeljenju Vankuverske policije, prateći ovaj pristup dovelo je do razvoja CGT algoritma (Criminal Geographic Targeting - Kriminalističko geografsko određivanje mete) koji je patentiran i ugrađen u Rigel softver za geografsko profiliranje, proizvođača Environmental Criminology Research Inc. (ECRI). Ovim algoritmom se analiziraju koordinate mjesta počinjenja krivičnih djela te se proizvodi površina koja predstavlja vjerovatnoću boravišta počinioca u području potrage. Trodimenzionalni prikaz ove vjerovatnoće se naziva površina ugrožavanja ili rizika (engl. jeopardy surface). Dvodimenzionalne prikaz integrisan sa mapama se naziva geoprofil.

CGT algoritam prati proces od četiri koraka:

1. Utvrđivanje granice mape područja kriminalnog djelovanja (engl. hunting area) proračunom na osnovu lokacija počinjenja krivičnih djela.
2. Za svaku tačku na mapi, određuje se Manhattan udaljenosti do svake lokacije gdje su počinjena krivična djela. (40000 piksela)
3. Udaljenost se koristi kao nezavisna varijabla funkcije za smanjivanje udaljenosti, ako je udaljenost manja od radijusa tampon zone, u suprotnom funkcija je negativna. Vrijednosti se izračunavaju za svaku lokaciju krivičnog djela (npr., 12 lokacija krivičnih djela jednako 12 vrijednosti za svaku tačku na mapi).
4. Ove vrijednosti se sabiraju da se proizvede konačan rezultat za svaku tačku na mapi. Što je veći konačan skor to je veća vjerovatnoća da je ta tačka lokacija boravišta počinioca krivičnih djela.



Slika 5. CGT algoritam

Slika 6. Proces geografskog profiliranja

Uspješan razvoj algoritma u velikoj mjeri je rad dr. Kim Rossmo, koji je uspostavio njegove osnovne principe. Rossmo je osmislio formulu geografskog profiliranja za predviđanje lokacije boravišta počinioca krivičnih djela.

Vjerovatnoća boravišta osumnjičenog u određenom sektoru (pikselu) je navedena kao:

$$P_{i,j} = \alpha \sum_{k=1}^N \left[\frac{\lambda}{\left(|X_i - x_k| + |Y_j - y_k| \right)^\xi} + \frac{(1-\lambda)(\beta^{\eta-\xi})}{\left(2\beta - |X_i - x_k| + |Y_j - y_k| \right)^\eta} \right] \quad [88]$$

P_{ij} je rezultujuća vjerovatnoća za tačku ij ;

λ je težinski faktor;

α je empirijski određena konstanta;

β je poluprečnik tampon zone;

N je broj lokacija krivičnih djela;

η je empirijski određen eksponent;

ξ je empirijski određen eksponent;

X_i, Y_j su koordinate tačke ij ;

x_k, y_k su koordinate mjesta k -tog krivičnog djela.

Suma je zbir N počinjenih krivičnih djela na lokacijama čije koordinate (x_k, y_k) ispunjavaju dva uslova. Fenomen smanjenja vjerovatnoće sa povećanjem udaljenost je sadržan u prvom uslovu (smanjenje broja krivičnih djela sa povećanjem udaljenosti od baze), a drugi uslov se odnosi na tzv. koncept "tampon zone" - neposrednog okruženje oko mjesta boravišta gdje počinioci izbjegavanju kriminalne aktivnosti. Ova formula pokazuje da vjerovatnoća počinjenja nekog krivičnog djela se povećava sa kretanjem od tampon zone prema aktivnim kriminogenim mjestima, ali se smanjuje nakon toga. [85], [86], [91]

CGT u suštini dijeli područje kriminalnog djelovanja u mrežu konačnog broja piksela i izračunava vjerovatnoću da svaki pojedini piksel sadrži lokaciju boravišta počinioca. (slika 6)

Međutim i dalje postoje određene vrste kriminala za koje podobnost geografskog profiliranja nije u potpunosti ispitana, a jedan od tih je kriminal povezan sa upotrebom elektronske pošte. [83], [92], [93]

Stoga je predmet ovog doktorskog rada istraživanje primjenjivosti i učinkovitosti geografskog profiliranja kriminala povezanog sa upotrebom elektronske pošte, odnosno kriminala na bazi e-maila.

