Centar Informacijske Sigurnosti

Otisak web preglednika

srpanj 2012.

CIS-DOC-2012-07-054

Upozorenje

Podaci, informacije, tvrdnje i stavovi navedeni u ovom dokumentu nastali su dobrom namjerom i dobrom voljom te profesionalnim radom CIS-ovih stručnjaka, a temelje se na njihovom znanju i petnaestak godina iskustva u radu u informacijskoj sigurnosti. Namjera je da budu točni, precizni, aktualni, potpuni i nepristrani.

Ipak, oni su dani samo kao izvor informacija i CIS ne snosi nikakvu izravnu ili posrednu odgovornost za bilo kakve posljedice nastale korištenjem podataka iz ovog dokumenta.

Ukoliko primijetite bilo kakve netočnosti, krive podatke ili pogreške u ovom dokumentu, ili imate potrebu komentirati sadržaj molimo Vas da to javite elektroničkom poštom na adresu info@CIS.hr.

O CIS-u

CIS izrađuje pregledne dokumente (eng. white paper) na teme iz područja informacijske sigurnosti koji će biti korisni zainteresiranoj javnosti, a u svrhu **podizanje njezine svijesti o informacijskoj sigurnosti** i **sposobnosti za čuvanje i zaštitu informacija i informacijskih sustava**. Pored toga, CIS razvija i održava mrežni portal www.CIS.hr kao referalnu točku za informacijsku sigurnost za cjelokupnu javnost; izrađuje obrazovne materijale namijenjene javnosti; organizira događaje za podizanje svijesti o informacijskoj sigurnosti u javnosti i pojedinim skupinama te djeluje u suradnji sa svim medijima.

CIS **okuplja mlade** zainteresirane za informacijsku sigurnost i radi na njihovom pravilnom odgoju i obrazovanju u području informacijske sigurnosti te pripremu za **profesionalno bavljenje informacijskom sigurnošću**.

Centar informacijske sigurnosti [CIS] nastao je 2010. godine na poticaj Laboratorija za sustave i signale[LSS] Zavoda za elektroničke sustave i obradbu informacija Fakulteta elektrotehnike i računarstva Sveučilišta u Zagrebu, a kao posljedica 15togodišnjeg rada na istraživanju, razvoju i primjeni informacijske sigurnosti. LSS je među ostalim potaknuo osnivanje CARNetovog CERTa i sudjelovao u izradi Nacionalnog programa informacijske sigurnosti RH.

Smisao CISa je da bude **referentno mjesto za informacijsku sigurnost** za javnost, informatičare i posebno za mlade te da sustavno podiže njihovu svijest i sposobnosti u području informacijske sigurnosti.

Rad CISa podržava Ministarstvo znanosti, obrazovanja i sporta Republike Hrvatske, a omogućuju sponzori.

Prava korištenja

Sadržaj

1. Uvod

Prisutnost osobe (ili životinje) na određenom području najčešće se otkriva pomoću nekih tragova koje je ta osoba ostavila iza sebe. Tragovi mogu biti ostaci hrane, otisci stopala, bačeno smeće itd. Isto tako, poznato je da ljudi mogu ostaviti otisak svojih prstiju prilikom dodira nekog predmeta. Za takve tragove možemo reći da su fizički jer se oni javljaju u materijalnom obliku, te ih je vrlo često moguće vidjeti i opipati.

Analogno tome, postavlja se pitanje ostavljamo li tragove i u digitalnom svijetu. Naravno, u ovome slučaju radi se o digitalnome tragu koji nije nekakav stvarni predmet. Upravo zbog te činjenice digitalni tragovi su nezgodni, jer osoba koja ih ostavlja najčešće uopće toga nije ni svjesna. U ovome dokumentu opisati će se kakve to tragove ostavlja web preglednik prilikom posjete nekoj web stranici te može li takav otisak biti jedinstven poput onoga kojeg ostavljamo dodirom ruke.

Treba napomenuti kako se ovdje radi o otisku web preglednika, a ne o otisku osobe koja koristi web preglednik. Isti web preglednik može koristiti više ljudi, no njegov otisak bi trebao ostati isti neovisno o tome koja osoba ga koristi.

Na prvi pogled ideja otiska web preglednika (eng. browser fingerprinting) izgleda poprilično nevjerojatno zbog malog broja različitih preglednika i njihove slične funkcionalnosti. Postoji pet Internet preglednika koji se najviše koriste, a to su: Internet Explorer, Mozilla Firefox, Google Chrome, Opera i Safari. Osim toga postoje još i njihove inačice za mobilne telefone. Čak i ako u to uključimo različite operacijske sustave koji pokreću te preglednike svejedno je poprilično teško zamisliti da svaki Internet preglednik na svijetu (broj u milijardama) može ostaviti jedinstveni otisak.

Pokazalo se da priča ipak nije toliko jednostavna. Kao što će se vidjeti u nastavku dokumenta, web preglednik tokom rada ostavlja puno više informacija nego što bi se to intuitivno reklo. S obzirom na to da su te informacije varijabilne, smanjuje se broj web preglednika s identičnim otiskom. Zašto je to uopće bitno?

Iako otisak web preglednika nije otisak osobe koja koristi web preglednik, pretpostavlja se da najčešće jedan web preglednik koristi jedna (ili manji broj) osoba. Dakle, ukoliko možemo prepoznati web preglednik vrlo često možemo prepoznati i osobu. Jasno je da je uz pomoć takve tehnologije moguće pratiti osobu svaki puta kada sjedne za svoje računalo. Na taj način je moguće napraviti bolji profil korisnika i ponuditi mu sadržaje kakvi ga zanimaju, te samim time i poboljšati kvalitetu usluge (eng. quality of service). S druge stane, svaki korisnik ima pravo zapitati se krše li se time njegova osnovna prava o privatnosti. Obzirom na to da se ne zna kada je netko (i tko) uzeo otisak možemo se pitati prate li nas već sada. Znaju li velike kompanije već danas više o nama nego što mi to mislimo? Kako je moguće zaštititi se?

Iako postoje istraživanja koja se bave temom otiska web preglednika teško je sa sigurnošću odrediti u kolikoj mjeri se to trenutno koristi u Internet svijetu. Kroz drugo i treće poglavlje detaljno se opisuje ideja otiska web preglednika i način na koji se može ostvariti. Jedan takav algoritam za ostvarivanje otiska web preglednika razvio se u sklopu Panopticlick projekta, te je opisan u četvrtom poglavlju. Peto i šesto poglavlje se bave posljedicama koje ova tehnologija ostavlja i načinom na koji se je moguće zaštititi od njezinog korištenja u zlonamjerne svrhe.

2. Principi identifikacije u digitalnom sustavu

2.1. Princip združenih informacija

Zamislite da trebate osobu imena Ivan Horvat. Samo ime te osobe ne govori previše zbog toga što je vrlo često na ovim područjima. Niti podatak da je ta osoba u Hrvatskoj ne pomaže mnogo jer ionako postoji vrlo malo Ivana Horvata izvan Hrvatske. No, ukoliko saznate da je tražena osoba Ivan Horvat iz Gospića, broj ljudi koji bi mogli odgovarati tome je podosta smanjen (u odnosu na cijelu Hrvatsku) iako i dalje postoji vjerojatnost da u Gospiću postoje dva Ivana Horvata. Ukoliko nakon toga saznate i datum rođenja tražene osobe, tada već s dosta velikom sigurnošću možete zaključiti o kojoj osobi se radi.

Važno je primijetiti da nam te informacije (ime i prezime, prebivalište, datum rođenja) same za sebe ne pomažu puno, no ako ih združimo one nas vode u određenome smjeru. Čak i ako nemamo dovoljno informacija da na kraju možemo sa sigurnošću točno odrediti traženu osobu, broj potencijalnih traženih osoba je smanjen u odnosu na početni broj.

Kod takvih informacija važno je pronaći onu koja najviše sužava izbor. Na primjer, ukoliko se traži osoba imena Ivan Horvat i netko kaže da je ta osoba muško, nije zapravo puno pomogao jer ne postoje osobe Ivan Horvat koje nisu muško. No, ukoliko saznate da je ta osoba iz Šiškovaca (maleno selo u Vukovarsko-srijemskoj županiji) broj potencijalnih traženih osoba je sveden na nekolicinu (ili na točno jednu). Iz ovih primjera vidljivo je kako neke informacije ukoliko se rijetko pojavljuju mogu skoro pa same dovesti do traženih rezultata.

Sličan princip može se primjenjivati i kod Internet preglednika. Važno je uočiti da nam najčešće jedna informacija koju nudi web preglednik neće biti dovoljna za njegovo jednoznačno određivanje, no više kombiniranih informacija sasvim sigurno će svesti broj identičnih preglednika na vrlo mali broj. Inačica web preglednika, operacijski sustav i rezolucija samo su neke od informacija koji utječu na različitost otiska web preglednika. Postoje još mnogo takvih podataka od kojih se većina nalazi u pozadini sustava i korisniku nije toliko očita. Neke informacije više govore od drugih, kao što je prikazano u primjeru s Ivanom Horvatom. Informacija da korisnik koristi Safari kao web preglednik „vrijedi“ više od informacije da korisnik koristi Internet Explorer zbog činjenice da Internet Explorer koristi više ljudi nego Safari. Zbog toga je važno prikupiti što više podatka koji se što rjeđe pojavljuju kako bi mogli napraviti što veću diferencijaciju. Kao što je prikazano u primjeru sa Šiškovcima, i u digitalnom svijetu postoje „netipične“ informacije koje se ne pojavljuju često i koje mogu uvelike pomoći u identificiranju web preglednika.

