DPI in legalni nadzor

Goran Grašič, Roman Kotnik, Franci Katrašnik, Andrej Kos

University of Ljubljana, Faculty of Electrical Engineering, Tržaška 25, 1000 Ljubljana, Slovenia
goran.grasic@ltfe.org

Deep packet inspection and lawful interception

Deep Packet Inspection or DPI is a computer networking term that refers to devices and technologies that inspect and take action based on the contents of the packet (commonly called the "payload") rather than just the packet header[4]. This article describes DPI technology in lawful interception, what it does, how it works, methods of analysis that it uses and ways to access data in the network.

1 Uvod

Vpogled v vsebino elektronskih komunikacij (angl. deep packet inspection ali DPI) se nanaša na naprave in tehnologije, ki analizirajo promet glede na vsebino paketa. To vsebino (angl. payload) pridobijo same. Ostale tehnologije gledajo samo glavo paketa in v tem je bistvena prednost DPI. [1]

DPI uporabljamo tudi za legalni nadzor. To je legalno urejen uradni dostop do privatnih komunikacij, kot so telefonski klici ali e-poštna sporočila. V osnovi je legalni nadzor varnostni proces, kjer mrežni operater ali ponudnik storitev omogoča organom kazenskega pregona dostop do komunikacij posameznikov ali organizacij.

2 Deep packet inspection

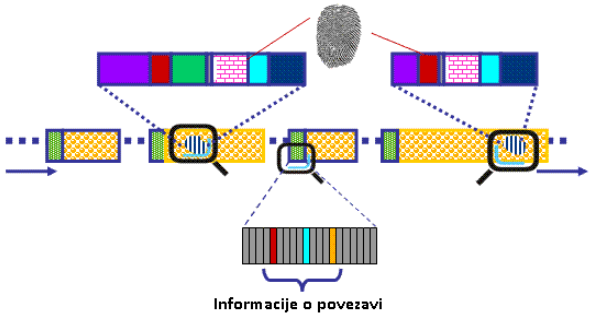
DPI je pomembna inovacija med omrežnimi tehnologijami, saj predstavlja temelj veliki količini sedanjih in prihodnjih storitev. Produkti, ki podpirajo DPI postajajo globalna prioriteta na področju informacijske tehnologije. V podjetjih to vključuje omrežne naprave za usmerjanje in komutacijo, požarne zidove naslednje generacije, sisteme za detekcijo in preprečevanje vdorov (angl. Intrusion Detection Prevention ali IDP), preprečitev uhajanja informacij, nadzor prometa in ostalo.

3 Metodologije analize prometa

Da bi naprave lahko obvladovale številne internetne, mrežne aplikacije in protokole, mora biti uveden nek metodičen in sistematičen način identifikacije.

3.1 Elektronski podpis

V najširšem smislu so podpisi vzorci, ki so izbrani za unikatno identifikacijo aplikacije ali protokola. Ko se srečamo z novo aplikacijo ali protokolom, ga analiziramo in mu dodamo primeren podpis, kar se zapiše v bazo (tipično se tej bazi reče knjižnica s podpisi, ali po angleško signature library).



Slika 1: Opazovanje vsebine paketa s pomočjo DPI.

Podpise aplikacij moramo preverjati redno, saj se spreminjajo pri posodobitvah. BitTorrent, eMule in Skype posodabljajo programsko opremo pogosto in vzpodbujajo (v nekaterih primerih celo prisilijo) uporabnike, da prenesejo novo različico. Uporaba novih različic z neposodobljenimi podpisi dramatično vpliva na hitrost kvalifikacije.

3.2 Napake pri podpisih

Čeprav je podpis razvit z namenom unikatne in kompletne identifikacije aplikacije ali protokola, pride do primerov, pri katerih podpis ni robusten (angl. weak signature). V takem primeru se pojavijo problemi s klasifikacijo. »False positive« je termin, ki se navezuje na napačno klasifikacijo, oziroma je verjetnost, da se bo aplikacija predstavljala kot nekaj kar ni. Ko razvijamo podpise moramo zagotoviti, da ne bo prihajalo do teh napak.
»False negative« se nanaša na tiste primere, kjer ni mogoče identificirati aplikacije – v nekaterih primerih je identifikacija zaupna, v nekaterih pa jo klasifikacijsko orodje preprosto spregleda. Najpogostejši razlog za ta fenomen je, da se aplikacije izvajajo na podoben način vendar na različnih sistemih.