Postojeće softverske aplikacije za geografsko profiliranje (najpoznatije: CrimeStat, DragNet, Predator, Rigel) se baziraju na različitim funkcijama za smanjivanje udaljenosti (engl. distance decay) i razvijene su za različite namjene i auditorij.

Njihovi nedostaci su: prvenstveno su namijenjene za istraživanje krivičnih djela nasilničkog kriminaliteta (serijska ubistva, serijska silovanja, serijska podmetanja požara, teroristički akti), zatim cijena njihova korištenja je prevelika za male policijske agencije (odnosi se na Rigel) dok ostale softverske aplikacije nemaju u potpunosti razvijeno korisničko okruženje jer su prije svega namijenjene za istraživače. [94]

4. Cilj istraživanja i fokus rješenja problema

Na osnovu teorijskih razmatranja i empirijskih istraživanja definisati model za poboljšano geolociranje počinioca sajber kriminala baziran na geografskom profiliranju i upotrebi informacija o lokacijama na kojima su počinjena ova krivična djela.

Novi pristup u modeliranju geolociranja počinilaca sajber kriminala će zahtijevati uvođenje određenih modifikacija, adaptacija, integracija i identifikacija.

Identifikacija:

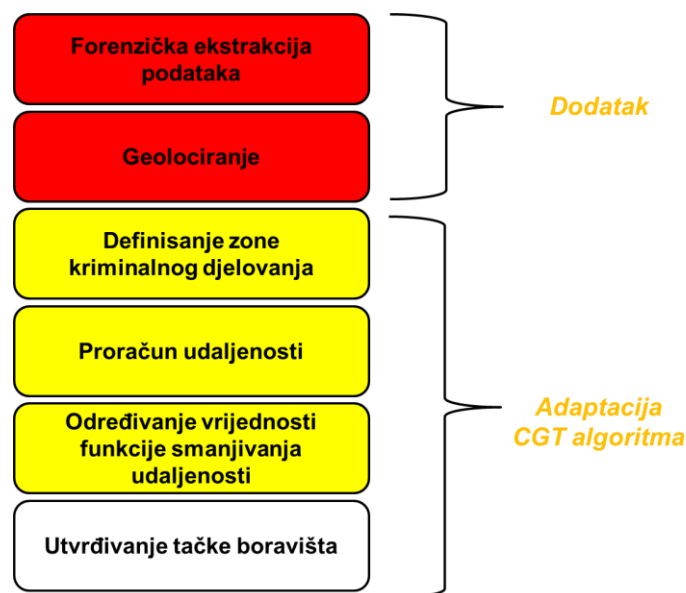
- Metapodataka u zaglavlju e-maila koji utiču na tačnost utvrđivanja originalnog izvora e-maila
- Metode za geolociranje IP adresa, bazirane na WBG strategiji za geolociranje Internet hostova i dajući prednost javno dostupnim i besplatnim geolokacijskim servisima i bazama podataka.
- Parametara modela čije podešavanje ključno utiče na efikasnost CGT algoritma za procesiranje ovog tipa kriminala

Adaptacija:

- Postojećih metoda za e-mail forenziku
- Parametara CGT algoritma za geografsko profiliranje kriminala na bazi e-maila
- Postojećih servisa za online geolociranje i mapiranje IP adrese

Integracija:

- Metode za e-mail forenziku
- Servisa za online geolociranje i mapiranje IP adrese



Slika 5. Novi model

5. Metode istraživanja

U skladu s postavljenim ciljevima, metode istraživanja se zasnivaju na teoretskim razmatranjima, te eksperimentalnoj provjeri dobivenih rezultata. U radu će se istraživati tehnike, metode i postupci temeljeni na analizi metapodataka u zaglavlju e-maila s kojima se može znatnije unaprijediti proces istrage sajber kriminala, predlažući rješenja za poboljšano geolociranje počinioaca krivičnih djela sajber kriminala. Takođe neophodna je upotreba matematičkih i statističkih metoda analize prostornog ponašanja kriminala, algoritama za analizu SMTP putanje, odgovarajućih metrika za mjerenje udaljenosti između tačaka (Euklidova ili Manhattan metrika) i funkcije za smanjivanje udaljenosti (engl. distance decay). Istraživanje će obuhvatiti i implementaciju metoda i tehnika pomoću softverskih razvojnih okruženja, geolokacijskih servisa i baza podataka, kao što su Microsoft .NET, Google Maps API V3, RIPE i InfoDB.