Slika 1. *Više poznatih informacija smanjuje broj jedinstvenih otisaka*
Izvor: *CIS*

Moguće je ovu priču s informacijama prikazati i grafičkim putem (Slika 1). Radi jednostavnosti uzete su tri osnovne informacije (inačica web preglednika, operacijski sustav, rezolucija ekrana) dok ih u stvarnosti ima puno više. Bitno je uočiti središnji dio dijagrama koji prikazuje presjek informacija. Taj dio prikazuje one preglednike koji imaju identične sve tri informacije. Ono što se uočava je da je taj dio dosta manji od početnoga (presjek sva tri kruga).

Više o informacijama koje se mogu prikupiti u trenutku kada web preglednik pristupa internetskoj stranici protumačeno je u idućem poglavlju ovog dokumenta[1].

2.2. Vrednovanje različitosti u digitalnom svijetu

Trenutno na svijetu živi nešto manje od sedam milijardi ljudi. Ukoliko bi željeli svakoj osobi dodijeliti jedinstveni broj koji će ga globalno karakterizirat bio bi nam potreban broj od deset znamenki od 0 do 9. U informatičkom svijetu se takvi podaci najčešće izražavaju u bitovima (dakle pomoću 1 i 0).

Neka se za primjer uzme jedan bit. On može biti 1 ili 0. S takvim bitom jednoznačno se može označiti 2 osobe (jedan dobije 0, a drugi 1). S dva bita može se jednoznačno označiti četiri osobe, zbog toga što postoje četiri kombinacije (00, 01, 10, 11). Pomoću formule 2^n moguće je dobiti broj ljudi koje je moguće označiti s n bitova. Formula je razumljivija ako kažemo da za svaki dodani bit povećavamo broj kombinacija za duplo. Za označiti sve ljude na svijetu potrebno nam je 33 bita, zbog toga što 2^{33} daje oko 8 milijardi kombinacija.

Broj različitih mogućnosti za neku varijablu naziva se entropija, te se najčešće označava u bitovima. Pojam entropija susreće se i u termodinamici, a u teoriju informacije uveo ju je Claude Shannon. [1] Radi se o matematičkoj definiciji, te zbog toga postoje različiti načini poimanja entropije. U ovom dokumentu se neće ulaziti u detalje i matematičku pozadinu entropije već je samo važno znati da entropija pokazuje različitost u vrijednostima neke varijable (npr. broj različitih ljudi na svijetu).

Za broj različitih ljudi na svijetu pokazali smo da postoje približno 33 bita entropije. Treba imati na umu da entropija nije cjelobrojna vrijednost te je moguće dobiti i decimalne vrijednosti bitova. Tako je točna vrijednost entropije za broj različitih ljudi na svijetu 32.7047 bita.

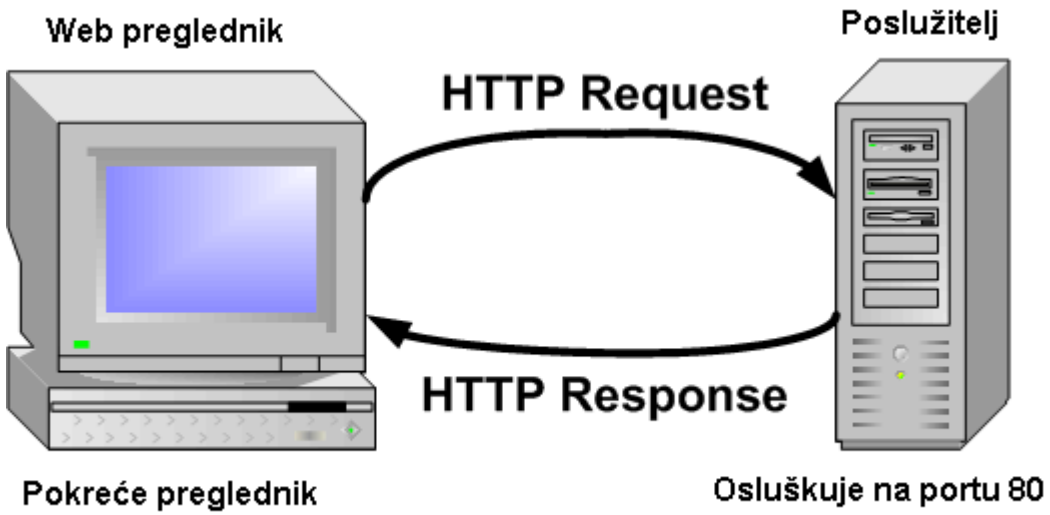
Ukoliko poznajete neku informaciju o nekoj osobi, broj ljudi za koje vrijedi ta informacija manji je od ukupnog broja ljudi kao što je objašnjeno u prethodnom poglavlju. Zbog toga je potreban manji broj bitova kako bi označili te ljude ili stručnije rečeno, broj bitova entropije se smanjuje. Entropija pomaže da se na jasan i jednostavan način matematički prikažu informacije o kojima se govorilo u prošlom odlomku. Isto tako, pomoću entropije moguće je izračunati koliko određena informacija „vrijedi“ odnosno koliko se smanjuje entropija ukoliko poznajemo neku informaciju. Detaljnije formule za izračunavanje entropije i bitova za koje se smanjuje biti će prikazane u poglavlju o matematičkoj pozadini otiska web preglednika.

U ovom poglavlju uzet je jednostavan primjer entropije na broju ljudi na svijetu, no analogno se može napraviti i s web preglednicima. Ključna stavka je već navedeno vrednovanje informacije, zbog činjenice da se entropija u slučaju otiska web preglednika želi najviše smanjiti.

3. Kako napraviti otisak web preglednika

3.1. Način rada web preglednika

Za razumijevanje načina na koji se radi otisak web preglednika prvo treba razumjeti na koji način web preglednik radi. Web preglednik je klijentska aplikacija (pokreće se na računalu korisnika) i služi za pregledavanje Internet sadržaja. Prilikom otvaranja Internet stranice web preglednik mora kontaktirati određeni poslužitelj (eng. *server*) na kojemu se nalazi traženi sadržaj.



Slika 2. Način rada web preglednika
Izvor:.Uregina

Unosom adrese web stranice (eng. *Uniform resource identifier*, URI) preglednik inicira komunikaciju sa poslužiteljem koji je odgovoran za tu adresu. Zadatak poslužitelja je osluškivati dolazne zahtjeve te ih posluživati određenim sadržajem. Komunikacija se odvija pomoću HTTP protokola (eng. *Hypertext Transfer Potocol*) koji je najčešće korišteni protokol u prijenosu informacija na Internetu. HTTP protokol definira različite metode kojima se obavlja komunikacija između preglednika i poslužitelja. Četiri osnovne metode su GET, POST, PUT i DELETE. Protokol definira da se GET metoda koristi za dohvat nekog udaljenog sadržaja, POST je namijenjen prijenosu parametara od klijenta prema poslužitelju, dok su PUT i DELETE namijenjeni za stvaranje odnosno brisanje resursa na poslužitelju. Prilikom dohvata Internet stranice prvo preglednik pošalje poslužitelju HTTP GET zahtjev za Internet stranicom. Nakon primitka zahtijeva poslužitelj šalje poruku o ispravnom primitku poruke, a nakon toga vraća klijentu traženi sadržaj. Iako je komunikacija na slici radi jednostavnosti svedena na jednog klijenta u stvarnosti najčešće više klijenata pristupa jednom poslužitelju.

Nakon primitka odgovora od poslužitelja web preglednik treba interpretirati ono što je dobio. Izgled Internet stranice definiran je jezikom HTML koji je zasnovan na oznakama (eng. *tag*) (npr. oznaka **
** predstavlja prelazak ispisa na novi red). Osim HTML-a, unutar odgovora može biti i Javascript kod. Javascript je skriptni jezik koji se interpretira unutar web preglednika i služi za proširivanje mogućnosti jezika HTML. Osim što omogućuje Internet stranicama da dinamički mijenjaju izgled, mogu izvoditi i neke pozadinske programe unutar preglednika. Jedna od najvažnijih mogućnosti Javascripta je asinkroni prijenos informacija (eng. *Asynchronous Javascript*, AJAX). AJAX omogućuje pregledniku da komunicira sa poslužiteljem u pozadini dok korisnik pregledava stranicu.

Upravo je Javascript „oružje“ pomoću kojega se može načiniti otisak web preglednika.

Prilikom slanja HTTP GET zahtijeva poslužitelju u zaglavlju poruke sadržane su neke informacije koje se mogu iskoristiti za stvaranje otiska preglednika[7].

```
GET / HTTP/1.
Host: www.google.com
Connection: close
User-Agent: Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US)
AppleWebKit/525.13 (KHTML, like Gecko) Chrome/0.2.149.29
Safari/525.13Accept-Charset: ISO-8859-1,UTF-8;q=0.7,*;q=0.7
Cache-Control: no-cache
Accept-Language: de,en;q=0.7,en-us;q=0.3
```

U navedenom primjeru HTTP GET zahtijeva moguće je iskoristiti polja poput User-Agent, Accept-Language i Cache-Control za stvaranje otiska web preglednika. Očito je da takvih informacija u samom zahtjevu nema dovoljno za napraviti jedinstvenu sliku preglednika pa se zbog toga iskorištava Javascript koji uz pomoć AJAX-a asinkrono dostavlja dodatne podatke poslužitelju. S obzirom na to da se sve odvija u pozadini, korisnik ne može primijetiti je li uzet otisak s njegovog preglednika[7].

3.2. Informacije korištene za stvaranje otiska web preglednika

3.2.1. Informacije prikupljene iz HTTP GET zahtijeva

Kao što je pokazano u prethodnom poglavlju, unutar HTTP GET zahtijeva nalaze se informacije koje mogu poslužiti za stvaranje otiska preglednika.