3.3 Metode analize podpisov

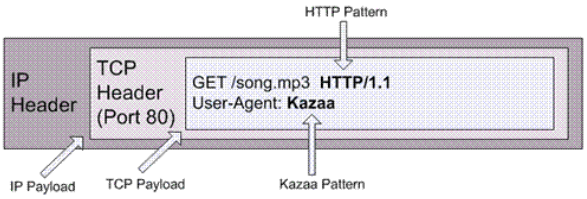
Poznamo več metod za identifikacijo in klasifikacijo prometa. Lahko analiziramo glede na vrata (angl. port), glede na ujemanje določene besedne zveze, števil ali obnašanja in hevristike.

3.3.1 Analiza glede na vrata (port)

Ta način je verjetno najlažji in najbolj poznan od vseh. Razlog za to je preprosto dejstvo, da aplikacije večinoma uporabljajo privzeta vrata. Oglejmo si to na primeru POP3, ki se ga uporablja pri e-poštnih aplikacijah. Dohodni POP3 tipično uporablja vrata 110 (v varnem načinu 995), odhodni (SMTP) pa vrata 25.
Zelo lahko je zaznati aplikacijo glede na vrata preko katerih deluje, zato je to tudi slabost. Veliko aplikacij se tako pretvarja, kot da so neka druga aplikacija. Tipično je to pri vratih 80, saj se številne aplikacije predstavljajo, kot da so HTTP promet, ki deluje na teh vratih.
Kot omenjeno zgoraj, nekatere aplikacije uporabljajo naključna vrata namesto privzetih. V tem primeru je pogosto vključen nek vzorec, saj so ponavadi prva vrata naključna, naslednja pa ne več. Zaradi teh razlogov analiza glede na vrata ni uporabno orodje za identifikacijo aplikacij, ampak je bolj pomagalo pri drugih orodjih.

3.3.2 Analiza glede na ujemanje z besedno zvezo

Analiza glede na ujemanje z besedno zvezo deluje tako, da po vsebini paketa išče določeno besedo ali številko. Ta beseda je lahko tudi razporejena čez cel paket ali pa celo čez več paketov.
Veliko aplikacij sporoči njihovo ime znotraj samega protokola, kot je to pri Kazaa, kjer se beseda "Kazaa" pojavlja pri polju user-agent s tipično HTTP GET zahtevo. Ravno v tem primeru lahko vidimo kako pomemben je DPI za pravilno klasifikacijo. Če bi v zgornjem primeru uporabili analizo glede na vrata, bi ugotovili, da gre za neko HTTP zahtevo, saj gre promet preko vrat 80. Ta ugotovitev bi bila seveda napačna.

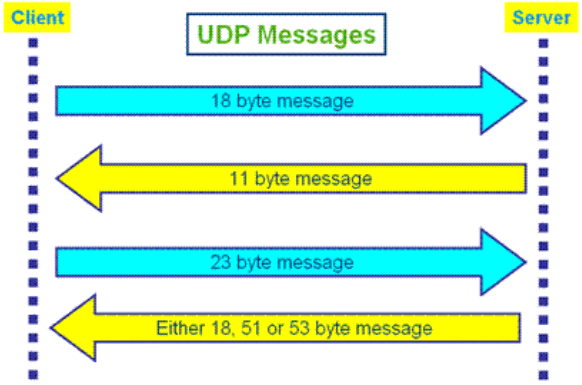


Slika 2: Analiza glede na ujemanje z besedno zvezo (Kazaa) [4]

Ta primer spet poudarja, da potrebujemo več orodij za analizo, da zagotovimo pravilno klasifikacijo.

3.3.3 Analiza glede na numerične lastnosti

Analiza glede na numerične lastnosti zajema preiskavo aritmetičnih in numeričnih karakteristik znotraj enega ali več paketov. Nekatere izmed teh preiskav zajemajo analizo dolžine paketa, število paketov poslanih pri odgovoru na določeno transakcijo in numerični zamik pri nekaterih besednih zvezah oziroma bitih znotraj paketa.