Realizacijom definisanog modela unutar softverske aplikacije će se provjeriti valjanost pristupa i rezultata istraživanja i doći do jednog koherentnog alata koji mogu koristiti istražitelji sajber kriminala.

Rezultati istraživanje neće biti ocjenjeni samo na osnovu laboratorijskih eksperimenata nego i eksperimentisanje na realnom sistemu, gdje će biti moguće napraviti određena poboljšanja na osnovu povratnih informacija od strane stručnjaka iz policijskih agencija naše zemlje i okruženja.

6. Plan istraživanja

Prije svega je potrebno eksperimentalno i teorijski, u realnom okruženju, identificirati metapodatke u zaglavlju e-maila koji mogu biti ključni za praćenje i utvrđivanje lokacije pošiljaoca zlonamjernih e-maila. Nakon toga slijedi istraživanje proces geolociranja i mapiranja niza povezanih krivičnih djela počinjenih upotrebom e-maila u skladu sa metodologijom geografskog profiliranja.

Plan istraživanja bi se u najopštijem smislu mogao predstaviti u nekoliko koraka:

- Akvizicija podataka
- Predprocesiranje podataka
- Izrada modela
- Eksperimentalno podešavanje parametara modela
- Implementacija
- Validacija u laboratorijskim uslovima
- Validacija u u realnim eksploatacionim uslovima

U radu će biti dat intenzivan fokus na geografiju kriminaliteta, pružajući kvantitativnu i kvalitativnu procjenu kriminala vezanog za prostorno-vremensko ponašanje. Prikaz sličnih dosadašnjih istraživanja će poslužiti kao poređenje pri analizi dobivenih rezultata.

7. Očekivani izvorni naučni doprinos disertacije

Očekivani izvorni naučni doprinos rada je razvoj novog model vizuelizacije i analize uzoraka kriminala na bazi geografskog profiliranja koji će pomoći razumijevanju i poboljšanju postojećih metoda istrage i sistema za kontrolu sajber kriminala.

Očekivani elementi doprinosa su:

- Definisanje novog pristupa za istraživanje kriminala na bazi e-maila pomoć geografskog profiliranja, kroz objedinjavanje teorijskih i praktičnih znanja i integraciju svih funkcionalnosti potrebnih za obradu podataka datih krivičnih djela.
- Adaptacija primijenjenih tehnika IP geolociranja koje će doprinijeti poboljšanju pouzdanosti i tačnosti lociranja IP adrese hosta.
- Adaptacija primijenjenih metoda e-mail forenzike koje će doprinijeti poboljšanju rezultata analize e-mail zaglavlja.