Najvažnije polje u zahtjevu je **User-Agent** polje. To polje u sebi sadrži informacije o inačici Internet preglednika, inačici operacijskog sustava, jeziku, te (opcionalno) još neke informacije o računalu korisnika. U sljedećem primjeru vidi se kako se User-Agent mijenja ovisno o operacijskom sustavu i inačici web preglednika.

```
Microsoft Windows XP

Internet Explorer 6:
Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)

Firefox 3.0:
Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.9.0.1)
Gecko/2008070208 Firefox/3.0.1

Firefox 2.0:
Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.8.1.14)
Gecko/20080404 Firefox/2.0.0.14

Linux (Ubuntu)

Firefox 2.0.0.19 (Ubuntu):
Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.8.1.19) Gecko/20081216
Ubuntu/8.04 (hardy) Firefox/2.0.0.19
```

Osim User-Agent polja, iz zahtjeva je još moguće pročitati informacije o omogućenim kolačićima (eng. *cookies*). Osim toga GET zahtjev može sadržavati i polja poput **Accept-Language** i **Accept-Charset,** te još neka **Accept** zaglavlja HTTP protokola iz kojih se pokušavaju izvući informacije kako bi se dobio što bolji otisak[6].

3.2.2. Kolačići

HTTP protokol je protokol koji ne pamti „stanja“ (tj. povijest interakcije klijent-poslužitelj). Dakle, nakon što se ostvari komunikacija i poslužitelj dostavi sadržaj stranice klijentu zatvara se komunikacija i poslužitelj više ne zna ništa o toj prošloj komunikaciji. Postoje situacije kada takvo rješenje nije idealno kao u slučaju autentikacije korisnika. Ne bi bilo praktično da korisnik svaki puta kada pristupa određenoj stranici mora ponovno upisati korisničko ime i lozinku. Zbog toga se uvode kolačići koji omogućuju da korisnik jednom upiše svoje autentikacijske podatke, te ih nakon toga više ne treba upisivati.

Za kolačiće možemo reći da su nizovi znakova spremljeni u datoteku na računalu korisnika. Pomoću kolačića moguće je sačuvati sjednicu sa poslužiteljem unatoč tome što HTTP protokol ne pamti stanja. Unutar kolačića spremljen je identifikator sjednice koji omogućuje poslužitelju da prepozna korisnika.

Kolačići se postavljaju prilikom HTTP odgovora poslužitelja klijentu u polju **Set-Cookie**. Nakon toga poslužitelj prilikom iduće komunikacije s tim poslužiteljem šalje u GET zahtjevu polje **Cookie**. Kolačić može isteći ili se promijeniti (ovisno o parametrima postavljenim prilikom postavljanja kolačića).

Svrha kolačića je zapravo ista kao i kod otiska web preglednika, a to je prepoznati korisnika. Zbog toga je pomoću kolačića moguće pratiti korisnikove aktivnosti na Internetu. Svaki web preglednik podržava opciju isključivanja kolačića kako bi onemogućio njihovo postavljanje i korištenje. To rješava problem praćenja korisnika, no zbog toga neke Internet stranice neće raditi ispravno (jer svoj rad temelje na poznavanju korisnika, tj. korištenju kolačića). Otisak web preglednika je još jedna moćnija metoda praćenja korisnika zbog toga jer ju nije jednostavno isključiti kao kolačić. No i sami kolačići mogu dati svoj doprinos otisku web preglednika. Čak i ukoliko isključimo kolačiće omogućili smo filtriranje preglednika samo na one koji imaju isključene i one koji imaju uključene kolačiće[8].

3.2.3. Superkolačići

Kao što i samo ime kaže, superkolačići (eng. *supercookies*) su naprednija inačica običnih kolačića. Naprednija u smislu da se identifikatori ne spremaju na isti način i na ista mjesta kao što se to radi kod običnih kolačića. Zbog toga ih nije moguće obrisati klasičnim metodama brisanja kolačića poput onih što ih nude web preglednici. U početku su superkolačići bili drugo ime za Flash kolačiće[1] zbog toga što su se spremali u prostoru namijenjenom za Flash sadržaj, no s vremenom su se superkolačići počeli spremati i na druga mjesta poput HTML5[2] spremničkih prostora i u arhivu namijenjenu Silverlight platformi. Zbog toga su vrlo robusni i najčešće ih je potrebno ručno brisati, što zahtjeva veću razinu informatičke pismenosti. Osim toga postoji vjerojatnost regeneracije superkolačića nakon što su bili djelomično obrisani zbog činjenice da mogu biti spremljeni na više različitih mjesta. Namjena im je slična kao i kod običnih kolačića, te se uglavnom odnosi na prepoznavanje korisnika. No, zbog svojih osobina superkolačići su odlična metoda praćenja i zbog toga su često na meti kritika. Iako je superkolačiće moguće koristiti kao zasebnu metodu praćenja korisnika, moguće ju je također i kombinirati s otiskom web preglednika kako bi se napravio što učinkovitiji algoritam za identifikaciju preglednika.

[1] Podaci koje stranice koje koriste Adobe Flash spremaju na korisničko računalo. Koriste ga sve inačice Adobe Flash Playera najčešće za spremanje korisničkih postavki. Moguće je iskoristiti prostor namijenjen Flashu za spremanje nekog drugog sadržaja što predstavlja opasnost za sigurnost web preglednika.

[2] Naziv za petu reviziju HTML standarda. Standard je još uvijek u izradi.

3.2.4. Rezolucija ekrana, Vremenska zona, Fontovi

Ove informacije nije moguće prikupiti iz HTTP zahtijeva nego se prikupljaju izravno s korisničkog računala. Kao što je već rečeno u prethodnom poglavlju postoji jezik Javascript pomoću kojeg je moguće prikupiti neke podatke s klijentskog računala. S obzirom na to da se Javascript izvodi unutar web preglednika, ima ograničeno područje djelovanja. No, bez obzira na to može prikupiti podatke poput rezolucije ekrana, vremenske zone (eng. *Timezone*) i korištenih fontova.

Rezolucija ekrana ovisna je o veličini i obliku korisnikovog ekrana. Pretpostavlja se da se monitori ne mijenjaju često te je zbog toga ovaj parametar vrlo stabilan.

Vremenska zona predstavlja informaciju o konkretnoj zoni na kojoj se nalazi korisnik. Postoji 24 različite vremenske zone i one odgovaraju geografskim vremenskim zonama.

Fontovi predstavljaju instalirane kolekcije znakova na klijentsko računalo. Ukoliko korisnik ima instalirane fontove s rjeđe korištenim znakovima diferencijacija između korisnika je veća[6].

3.2.5. Dodaci namijenjeni web preglednicima

Postoji čitav niz različitih dodataka (eng. plug-in) namijenjenih preglednicima. Dodaci proširuju mogućnost preglednika te omogućuju korisniku neke dodatne aktivnosti na Internetu. Takvi dodaci ne moraju davati dodatne informacije o korisniku, već je i samo njihovo postojanje dodatna informacija za ostvarivanje otiska. Zbog velikog broja postojećih dodataka moguće je napraviti vrlo velik broj različitih kombinacija korištenja istih. Zbog toga se smanjuje broj korisnika s istim skupom dodataka. Neki dodaci su češće korišteni u odnosu na neke druge, te je zbog toga lako prepoznati korisnike koje koriste specifične dodatke.

3.2.6. Ostale informacije pomoću kojih je moguće napraviti otisak

Teško je definirati sve informacije koje se koriste prilikom izrade otiska. U dosadašnjem dijelu dokumenta govorilo se o ideji otiska web preglednika, a ne o konkretnom algoritmu. Svaki algoritam ima svoje posebnosti i najčešće iskorištava samo neke od navedenih informacija koje su mu ponuđene.

Osim navedenih, moguće je iskorištavati i neke druge informacije koje je moguće saznati. Činjenica da web preglednik pristupa stranici preko posrednika (eng. *proxy*) se može iskoristiti za diferencijaciju. Isto tako, ukoliko preglednik ima onemogućenu neku od standardnih opcija poput prijenosa datoteka ili izvršavanje Javascrip koda spada u vrlo mali krug preglednika s istim postavkama, što je moguće kvalitetno iskoristiti u izradi otiska.