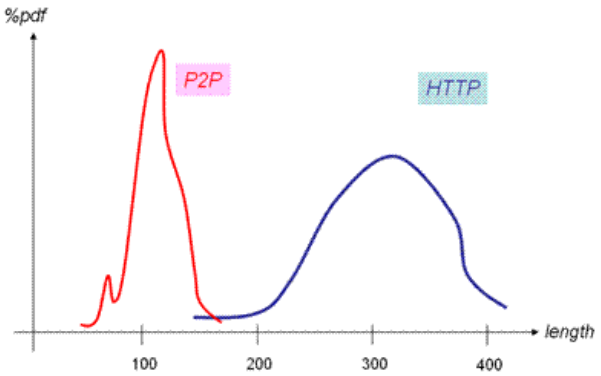


Slika 3: Analiza glede na numerične lastnosti (Skype) [4][4]

Kot primer si poglejmo proces vzpostavitve TCP povezave z uporabo protokola UDP pri aplikaciji Skype (verzije do 2.0). Odjemalec pošlje 18 oktetov dolgo sporočilo in pričakuje 11-okteten odgovor. Temu sledi pošiljanje 23-oktetnega sporočila in pričakovanje 18, 51 ali 53-oktetnega odgovora. Podobno kot pri analizi glede na vrata in glede na besedno zvezo, analiza glede na numerične lastnosti ni samozadostna in lahko vodi v napačne analize.

3.3.4 Analiza glede na obnašanje in hevristiko

Analiza glede na obnašanje se nanaša na načine delovanja posameznega protokola, hevristična analiza pa na analizo statističnih parametrov paketnega prenosa. Pogosto so združene v eno, da omogočijo še boljše rezultate.
Dogodki, ki jim sledijo posledice kažejo na nek vedenjski vzorec, ki mu lahko sledimo. Lep primer tega je, ko se UDP povezava pretvori v TCP (pri uporabi istega IP-ja in vrat).



Slika 4: P2P proti HTTP [4]

Še en primer behavioristične in hevristične analize lahko vidimo na sliki 10, ki kaže primerjavo med tipično HTTP in P2P aplikacijo. Iz grafa je razvidno, da imajo HTTP paketi velikosti okoli 300 oktetov, medtem ko P2P raje uporablja krajše pakete. Na ta način lahko ugotovimo, ali se preko vrat 80 prenaša HTTP ali nek drug P2P promet.

4 DPI za legalni nadzor v omrežjih IP

Legalni nadzor v omrežjih je zajemanje in posredovanje podatkov oblastem z namenom analize ali pridobitve dokazov. Ti podatki načeloma vključujejo signalizacijske informacije in informacije o upravljanju omrežja, v nekaterih primerih celo vsebino komunikacije.

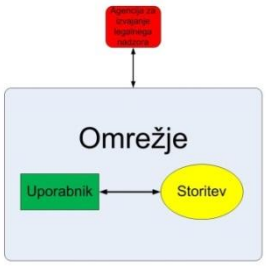
Razlogov za izvajanje legalnega nadzora je več, glavni izmed njih je zagotavljanje varnosti in s tem zaščita infrastrukture. Javni mrežni operater lahko izvaja legalni nadzor v te namene, kakor tudi nadzornik privatnega omrežja za svoje omrežje, razen če mu je to preprečeno iz drugih razlogov. [5]

4.1 Koncepti legalnega nadzora

V tem poglavju bomo govorili o različnih načinih izvajanja legalnega nadzora.

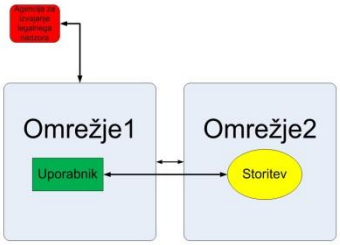
4.1.1 Notranje in zunanje prestrezanje

O notranjem prestrezanju podatkov govorimo, ko ima agencija, ki izvaja legalni nadzor, neposredni dostop do omrežja (njegovih elementov) v katerem se nahaja tako uporabnik telekomunikacijske storitve, kot storitev, ki jo ta uporabnik uporablja. [5]



Slika 5: Notranje zajemanje

Potreba po izvajanju zunanjega prestrezanja telekomunikacijskih podatkov se pojavi v primeru, ko nek uporabnik v omrežju uporablja storitev, ki se ne nahaja v le-temu. Omrežje v katerem deluje storitev pa v našem primeru ni podvrženo legalnemu nadzoru.



Slika 6: Zunanje zajemanje

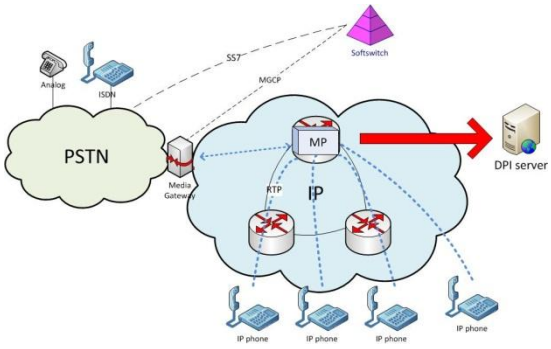
4.1.2 Centraliziran in decentraliziran pristop