8. Polazna literatura

- [1] Nirkhi, S.M. ; Dept. of Comput. Sci., G.H.R.C.E., Nagpur, India ; Dharaskar, R.V. ; Thakre, V.M., *Analysis of online messages for identity tracing in cybercrime investigation*, Cyber Security, Cyber Warfare and Digital Forensic (CyberSec), 2012 International Conference, June 2012, 300 – 305
- [2] Michael Leitner, *Crime Modeling and Mapping Using Geospatial Technologies*, 2013
- [3] Quist-Aphetsi Kester, *Visualization And Analysis Of Geographical Crime Patterns Using Formal Concept Analysis*, International Journal Of Remote Sensing And Geoscience 2013
- [4] International Association of Crime Analysts (IACA), *Crime Pattern Definitions For Tactical Analysis*, 2011
- [5] K. K. Sindhu, B. B. Meshram, *Digital Forensic Investigation Tools and Procedures*, IJCNIS Vol.4, No.4, May 2012
- [6] Hong Guo, Bo Jin , Wei Qian, *Analysis of Email Header for Forensics Purpose*, CSNT '13 Proceedings of the 2013 International Conference on Communication Systems and Network Technologies, 340-344
- [7] Sridhar N., D.Lalitha Bhaskari, P.S.Avadhani, *A Novel Graph Model For E-Mail Forensics: Evidence Activity Analysis Graph*, International Journal of Engineering Science and Technology, 2013
- [8] Pranjal S. Bogawar, Kishor. K. Bhoyar, *E-mail Mining A Review*, IJCSI International Journal of Computer Science 2012
- [9] Uttam Mande, Y.Srinivas, J.V.R.Murthy, *Criminal Mapping Based on Forensic Evidences Using Generalized Gaussian Mixture Model*, 2012
- [10] Richard Wortley, *Environmental Criminology And Crime Analysis*, 2008
- [11] Daniel A. Keim, *Information Visualization and Visual Data Mining*, IEEE Transactions On Visualization And Computer Graphics, Vol. 7, No. 1, 2012

- [12] Raffael Marty, *Applied Security Visualization*, 2008
- [13] Butković Asmir, Mrdović Saša, Mujačić Samra, *IP geolocation suspicious e-mail messages*, Telfor 2013
- [14] Aurel Gross, *Using Geolocation In Authentication And Fraud Detection For Web-Based Systems*, 2011
- [15] Guanli Huang, *Dynamic Analysis for Geographical Profiling of Serial Cases Based on Bayesian-Time Series*, Journal of Software 2012
- [16] Srikanth Palla , Ram Dantu, João W. Cangussu, *Spam Classification Based On E-Mail Path Analysis*, 2011
- [17] Ickin Vural, Hein S. Venter, *Investigating Identity Concealing And E-mail Tracing Techniques*, 2009
- [18] Chuck Easttom, Jeff Taylor, *Computer Crime, Investigation, and the Law*, 2010
- [19] K. K. Sindhu, B. B. Meshram, *Digital Forensics And Cyber Crime Datamining*, Journal of Information Security, 2012
- [20] Ferebee, D., Dasgupta, D. Schmidt, M. Qishi Wu, *Security Visualization: Cyber Security Storm Map and Event Correlation*, Computational Intelligence in Cyber Security (CICS), 2011 IEEE
- [21] H. Jahankhani, Ameer Al-Nemrat , *Examination of Cyber-criminal Behaviour*, International Journal of Information Science and Management (IJISM) 2010
- [22] O. B. Longe, V. Mbarika, M. Kourouma, F. Wada, R. Isabalija, *Seeing Beyond The Surface Understanding And Tracking Fraudulent Cyber Activities*, International Journal of Computer Science and Information Security, IJCSIS 2010