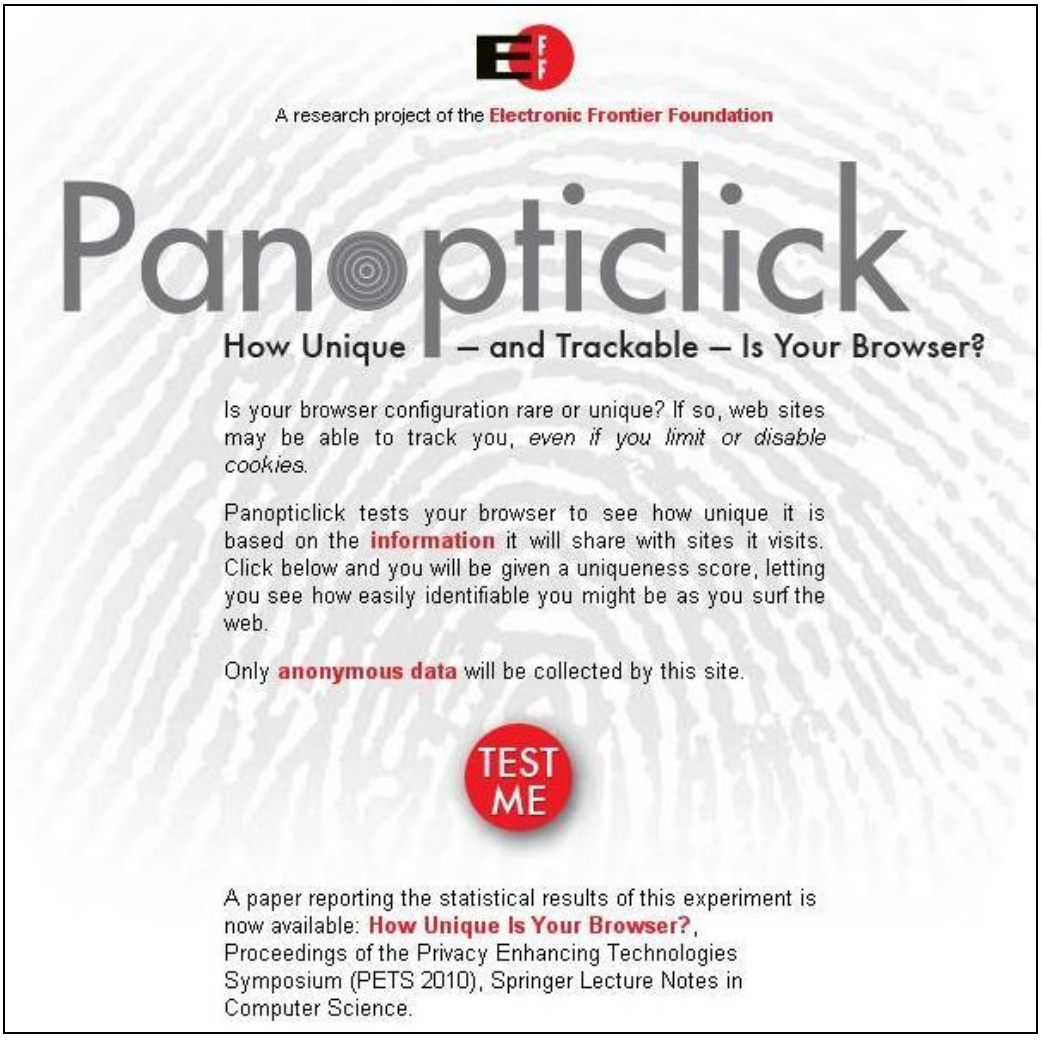
4. Projekt Panopticlick

Projekt Panopticlick pokrenula je internacionalna neprofitna organizacija Electronic Frontier Foundation (EFF). Organizacija EFF smještena je u San Franciscu (SAD), te joj je osnovna namjena promicanje osnovnih ljudskih prava bez obzira na tehnologiju. Projekt Panopticlick temeljen je na ideji otiska web preglednika obzirom na informacije koje je moguće prikupiti prilikom komunikacije. Cilj projekta je prikupiti što veći broj otisaka kako bi se, korištenjem ove analize, moglo zaključiti koliku opasnost privatnosti korisnika otisak predstavlja.

Jedinstvenost web preglednika moguće je provjeriti na službenim stranicama Panopticlick-a:

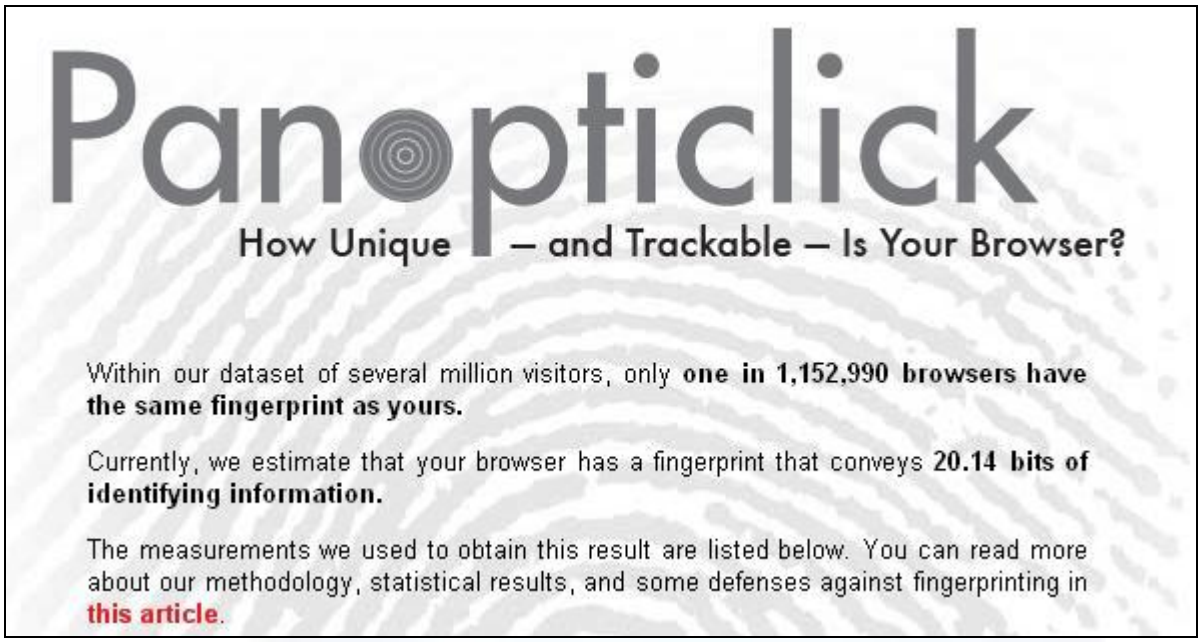
https://panopticlick.eff.org/

Također, na svom portalu EEF educira korisnike te objavljuje rezultate svojih istraživanja. Ovaj portal iznimno je važan za istraživački rad jer služi za prikupljanje otisaka pomoću kojih se obavljaju daljnja istraživanja. Svaki korisnik može provjeriti otisak svog web preglednika i podržati projekt odabirom „Test Me“ opcije.



Slika 3. *Internet stranica projekta Panopticlick*
Izvor: *Panopticlick*

Prilikom uzimanja otiska EEF se obavezuje na anonimno prikupljanje podataka, tj. obećava da te podatke neće koristiti za daljnje praćenje korisnikovih aktivnosti. Prilikom davanja otiska postajete dio statistike koja može biti izložena u javnosti.



Slika 4. *Otisak web preglednika*
Izvor: *Panopticlick*

Panopticlick stranica nudi korisniku uvid u informacije koje se mogu prikupiti iz njegovog web preglednika. Glavni elementi po kojima se radi otisak web preglednika na Panopticlick stranici su User-Agent zaglavlje, Accept polja kog HTTP protokola, vremenska zona, rezolucija ekrana, uključeni ili isključeni kolačići, sustavski fontovi i određeni testovi na superkolačiće.

Kao što se vidi na prethodnoj slici (Slika 4) testirani Internet preglednik jedinstven je u usporedbi s 1.152.990 preglednika u njihovoj bazi. Dakle, oko milijun preglednika je identično testiranom što je relativno mali broj u odnosu na desetke milijuna prikupljenih. Nadalje, za svaki od navedenih polja navedena je njegova vrijednost u bitova informacije koju donosi, te podatak na koliko preglednika se javlja jedan identičan testiranome.

4.1. Algoritam korišten u projektu Panopticlick

Za razliku od dosadašnjih poglavlja u kojima se govorilo o ideji i principu na kojima rade algoritmi za uzimanje otiska u ovom poglavlju govoriti će se o konkretnom algoritmu koji je razvila tvrtka Panopticlick. Taj algoritam se temelji na navedenim idejama, iako ih ne primjenjuje u potpunosti.

Cilj algoritma je pomoću HTTP protokola i Javascript jezika (u kombinaciji s AJAX-om) prikupiti sve česte i manje česte informacije koje web preglednici pružaju Internet stranicama. Prikupljene informacije dijele se na osam osnovnih nizova znakova (eng. *strings*) koji predstavljaju različite podatke prikupljene iz preglednika. Svaki niz predstavlja podatke prikupljene iz jednog od navedenih izvora informacija prikazanih u tablici 1. Otisak web preglednika predstavljaju svi ti nizovi znakova zajedno.

Varijabla	Metoda prikupljanja	Komentar
User Agent	HTTP protokol	Inačica preglednika, inačica operacijskog sustava, jezici, alati i dodatne informacije
HTTP ACCEPT zaglavlja	HTTP protokol	
Omogućeni kolačići ?	HTTP protokol	
Rezolucija ekrana	Javascrip(AJAX)	
Vremenska zona	Javascrip(AJAX)	
Dodaci instalirani u web preglednik	Javascrip(AJAX)	Koristi se knjižnica PluginDetect jezika Javacript, te dodatan kod za detekciju inačice Adobe Acrobat Readera.
Fontovi	Flash applet, Java applet, Javascrip(AJAX)	
Djelomični test superkolačića	Javascrip(AJAX)	Nije u potpunosti implementirano

Tabela 1. *Informacije prikupljene u projektu Panopticlick*
Izvor: *Panopticlick*

Problem u izvođenju algoritma može biti ukoliko je isključen Javascript. No, moguće je prepoznati isključeni Javascripta uz pomoć vrijednosti za dodatke, fontove i superkolačiće (u slučaju isključenog Javascripta njihove vrijednosti su inicijalne). Isto tako moguće je imati onemogućen Flash (pomoću određenih dodataka za njegovu blokadu) što je moguće prepoznati jer Flash postoji u listi dodataka, no nije moguće dohvatiti listu fontova putem Flash-a.

Iako onemogućenim Flash-om nije moguće dohvatiti informacije o fontovima, svejedno je moguće napraviti različit otisak zbog činjenice da postoji mali broj preglednika s takvim svojstvima.

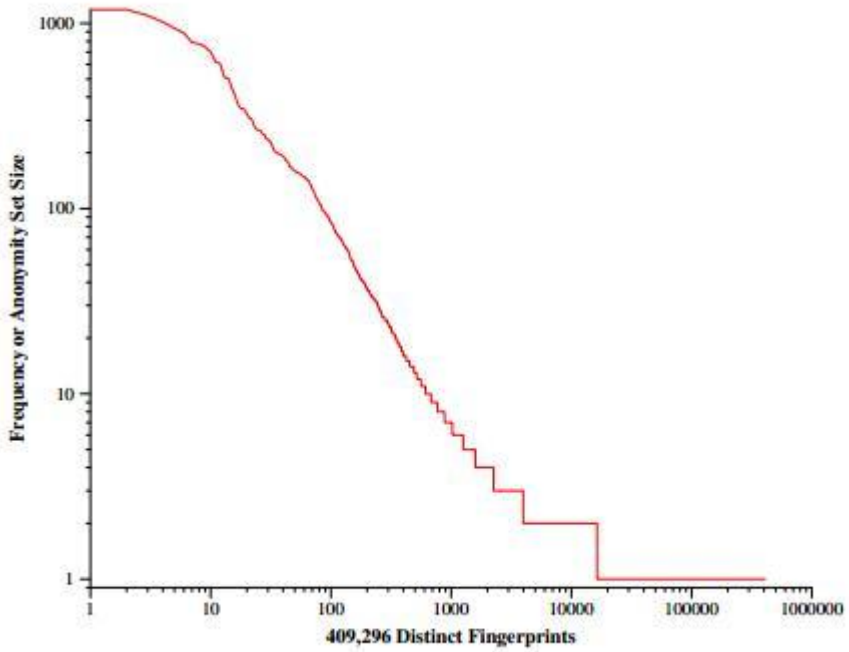
Na sličan način moguće je otkriti i krivotvoreno User-Agent zaglavlje, zbog toga što se njegova vrijednost ne slaže s drugim parametrima.

Za svaku osobu koja je pristupila stranici Panopticlika i odabrala opciju „Test Me“ osim izrade otiska postavlja se tromjesečni kolačić (ukoliko su podržani). Sprema se i sažetak vrijednosti IP adrese s koje je pristigao zahtjev, te sažetak vrijednosti IP adrese bez posljednjeg okteta.

Treba imati na umu da je ovaj algoritam rađen u istraživačke svrhe te izbjegava ulaženje u implementaciju zahtjevnih programa za prikupljanje podataka poput potpunog iskorištavanja mogućnosti Silverlight tehnologije, potpune provjere superkolačića, otkrivanja dodataka u Internet Explorer web pregledniku itd. Zbog toga se pretpostavlja da komercijalni algoritmi za ovu namjenu imaju još preciznije rezultate od onih navedenih ovdje[2].

4.2. Rezultati istraživanja

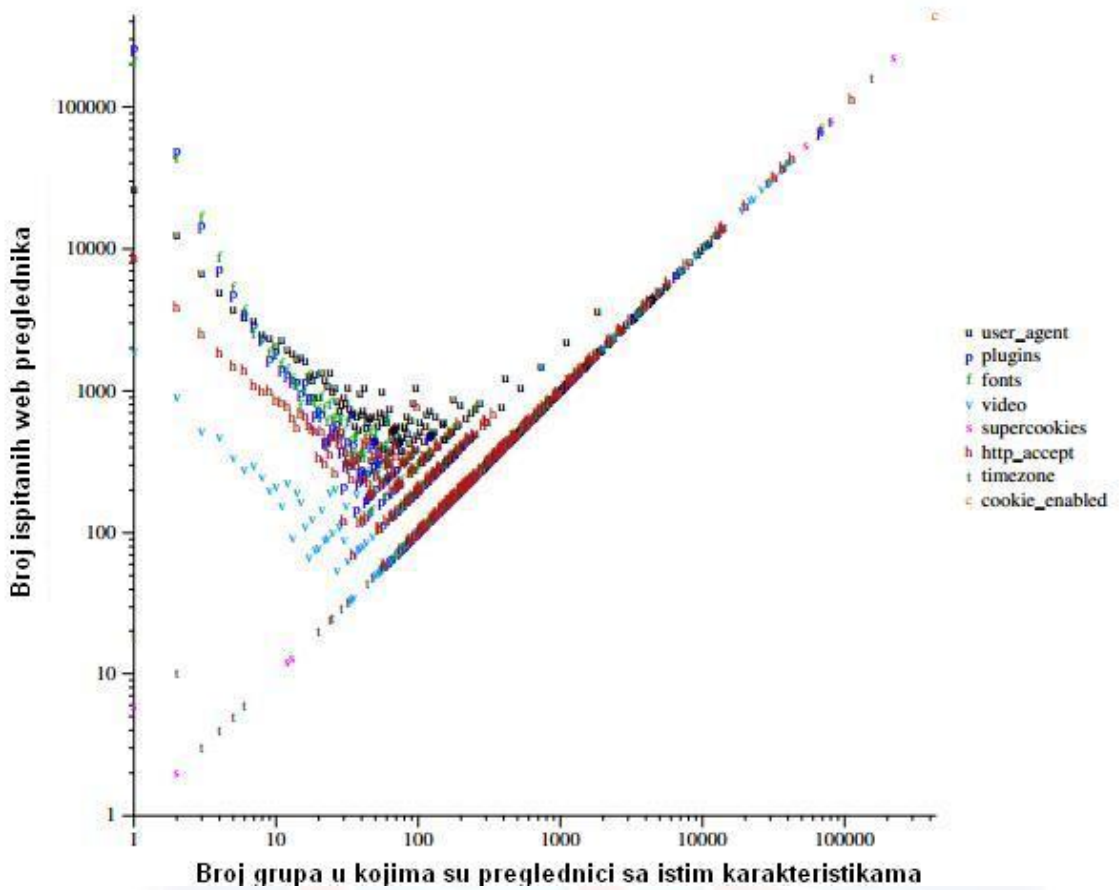
Izneseni podaci prikupljeni su pomoću Panopticlick Internet stranice te su izneseni u dokumentu „How Unique Is Your Web Browser?“ objavljenom na istoj stranici. Podaci su prikupljeni u periodu između 27. siječnja i 15. veljače 2010. godine. Zbog postavljanja kolačića u preglednik korisnika koji su pristupili provjeri moguće je prepoznati i odbaciti otiske koji već postoje. U dodatnoj kombinaciji sa spremljenom IP adresom moguće je prepoznavati i računala s isključenim kolačićima i otkriti više računala iste računalne mreže „sakrivenih“ iza usmerivača (eng. *router*).



Slika 5. *Broj različitih preglednika s obzirom na anonimnost*
Izvor: *Panopticlick*

Na slici 5. grafički je prikazan odnos različitih preglednika s razinom anonimnosti. Na osi apscisa raspoređen je broj različitih preglednika, dok je na ordinati prikazana anonimnost korisnika na način da 1 predstavlja najmanju anonimnost (jedinstveni preglednik). Iz grafa vidimo da od ispitanih preglednika 83.6% pripada jedinstvenim preglednicima (korisnike se može jedinstveno identificirati!), dok skupini koju je teže identificirati pripada svega 8.1% preglednika.

Pokazalo se da oko 90% preglednika namijenjenih osobnim računalima ima jedinstven otisak, dok su najmanje jedinstveni otisak imali oni preglednici kojima je onemogućen Javascript. Preglednici namijenjeni pametnim telefonima poput iPhonea i Androida su zbog svojih ograničenih mogućnosti sličniji i samim time im je teže uzeti otisak. Kod tih preglednika najčešće je ograničena i kontrola kolačića te su zbog toga podložni praćenju i bez metode otiska.



Slika 6. *Broj grupa web preglednika u ovisnosti o svakoj pojedinoj varijabli*
Izvor: *Panopticlick*

Na prethodnoj slici (Slika 6) prikazane su varijable iz kojih se prikupljaju informacije prilikom izrade otiska. Vidi se da su prije instalirani dodaci, User-Agent zaglavlje te korišteni fontovi osnovni podaci po kojima je moguće raditi diferencijaciju preglednika. Nakon toga slijede HTTP Accept zaglavlja i rezolucija ekrana kao dodatni parametri koji rade veću razliku među preglednicima.

4.3. Matematička pozadina

U uvodnom razmatranju uveden je pojam entropije pomoću kojega se može odrediti količina informacije. Pokazano je kako postoji 32.7 bitova entropije kod ljudi na svijetu, no nije pokazano kako se taj broj može mijenjati s obzirom na poznavanje nekih informacija. O tome će biti više riječi u ovom poglavlju.

Kada se otkrije nova informacija, entropija se smanjuje za određeni iznos. Formula kojom se računa za koliko bitova se smanjuje entropija glasi:

$$\Delta S = - \log 2 * p(X=x)$$

, gdje P(X=x) predstavlja vjerojatnost za koju je informacija točna za neku nasumičnu osobu.

Na primjer, ukoliko je poznato da je datum rođenja osobe 1.7., to znači da je to jedan dan od 365 u godini. Stoga je vjerojatnost da nasumično izabrana osoba rođena na taj datum iznosi 1/365=0,002739726. Ako se to uvrsti u navedenu formulu, dobiva se :

$\Delta S = - \log 2 * (1/365) =$ **8.51 bitova**

Dakle, ukoliko je poznat datum rođenja osobe, entropija s 32.7 bitova se smanjuje na 24.19 bita. No, to još nije dovoljno kako bi jednoznačno odredili osobu (iako je entropija smanjena). Za jednoznačno određivanje osobe potrebno je maksimalno smanjiti entropiju, što znači prikupiti dovoljno informacija čije će smanjivanje entropije (ΔS) biti barem 32.7 bita. Ukoliko je, prema prethodnom primjeru, poznato da osoba dolazi iz Šiškovaca smanjenje entropije se računa po formuli:

$$\Delta S = - \log 2 * (841 / 6{,}625{,}000{,}000) = 22.909 \text{ bitova}$$

,gdje je 841 populacija ljudi u Šiškovcima, dok je 6,625,000,000 broj ljudi na planeti.

Vidljivo je da ukoliko znamo osobu rođenu 1.7. koja živi u Šiškovcima imamo 31.41 bita entropije što je vrlo blizu broju 32.7 bita. Dakle, još uvijek nije moguće sa sigurnošću jednoznačno identificirati osobu, što je zapravo i logično jer i u takvo malom selu postoje ljudi rođeni na isti datum. No ukoliko je poznata još neka informaciju (poput prezimena), entropija će proći 32.7 bitova, čime će osoba biti jednoznačno identificirana.