- [23] Tommy Stallings, Brad Wardman, Gary Warner, and Sagar Thapaliya, *Whois Selling All The Pills*, International Journal of Forensic Computer Science 2012
- [24] Paglierani, J. , Mabey, M. , Ahn, G.-J., *Towards comprehensive and collaborative forensics on email evidence*, Collaborative Computing: Networking, Applications and Worksharing (Collaboratecom), 2013
- [25] Sobiya R. Khan, Smita M. Nirkhi, R. V. Dharaskar, *E-mail Data Analysis for Application to Cyber Forensic Investigation using Data Mining*, NCIPET 2013
- [26] Gregory Roussas, *Visualization Of Client-Side Web Browsing And E-mail Activity*, Master's Thesis. 2007
- [27] Samuel C. McQuade, *Encyclopedia of Cybercrime* , 2008
- [28] U.S. Department of Justice, Office of Justice Programs, National Institute of Justice, *Investigations Involving the Internet and Computer Networks*, 2012
- [29] Satheesaan Pasupatheeswaran, *E-mail 'Message-IDs' helpful for forensic analysis?*, Australian Digital Forensics Conference, Security Research Institute Conferences 2008
- [30] Timothy Rooney, *IP Address Management Principles and Practice*, 2011
- [31] Soren Riise, Devesh Patel, *Method of determining geographical location from IP address information*, United States Patent 2010
- [32] Venkata N. Padmanabhan, Lakshminarayanan Subramanian, *An Investigation of Geographic Mapping Techniques for Internet Hosts*, SIGCOMM 2001, pp. 173-185
- [33] Barry Leiba, Joel Ossher, V. T. Rajan, Richard Segal , Mark Wegman, *SMTP Path Analysis*, CEAS 2005
- [34] Fernando Sanchez, Zhenhai Duan, *A Sender-Centric Approach to Detecting Phishing Emails*, International Conference on Cyber Security, 2012
- [35] M. Tariq Bandy, *Techniques And Tools For Forensic Investigation Of E-Mail*, International Journal of Network Security & Its Applications (IJNSA), Vol.3, No.6, November 2011
- [36] Li Ding, *A Data Mining Approach To Identify Perpetrators: An Integration Framework And Case Studies*, A Dissertation 2010
- [37] Van Staden; Shlomo Hershkop; Eleazar Eskin, *Adding digital forensic readiness to the email trace header*, Information Security for South Africa (ISSA), 2010
- [38] Farkhund Iqbal, Liaquat A. Khan, Benjamin C. M. Fung, Mourad Debbabi, *E-mail Authorship Verification for Forensic Investigation*, Proceeding SAC '10 Proceedings of the 2010 ACM Symposium on Applied Computing Pages 1591-1598
- [39] Bogawar, Pranjal S.; Bhoyar, Kishor K., *E-mail Mining A Review*, International Journal of Computer Science Issues (IJCSI) . Jan2012, Vol. 9 Issue 1, p429-434. 6p.
- [40] B. Huffaker, M. Fomenkov, k. Claffy, *Geocompare: a comparison of public and commercial geolocation databases*, Cooperative Association for Internet Data Analysis (CAIDA), May 2011
- [41] Srikanth Palla , Ram Dantu, João W. Cangussu, *Spam Classification Based On E-Mail Path Analysis*, 2011
- [42] Srikanth Palla and Ram Dantu, *Detecting Phishing in E-mails*, IT Spam Conference, pp.1-7, MIT, Boston, March 2006
- [43] Patrick Dwyer , Zhenhai Duan, *MDMap: Assisting Users in Identifying Phishing Emails*, CEAS 2010
- [44] Omar Al-Jarrah, Ismail Khater, Basheer Al-Duwairi, *Identifying Potentially Useful E-mail Header Features for E-mail Spam Filtering*, IARIA, 2012