Cijela priča oko otiska web preglednika krije se iza ove matematike. Dakle, potrebno je skupiti dovoljno informacija s kojima će se smanjiti entropija preglednika. Stvari nisu jednostavne kao kod ljudi zbog toga što nije jednostavno odrediti entropiju preglednika, no to se postiže aproksimacijama na temelju statističkih podataka.

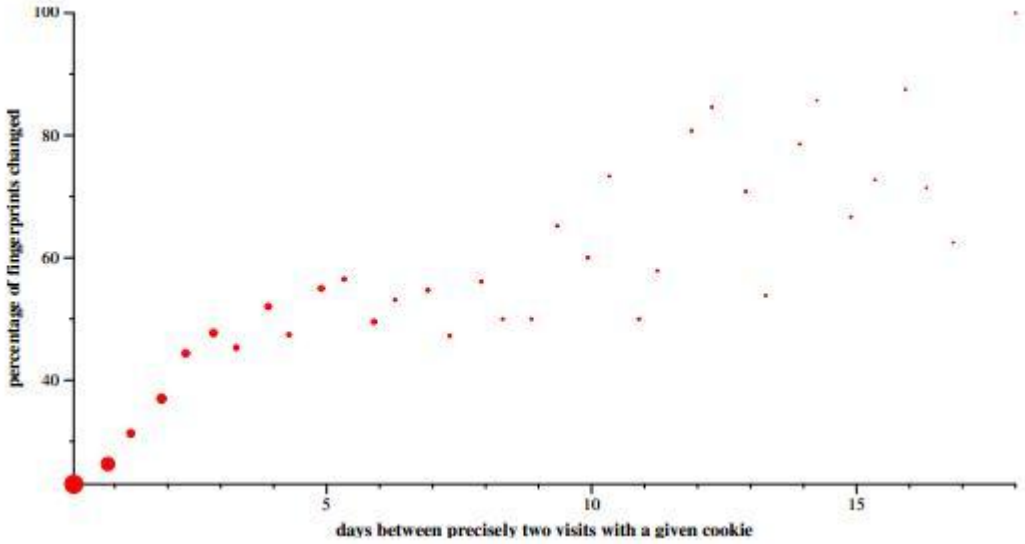
Na primjer, User-Agent sadrži 10.5 bitova informacije u prosjeku što zapravo znači da samo 1500 ljudi dijeli isti User-Agent [1][2].

4.4. Stabilnost otiska web preglednika

Za razliku od ljudskog otiska prsta koji je uvijek isti, otisak web preglednika može se promijeniti. To je zapravo i logično jer se informacije koje preglednik pruža mogu promijeniti. Promjene mogu biti uzrokovane nekom korisničkom aktivnosti, no isto tako računalo samo može promijeniti otisak. Promjene u otisku dogodit će se prilikom svake nadogradnje web preglednika ili njegovih dodataka (neovisno o tome je li inicirano s korisničke strane ili su nadogradnje automatske). Isto tako, otisak se mijenja ukoliko na primjer korisnik isključi ili uključi Javascript ili kolačiće. Ukoliko korisnik, na primjer, promjeni monitor i rezoluciju na njemu, opet će doći do promjene otiska.

Te promjene parametara su zapravo relativno česte, te je za očekivati da se otisak mijenja s vremenom. Najčešće korisnik nije ni svjestan promjene u svom otisku, no isto tako moguće je da napredni korisnici namjerno mijenjaju svoj otisak kako bi prikrili svoje prisustvo.

Projekt Panopticlick ostavlja kolačiće u pregledniku kako bi lakše mogli pratiti promjene u otiscima. Napredni algoritmi za uzimanje otisaka trebali bi moći naknadno prepoznati preglednike s promijenjenim otiskom bez obzira na kolačiće.



Slika 7. *Promjene otiska web preglednika u vremenu*
Izvor: *Panopticlick*

Slika 7 prikazuje kako su se mijenjali otisci prikupljeni na Panopticlick stranici kroz vrijeme. Vidljivo je da se otisci mijenjaju između 40% i 60% kroz pet dana, što zapravo znači da otisak nije stabilan u vremenu. Kroz period od 15 dana otisak se mijenja preko 80% zbog čega ga je vrlo teško pratiti. Takva brza promjena uzrokovana je čestim nadogradnjama koje su gotovo svakodnevno dostupne za preglednike i njihove dodatke.

Ovo istraživanje provedeno je nad preglednicima koji su pristupili Panopticlick stranici u dvama različitim vremenima. Na taj način pokušalo se spriječiti uzimanje u obzir onih korisnika koji su nakon pristupa stranici višestruko ispitivali promjene u svom pregledniku.

Česte promjene otiska najveća su zaštita korisnika od praćenja metodom uzimanja otiska. U sklopu Panopticlick projekta razvijen je jednostavan algoritam koji uz pomoć heurističkih metoda pokušava prepoznati promijenjeni otisak preglednika. Algoritam provjerava osam osnovnih parametara kod uzimanja otiska te ih uspoređuje s već prikupljenima. Na temelju kolačića moguće je zaključiti je li algoritam ispravno odredio promjenu ili je pogriješio. Zbog svoje jednostavnosti algoritam ne može odrediti promjenu ukoliko je u novom otisku isključen Javascript ili Flash (što korisnici često rade nakon posjeta Panopticlick stranici, no to nije česta pojava u Internet svijetu).

Nakon provedenog istraživanja nad korisnicima koji su imali promijenjen otisak, algoritam je u 65% slučajeva točno prepoznao promijenjeni otisak. Pogriješio je u svega 0.56%, dok u 35% slučajeva nije uspio odrediti rezultat (zbog isključenog Javascripta ili Flasha). Unatoč vrlo jednostavnom algoritmu, postotak pogodaka je vrlo visok. Ako se uzme u obzir da je algoritam moguće poboljšati, može se zaključiti kako je promijenjeni otisak moguće prepoznati u vrlo velikom broju slučajeva. Najteže je prepoznati otiske s isključenim Javascriptom zbog ograničenih podataka koje pruža i to je jedan od osnovnih načina zaštite od uzimanja otiska[2].

5. Prednosti i nedostaci uzimanja otiska web preglednika

Prednosti tehnike uzimanja otiska su višestruke. Najveću iskoristivost ova tehnika ima na području računalne sigurnosti gdje se može upotrebljavati na ostvarivanju autentikacije ili otkrivanju sigurnosnih prijetnji.

Prilikom ostvarivanja autentikacije otisak web preglednika može služiti kao dodatan mehanizam provjere korisnika ili kao drugi sloj u dvofaktorskoj (eng. *two factor*) autentikaciji. Otisak web preglednika u takvim slučajevima služi kao način da se otkriju prijave korisnika s drugačijeg preglednika od onoga koji je do tada korišten. U takvim slučajevima moguće je otkriti nepravilnosti i preventivno djelovati na potencijalni napad. Ukoliko se korisnik prijavljuje uvijek s istog preglednika, svaka prijava s drugačijega može biti naznaka pokušaja napada.

Tehnika otiska web preglednika pokazala se vrlo korisnom u kolovozu 2009. kada su srpski kriminalci pokušavali s ukradenih kreditnih kartica svaki dan prebaciti iznos od 1.99$ na iReel.com Internet stranicu. Unatoč pokušajima prikrivanja koristeći lažne IP adrese (pomoću web proxya), te onemogućavanja kolačića, prevarante su pomoću otiska web preglednika uspjeli prepoznati i onemogućiti im daljnje transakcije.
Ova tehnika također se može koristiti na stranicama na kojima je važno da samo jedan korisnik ima pristup jednom korisničkom računu. Primjer toga mogu biti Internet forumi ili stranice koje se naplaćuju, jer njima nije u interesu da se s jednog stvorenog računa prijavljuje više ljudi. Otisak web preglednika je vrlo dobra tehnika za otkrivanje takvih korisnika. Prilikom registracije korisnika uzme mu se otisak web preglednika, te se nakon toga otisci uzimaju i prilikom svake prijave. Vrlo lako je utvrditi s kojih preglednika se korisnik najčešće prijavljuje i kada prijavu radi preglednik koji nije do tada korišten. U takvim slučajevima to može biti indikacija da jedan korisnički račun koristi više ljudi [4].
Iako se pokazalo kako je tehnika uzimanja web preglednika korisna i moćna, uglavnom se smatra negativnom. S obzirom na to da ne ostavlja nikakve tragove na korisničkom računalu (za razliku od kolačića) gotovo nemoguće je otkriti stranice koje ju koriste. To ju čini savršenim alatom za zadiranje u privatnost korisnika. Smatra se da tehniku koriste razne banke za identifikaciju svojih korisnika, društvene mreže, razne komercijalne stranice itd. Kada se preglednik poveže s korisničkim računom poput facebooka ili Gmaila, moguće je identificirati korisnika koji ga koristi. Ne pomaže ni činjenica da nerijetko iz socijalnih mreža „cure“ informacije o korisnicima. Nakon toga je moguće pratiti što ti korisnici rade na Internetu, što ih zanima i kako se ponašaju. Sama činjenica da je takvo što moguće izaziva mnogo kontroverza što pokazuje broj ljudi koji pokušavaju promijeniti svoj otisak nakon posjeta Pantopticlikovoj stranici. Praćenje korisnika moguće je iskoristiti za prikupljanje informacija o korisniku kako bi mu se kasnije mogli prikazivati sadržaji ili proizvodi koji ga zanimaju. Isto tako ova metoda može biti korištena u prikupljanju informacija o korisnicima od strane Vlade ili sličnih organizacija, što je posebno popularno u zemljama s ograničenom slobodom govora. Za Internet se govori da je jedini slobodni medij, a ova tehnika mogla bi zadati ozbiljan udarac toj tezi.

Pitanja koje si svaki korisnik postavlja su: Koliko je ta tehnologija razvijena? Koristi li se u stvarnosti? Koje stranice je koriste? Prate li me sada?

Na sve to vrlo teško je dati konkretne informacije baš zbog činjenice da vrlo često iza takvih aktivnosti stoje velike (često vojne) organizacije ili Vlade. Pouzdano se može reći jedino da takva tehnologija postoji i da se koristi i više nego što bi se na prvi pogled moglo pretpostaviti [5].

6. Kako se zaštititi od uzimanja otiska web preglednika

Iz prethodnog poglavlja vidi se kako otisak web preglednika ne mora uvijek biti pozitivna stvar. Korisnici žele postići što veću anonimnost na Internetu, te im otisak web preglednika predstavlja izrazito veliku prijetnju. Zbog toga se pokušavaju naći razni načini kako anonimizirati web preglednik na način da se utopi u moru drugih sličnih preglednika. U praksi se pokazalo da to baš i nije jednostavan zadatak, već su stvari puno složenije od brisanja kolačića iz web preglednika.

Ponekad su alati koji se pokušavaju iskoristiti kontraproduktivni, te samo povećavaju jedinstvenost web preglednika. Primjer toga su pokušaji promjene User-Agent zaglavlja, te blokiranje Flasha koji ne samo da ne pomažu, već čine otisak takvog preglednika karakterističnim.

Agent je polje koje sa sobom donosi 5-15 bitova (10.5 u prosjeku) informacija. Što znači da samo 1500 ljudi dijeli isti User-Agent, što je jedna trećina od potrebnih informacija za identifikaciju preglednika [3].

Browser class	Avg. identifying information	Minimum identifying information	(Least identifying user agent)
Modern Windows Desktops	10.3–11.3 bits	4.6 – 5.0 bits	Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.9.1.3) Gecko/20090824 Firefox/3.5.3 (.NET CLR 3.5.30729)
Internet Explorer	13.2–13.5 bits	6.3 – 7.2 bits	Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1)
Firefox	8.6 – 9.4 bits	4.6 – 5.0 bits	Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.9.1.3) Gecko/20090824 Firefox/3.5.3 (.NET CLR 3.5.30729)
Chrome	7.5–8.5 bits	5.7 – 6.2 bits	Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US) AppleWebKit/532.0 (KHTML, like Gecko) Chrome/3.0.195.27 Safari/532.0
Linux	11.8–13.15 bits	6.6–7.9 bits	Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.0.14) Gecko/2009090216 Ubuntu/9.04 (jaunty) Firefox/3.0.14
Ubuntu	9.6 – 11.7 bits	6.6 – 7.8 bits	Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.0.14) Gecko/2009090216 Ubuntu/9.04 (jaunty) Firefox/3.0.14
Debian	13.5–15.3 bits	10.50 – 11.7 bits	Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.9.0.14) Gecko/2009091010 Iceweasel/3.0.6 (Debian-3.0.6-3)
Macintosh	8.8–9.3 bits	5.8–5.8 bits	Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10.5; en-US; rv:1.9.1.3) Gecko/20090824 Firefox/3.5.3
iPhone	10.8 – 11.3 bits	8.7 – 9.3 bits	Mozilla/5.0 (iPhone; U; CPU iPhone OS 3_1 like Mac OS X; en-us) AppleWebKit/528.18 (KHTML, like Gecko) Version/4.0 Mobile/7C144 Safari/528.16
Blackberry	14.7 – 15.5 bits	12.0 – 12.7 bits	BlackBerry9530/4.7.0.148 Profile/MIDP-2.0 Configuration/CLDC-1.1 VendorID/105

Tabela 2. Informacije koje donose popularni preglednici
Izvor: EEF

U tabeli 2 prikazano je koliko bitova informacija sa sobom donosi User-Agent pojedinog preglednika. Zanimljivo je primijetiti da sustavi koji koriste neku od inačica linux/unix operacijskog sustava, unatoč tome što imaju manji broj korisnika, i dalje u prosjeku odaju sličan broj informacija kao i ona računala s Windows operacijskim sustavom (koji imaju puno više korisnika). Isto tako metode poput korištenja proxya, promjene IP adresa ili brisanje kolačića ne pomažu kod uzimanja otiska preglednika [3].

Postoje neke preporuke koje je izdao Panopticlick s ciljem povećanja anonimnosti na Internetu. Za početak je potrebno imati što Internet preglednik koji se često koristi. To može biti na primjer posljednja inačica Firefoxa instalirana na novoj inačici Windows operacijskog sustava (npr. Windows 7). No to samo po sebi nije dovoljno, jer je na takav preglednik i dalje moguće instalirati određene dodatke. Za razliku od preglednika namijenjenih stolnim računalima, preglednici za mobilne telefone su puno otporniji na otiske web preglednika. Razlog tome je manja varijabilnost kod takvih

preglednika. Oni nemaju mogućnost instaliranja dodataka, uvijek imaju iste rezolucije, te svi imaju iste instalirane iste fontove.

Onemogućavanje Javascripta je efektivna metoda u sprječavanju uzimanja otisaka iz razloga što se time onemogućuje prikupljanje velikog broja podataka nužnih za uzimanje otiska. Na žalost, mnoge stranice neće ispravno raditi s onemogućenim Javascriptom. Postoje dodaci poput „NoScript“, koji omogućuju korisniku da sam odluči hoće li ili neće dozvoliti Javascript na određenoj stranici. No i ta metoda nije u potpunosti sigurna jer korisnik mora imati dobru intuiciju prilikom omogućavanja Javascripta (mora jako dobro poznavati vlasnika stranice kojoj pristupa, što je u današnje vrijeme gotovo nemoguće).

TorButton[3] je način za efektivno povećanje anonimnosti na Internetu. Pomoću njega se kombinira blokiranje Javascripta i promjena svojstva preglednika. Jedini nedostatak je značajno usporavanje rada preglednika koji koristi TorButton.

Osim samih korisnika, ulogu u zaštiti od uzimanja otiska imaju i proizvođači web preglednika. Microsoftov Internet Explorer, koji je često je bio na meti kritičara zbog svojih sigurnosnih propusta, u ovom slučaju se pokazao boljim od ostalih preglednika. Za razliku od ostalih preglednika, Internet Explorer nema mogućnosti enumeracije dodataka što otežava prikupljanje informacija o njima. Čak i ako se uspiju prikupiti informacije, entropija koju donose je manja nego kod ostalih preglednika.

Proizvođači dodataka za web preglednike također vrlo često sami pogoduju u stvaranju jedinstvenog otiska. Razlog tome su njihova nastojanja da sa što većim brojem inačica i podinačica sebi olakšaju ispravljanje pogrešaka u svojim proizvodima. Primjer toga je zapis inačice dodatka DivX WebPlayer 1.4.0.233 koja otkriva puno više informacija od zapisa DivX Web Player 1.4. U ovakvim slučajevima postoje dva ekstrema. Jedan je onaj koji pogoduje proizvođačima dodataka, a to je zapis sa što većom razinom detalja. Na taj način je lakše izolirati u kojoj inačici se javlja pogreška i prije je ispraviti. No, takav način izrazito pogoduje stvaranju jedinstvenog otiska. Drugi ekstrem je suprotnost tome gdje dodaci pružaju što manje informacija, no to može stvoriti probleme njihovim proizvođačima. Trenutno u svijetu web preglednika prevladava prvi ekstrem, što bi vjerojatno trebalo promijeniti ako se želi povećati razina anonimnosti na Internetu.

Postoje još neke informacije koje se mogu iskoristiti za povećavanje entropije kod uzimanja otiska. Na primjer, kod dohvaćanja liste fontova oni se vraćaju uvijek u istom (ne posloženom) poretku. Manje promjene su primijećene samo kod Mac OS X operacijskog sustava. Na taj način se dodatno povećava količina informacija koju je moguće prikupiti iz fontova instaliranih na računalo.

[3] TorButton je komponenta Tor Browser Bundle alata koji brine za sigurnost korisničkih aplikacija i privatnost korisnika u pregledniku Mozilla Firefox. Zbog zaštite privatnosti TorButton onemogućuje mnoge aktivne komponente preglednika.

Otisak web preglednika danas i u budućnosti

Kao što je već navedeno, tehnologija otiska web preglednika nije samo ideja nego se već danas provodi u praksi. Zbog svojih svojstava očekuje se porast korištenja te tehnologije i njeno usavršavanje.

U budućnosti će se kombiniranjem otiska s drugim tehnikama napraviti profinjeni algoritmi koji će moći prepoznati i pratiti korisnike koji nemaju jedinstveni otisak.

Jedan od mogućih napada na privatnost korisnika je obnavljanje kolačića ili superkolačića putem otiska web preglednika. Ukoliko korisnik želi biti anoniman, te obriše kolačiće i superkolačiće, i dalje ih je moguće rekonstruirati iz otiska web preglednika u kombinaciji sa zabilježenom IP adresom. Za regeneraciju kolačića potrebno je svega 20 bita informacije ukoliko korisnik nastavi koristiti istu IP adresu ili željenoj web stranici pristupa iz iste lokalne mreže (npr. koristi istog pružatelja usluga pristupa Internetu). Moguće je prepoznati preglednike i povezati ih s prethodnim prijavama čak i ukoliko isključe kolačiće.

Pretpostavka je da će se preglednici i njihovi dodatci razvijati te će se broj različitosti još više povećati u odnosu na trenutno stanje. To će omogućiti bolju primjenu tehnika za uzimanje otisaka web preglednika te se pretpostavlja kako vrhunac ove tehnologije tek dolazi. Ukoliko se želi smanjiti ugrožavanje privatnosti koje ova tehnologija definitivno donosi sa sobom potrebna je suradnja između „običnih“ korisnika Interneta i proizvođača web preglednika.

Očekuje se ubrzan razvoj preglednika namijenjenih mobilnim telefonima te je pitanje vremena kada ni oni više neće biti otporni na otiske kao što su sada. Pretpostavka je da će se u budućnosti mobilnim web preglednicima moći uzeti otisak isto kao i onima namijenjenima stolnim računalima.