- [45] Wan Chung Cary Ho, *E-Mail Forensics: Tracing and Mapping Digital Evidence from IP Address*, Australian Digital Forensics Conference, thesis 2010
- [46] Li Zhuang , John Dunagan , Daniel R. Simon , Helen J. Wang , Ivan Osipkov , Geoff Hulten , J. D. Tygar , *Characterizing Botnets from E-mail Spam Records*, LEET'08 2008
- [47] Sarwat Nizamani, Nasrullah Memon, Uffe Kock Wiil, Panagiotis Karampelas, *Modeling Suspicious E-mail Detection Using Enhanced Feature Selection*, IJMO 2012 Vol.2(4): 371-377
- [48] Muallem, A., Shetty, S., Hargrove, S.K, *Visualizing Geolocation of Spam Email*, Computing Communications and IT Applications Conference (ComComAp), 2013
- [49] Robert K., Mario G., Gabi D. Rodosek, *Advanced Geolocation of IP Addresses*, World Academy of Science, Engineering and Technology International Journal of Electrical, Electronic Science and Engineering Vol:7 No:8, 2013
- [50] Ziqian Donga, Rohan D.W. Pererab, Rajarathnam Chandramoulib, K.P. Subbalakshmi, *Network measurement based modeling and optimization for IP geolocation*, Journal Computer Networks: The International Journal of Computer and Telecommunications Networking archive Volume 56 Issue 1, 2012
- [51] IANA (Internet Assigned Numbers Authority) URL <http://www.iana.org/>
- [52] Gueye, B., Ziviani, A., Crovella, M. , Fdida S., *Constraint-Based Geolocation of Internet Hosts*, IEEE/ACM 2006
- [53] P. Gill, Y. Ganjali, B. Wong , *Dude, where's that IP circumventing measurement-based IP geolocation*, Usenix Security Symposium, Aug 2010
- [54] Shavitt, Y. , Zilberman, N., *A Geolocation Databases Study*, IEEE Journal on Selected Areas in Communications - JSAC 2011
- [55] Singh, Satinder Pal, *IP Geolocation in Metropolitan Areas*, ACM SIGMETRICS 2011 Pages 347-348
- [56] Jeffrey Carr, *Inside Cyber Warfare 2nd Edition*, 2011
- [57] Clement A., *IXmaps – Tracking your personal data through the NSA's warrantless wiretapping sites*, Technology and Society (ISTAS), 2013 IEEE
- [58] Ingmar Poese, Mohamed Ali Kaafar, Benoit Donnet, Bamba Gueye, *IP geolocation databases: unreliable?*, ACM SIGCOMM Computer Communication Review, Volume 41 Issue 2, April 2011 Pages 53-56.
- [59] A. Butkovic, F. Orucevic, A. Tanovic, *Using Whois Based Geolocation and Google Maps API for support cybercrime investigations*, WEAS pp. 194-200, 2013
- [60] Brian Eriksson , Paul Barford , Joel Sommers , Robert Nowak, *A Learning-based Approach for IP Geolocation*, PAM'10 Proceedings of the 11th international conference on Passive and active measurement 2010
- [61] Ziqian Donga, Rohan D.W. Pererab, Rajarathnam Chandramoulib, K.P. Subbalakshmi, *Network measurement based modeling and optimization for IP geolocation*, Journal Computer Networks: The International Journal of Computer and Telecommunications Networking archive Volume 56 Issue 1, 2012
- [62] Gaurav Singhal, S. R. Tandan, Rohit Miri, *LAA (Internet Access Account) Based Security Modal For Detection And Prevention Of Cyber Crime*, International Journal of Engineering Research & Technology, 2013
- [63] Chuanxiong Guo, Yunxin Liu, Wenchao Shen, Helen Wang, Qing Yu, and Yongguang Zhang, *Mining the Web and the Internet for Accurate IP Address Geolocations*, IEEE INFOCOM 2009
- [64] Feng-Yu Lin, Yeali S. Sun, Meng Chang Chen, *Forensics Tracking for IP User using the Markov Chain Model*, Applied Mathematics & Information Sciences An International Journal 2013
- [65] Ethan Katz-bassett , John P. John , Arvind Krishnamurthy , David Wetherall , Thomas Anderson , Yatin Chawathe, *Towards IP Geolocation Using Delay and Topology Measurements*, ACM SIGCOMM 2006
- [66] Gabriel Svennerberg, *Beginning Google Maps API 3*, 2010
- [67] Jerry H. Ratcliffea, *Geocoding Crime And A First Estimate Of A Minimum Acceptable Hit Rate*, International Journal of Geographical Information Science Volume 18, Issue 1, 2004
- [68] Patricia Takako Endo, Djamel Fawzi Hadj Sadok, *Whois Based Geolocation a strategy to geolocate*, Advanced Information Networking and Applications (AINA), 2010
- [69] Kevin Curran, *Bringing location to IP Addresses with IP Geolocation*, JETWI - Journal of Emerging Technologies in Web Intelligence 2012
- [70] Sherri Davidoff, *Network Forensics - Tracking Hackers Through Cyberspace*, 2012
- [71] Nagender P Vedula, Aditya G Bhandarkar, Dharma K Shukla, William R Taylor, *Mapping Tool Graphical User Interface*, United States Patent 2010