Vjeruje se da će se zbog pritiska javnosti proizvođači web preglednika i njihovih dodataka odlučiti na određene korake u sprječavanju praćenja korisnika. Već danas svi popularni web preglednici imaju opciju privatnog korištenja, u kojem su onemogućeni kolačići i slanje privatnih informacija. U takvu mogućnost potrebno je nadograditi opciju za isključivanje prikaza punih inačica dodataka web pregledniku. Isto tako moguće je postavljati User-Agent polja na neke predefinirane vrijednosti koje će biti iste kod svih korisnika kada se ova opcija (privatno korištenje) koristi. No, unatoč tome, potreban je i veći broj korisnika koji će koristiti opisane mogućnosti jer će u suprotnom one biti kontraproduktivne.

7. Zaključak

Iako sama ideja otiska web preglednika ne zvuči toliko moćno, istraživanja su pokazala da je to snažno oružje u rukama velikih organizacija. Zbog svojih osobina robusnija je od bilo kojih drugih metoda prepoznavanja korisnika, te je gotovo nemoguće oduprijeti se bez smanjivanja kvalitete korištenja Internet usluga. Opravdano je glavna bojazan da će ova tehnologija omogućiti praćenje korisnika, te samim time narušavati njihova osnovna prava. Zbog nemogućnosti otkrivanja napada, te teškom sprječavanju korisnici su izloženi opasnostima bez pravog upozorenja. Još uvijek ne postoje efektivne tehničke protumjere kojima će se svakodnevni korisnici moći efektivno braniti od kršenja njihove privatnosti. Isključivanje Javascripta je učinkovita metoda, no nije realno očekivati da da će korisnici pristupati Internet sjedištima s isključenim Javascriptom upravo zbog drastičnog smanjenja komfora i mogućnosti rada s pojedinim sustavima (npr. bankarskim). TorButton i korištenje privatnih opcija mogu pomoći zaštiti od uzimanja otiska, no zahtijeva nešto više informatičkog znanja za korištenje. Zbog toga su glavna žrtva ove metode obični korisnici ne-tehničkog zanimanja koji nisu ni svjesni opasnosti. Jedino rješenje za nastalu situaciju je edukacija korisnika, no u medijima se rijetko spominju ovakve opasnosti, a još rjeđe se nude konkretna rješenja. Isto tako očekuje se od država da reguliraju zakone koji se odnose na privatnost korisnika te da zakonski ograniče primjenu tehnologije uzimanja otiska preglednika. Dok se ne ispune ti uvjeti, te dok se ne nađe neko odgovarajuće tehničko rješenje, korisnicima preostaje biti na oprezu prilikom posjeta stranicama za koje se pretpostavlja da uzimaju otisak. Svjesnost problema je korak ka rješenju, te zbog toga treba surađivati s drugim korisnicima i istraživačima koji pokušavaju istražiti/riješiti problem. Jedan od načina na koji svi mogu sudjelovati je testiranjem na stranici Panopticlick projekta koja će korisnika obavijestiti o vlastitim propustima. S druge strane, korisnik će pomoći (slanjem informacija o svom pregledniku) u detaljnijem istraživanju ove tehnologije.

8. Leksikon pojmova

Kolačić

Datoteka koja sadrži podatke o posjeti web stranici. Na taj način vlasnici web stranice rade statistiku posjeta. Cookie također pamti neke postavke koje ste namjestili i podatke koje ste upisali na posjećenoj stranici (npr. lozinku). cookie datoteka.

Reference: http://www.httpwatch.com/httpgallery/cookies/

Ostale poveznice:

http://webdesign.about.com/cs/cookies/a/aa082498a.htm
http://www.nczonline.net/blog/2009/05/05/http-cookies-explained/

IP protokol

IP je jedan od glavnih protokola u Internetu, a koristi se za usmjeravanja paketa kroz Internet. U tu namjenu, dodjeljuje IP adrese izvora paketa i njegovog odredišta na temelju kojih će se paketi usmjeravati kroz nekoliko računalnih mreža.

Reference: http://compnetworking.about.com/od/networkprotocolsip/g/ip_protocol.htm

Ostale poveznice:

http://en.wikipedia.org/wiki/Internet_Protocol http://www.ietf.org/rfc/rfc791.txt

HTTP protokol

HyperText Transfer Protocol - Osnovna i najčešća metoda prijenosa informacija na Webu. Predstavlja protokol na aplikacijskom sloju OSI modela, a osnovna namjena je prijenos HTML dokumenata (tj. web stranica). HTTP je request/response protokol za komunikaciju između poslužitelja i klijenta. HTTP klijent, kao što je web preglednik najčešće inicira prijenos podataka nakon što uspostavi TCP vezu s udaljenim web poslužiteljem na određenom priključku. Poslužitelj konstantno osluškuje zahtjeve na određenom mrežnom komunikacijskom priključku (tipično priključak 80), čekajući da klijent inicira komunikaciju. - Osnovna i najčešća metoda prijenosa informacija na Webu. Predstavlja protokol na aplikacijskom sloju OSI modela, a osnovna namjena je prijenos HTML dokumenata (tj. web stranica). HTTP je request/response protokol za komunikaciju između poslužitelja i klijenta. HTTP klijent, kao što je web preglednik najčešće inicira prijenos podataka nakon što uspostavi TCP vezu s udaljenim web poslužiteljem na određenom priključku. Poslužitelj konstantno osluškuje zahtjeve na određenom mrežnom komunikacijskom priključku (tipično priključak 80), čekajući da klijent inicira komunikaciju.

Reference: http://hr.wikipedia.org/wiki/HTTP http://www.w3.org/Protocols/

Ostale poveznice:

http://www.ietf.org/rfc/rfc2616.txt
http://compnetworking.about.com/od/networkprotocols/g/bldef_http.htm

Ajax

Asynchronous JavaScript and XML - Tehnologija weba koja omogućuje transparentnu komunikaciju između klijenta i poslužitelja bez osvježavanja trenutne stranice. Kao i druge slične tehnologije (DHTML, LAMP), Ajax nije jedna tehnologija već skupina tehnologija. Koristi kombinaciju HTML i CSS jezika kako bi prenijela informaciju. DOM (eng. *Document Object Model*) modelu se pristupa putem JavaScript isječaka kako bi se dinamički izmijenio sadržaj web stranice.

Reference: http://www.w3schools.com/ajax/default.asp

Ostale poveznice:

http://searchwindevelopment.techtarget.com/definition/Ajax
http://webtrends.about.com/od/web20/a/what-is-ajax.htm http://www.wisegeek.com/what-is-ajax.htm

JavaScript

Programski jezik JavaScript - JavaScript je skriptni programski jezik, koji se izvodi u web pregledniku na strani korisnika. Napravljen je da bude sličan Javi, zbog lakšega korištenja, ali nije objektno orijentiran kao Java, već se temelji na prototipu i tu prestaje svaka povezanost s programskim jezikom Java. Izvorno ga je razvila tvrtka Netscape (www.netscape.com). JavaScript je izrađen primjenom ECMAScript standarda. - JavaScript je skriptni programski jezik, koji se izvodi u web pregledniku na strani korisnika. Napravljen je da bude sličan Javi, zbog lakšega korištenja, ali nije objektno orijentiran kao Java, već se temelji na prototipu i tu prestaje svaka povezanost s programskim jezikom Java. Izvorno ga je razvila tvrtka Netscape (www.netscape.com). JavaScript je izrađen primjenom standarda ECMAScript.

Reference: http://javascript.about.com/od/reference/p/javascript.htm

Ostale poveznice: http://www.w3schools.com/js/default.asp

Autentikacija

Autentikacija je proces potvrđivanja identiteta podatka ili osobe. - Autentikacija je proces određivanja identiteta nekog subjekta, najčešće se odnosi na fizičku osobu. U praksi subjekt daje određene podatke po kojima druga strana može utvrditi da je subjekt upravo taj kojim se predstavlja. Najčešći primjeri su: uz korištenje kartice na bankomatu i upisivanje PIN-a, ili upisivanje (korisničkog) imena i zaporke.

Reference: http://searchsecurity.techtarget.com/definition/authentication

Ostale poveznice: http://en.wikipedia.org/wiki/Authentication

XML

XML je kratica za EXtensible Markup Language, odnosno jezik za označavanje podataka. Ideja je bila stvoriti jedan jezik koji će biti jednostavno čitljiv i ljudima i računalnim programima. U XML-u se sadržaj uokviruje odgovarajućim oznakama koje ga opisuju i imaju poznato, ili lako shvatljivo značenje.

Reference: http://webdesign.about.com/od/xml/a/aa091500a.htm

Ostale poveznice: http://www.w3schools.com/xml/default.asp http://www.w3.org/XML/

9. Reference

[1] A Primer on Information Theory and Privacy, https://www.eff.org/deeplinks/2010/01/primer-information-theory-and-privacy, srpanj 2012.

[2] How Unique Is Your Web Browser?, https://panopticlick.eff.org/browser-uniqueness.pdf , srpanj 2012.

[3] Panopticlick, https://panopticlick.eff.org/, srpanj 2012.

[4] Browser Fingerprints: A Big Privacy Threat, http://www.pcworld.com/article/192648/browser_fingerprints_a_big_privacy_threat.html, srpanj 2012.

[5] EFF: Forget cookies, your browser has fingerprints, http://www.computerworld.com/s/article/9176904/EFF_Forget_cookies_your_browser_has_fingerprints, srpanj 2012.

[6] BrowserSpy, http://browserspy.dk/, srpanj 2012.

[7] Hypertext Transfer Protocol, http://en.wikipedia.org/wiki/Hypertext_Transfer_Protocol, srpanj 2012.

[8] HTTP Cookie, http://en.wikipedia.org/wiki/HTTP_cookie, srpanj 2012.