- [72] Ratcliffe, J. H., *Crime Mapping And The Training Needs Of Law Enforcement*, Crime mapping and the training needs of law enforcement. European Journal on Criminal Policy and Research 2004
- [73] Ratcliffe, J. H., *Crime Mapping And The Training Needs Of Law Enforcement*, Crime mapping and the training needs of law enforcement. European Journal on Criminal Policy and Research 2004
- [74] Michael Peterson, *Choropleth Google Maps*, Cartographic perspectives Number 60, Spring 2008
- [75] Venkata N. Padmanabhan, Lakshminarayanan Subramanian, *An Investigation of Geographic Mapping Techniques for Internet Hosts*, SIGCOMM 2001, pp. 173-185
- [76] R. Hofstede, T. Fioreze, *SURFmap a network monitoring tool based on the Google maps API*, Integrated Network Management, 2009
- [77] Shunfu Hu, Ting Dai, *Online Map Application Development Using Google Maps Api, Sql Database, And ASP.NET*, International Journal of Information and Communication Technology Research 2013
- [78] Syngress, Ed Tittel, *Scene of the Cybercrime - Computer Forensics Handbook*, 2002
- [79] Amit Nithianandan, *Cs 6604 Project Proposal- Spatiotemporal Visualization And Analysis Of Crime Data*, Virginia Tech 2008
- [80] Franz-Josef Behr, Astrit Rimayanti, Hui Li, *Opengeocoding.org: A Free, Participatory, Community Oriented Geocoding Service*, XXI Congress of the International Society for Photogrammetry and Remote Sensing, 2008
- [81] Thangavelu A, Sathyaraj S. R., Balasubramanian S., *Assessment of Spatial Distribution of Rural Crime Mapping in India: A GIS Perspective*, International Journal of Advanced Remote Sensing and GIS, Volume 2, Issue 1, pp. 69-84, 2013
- [82] Jeff Brantingham, Andrea Bertozzi, *Algorithm Development for Criminal Geographic Profiling*, LAPD Project Description 2009
- [83] Jerry Ratcliffe, *Crime Mapping Spatial And Temporal Challenges*, Handbook of Quantitative Criminology 2010, pp 5-24
- [84] Tonkin M, Woodhams J, Bond JW, Loe T., *A Theoretical and Practical Test of Geographical Profiling with Serial Vehicle Theft in a U.K. Context*, Behavioral Sciences and the Law, 28, 442-460 2010
- [85] Alessio Papini, Stefano Mosti, Ugo Santosuosso, *Tracking the origin of the invading Caulerpa (Caulerpa, Chlorophyta) with Geographic Profiling, a criminological technique for a killer alga*, Biological Invasions 2013, Vol. 15 Issue 7, p1613
- [86] Mike O'Leary, *The Mathematics of Geographic Profiling*, Journal of Investigative Psychology and Offender Profiling Special Issue: Bayesian Journey to Crime Modelling Volume 6, Issue 3, pages 253–265, October 2012
- [87] Feroz Shah Syed, Didong Li, Xun Zhang, Zhenhong Guo, *Mathematical Modelling in Criminology*, Malaysian Journal of Mathematical Sciences 2013
- [88] Feroz Shah Syed, Didong Li, Xun Zhang, Zhenhong Guo, *Mathematical Modelling in Criminology*, Malaysian Journal of Mathematical Sciences 2013
- [89] Wesley J. English, *Geoprofile: Developing And Establishing The Reliability Of A New Geographic Profiling Software System*, Paper presented at the annual meeting of the American Psychology - Law Society, TBA, San Antonio 2009
- [90] Karla Emeno, Craig Bennell, *The effectiveness of calibrated versus default distance decay functions for geographic profiling: a preliminary examination of crime type*, Psychology, Crime & Law Volume 19, Issue 3, 2013
- [91] Mike O'Leary, *The Mathematics of Geographic Profiling*, Journal of Investigative Psychology and Offender Profiling Special Issue: Bayesian Journey to Crime Modelling Volume 6, Issue 3, pages 253–265, October 2012
- [92] Aida Cacan, *Geografsko profiliranje počnilaca u kriminalističkoj istrazi serijskih krivičnih djela*, Kriminalističke teme - Godište X, Broj 3-4, 2010
- [93] Ksenija Butorac, *Geografija kriminaliteta – kriminološki i kriminalistički diskursi*, Polic. sigur. (Zagreb), godina 20. (2011), broj 3, str. 363-379
- [94] A. Thangavelu, S. R. Sathyaraj and S. Balasubramanian, *Spatial distribution of crime in Coimbatore rural area: A GIS Analysis*, SASCV 2013