O **centraliziranem pristopu** govorimo, ko ves promet usmerimo na eno točko in ga iz te točke odvajamo na mesto, kjer izvajamo legalni nadzor. [6]
Prednost:

- vsi podatki dostopni na eni točki v omrežju

Slabost:

- zmožnost medijskega posredniškega strežnika



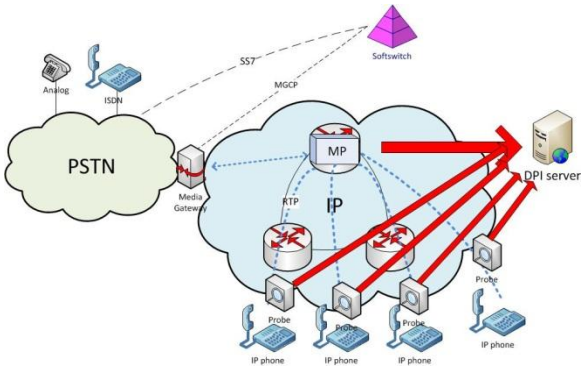
Slika 7: Centraliziran pristop

Pri **decentraliziranem pristopu** na robove omrežja postavimo sonde, ki odvajajo promet od centralnega strežnika za nadaljnjo obravnavo.
Prednost:

- enostavnejša izločitev prometa iz določenega dela omrežja.

Slabost:

- problematika namestitve sond na robove omrežja.



Slika 8: Decentraliziran pristop [6]

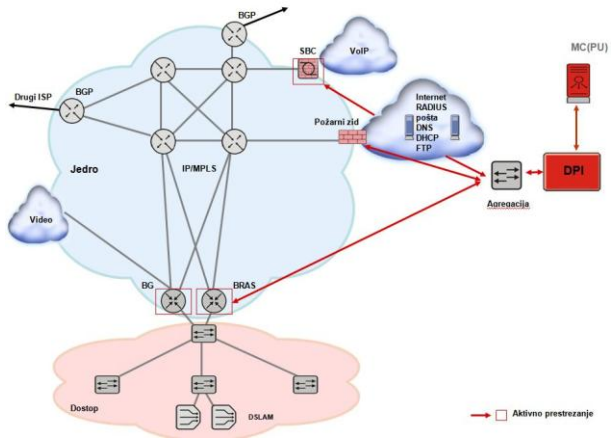
4.1.3 Aktivno, pasivno in hibridno prestrezanje

Izvedba **aktivnega nadzora** pomeni uporabo že **obstoječih mrežnih elementov**, ki aktivno sodelujejo pri identificiranju, podvojevanju in posredovanju komunikacijskega prometa funkciji posredovanja. To pomeni, da je oprema za legalni nadzor hkrati tudi oprema, ki sodeluje pri prenašanju komunikacijskih podatkov. Ta oprema mora imeti vmesnik, ki podpira protokolni sklad za legalni nadzor. Preko tega vmesnika lahko mediator dostopa do naročnika.
Prednosti:

- nizki stroški,
- uporaba omrežnih naprav operaterja.

Slabosti

- ni zavedanja ISP,
- potrebna prekonfiguracija operaterjevih mrežnih naprav.



Slika 9: Aktivno prestrezanje

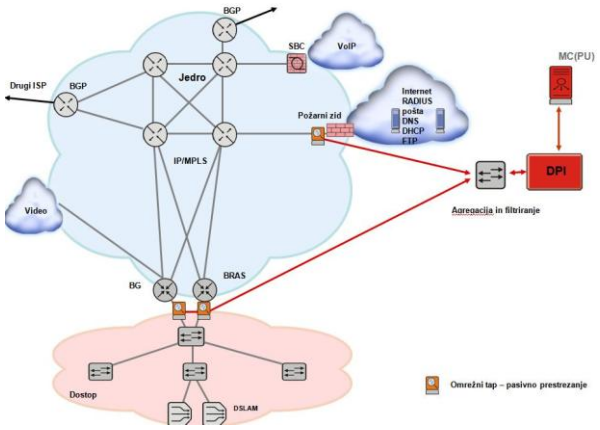
Nasprotje izvedbe aktivnega nadzora je **pasivno prestrezanje** komunikacij. Ta način prestrezanja je podoben decentraliziranemu pristopu, o katerem smo že nekaj povedali. V tej različici izvedbe sistema za legalni nadzor moramo v omrežje vstaviti **dodatno namensko opremo**, ki je sposobna zajemanja in nadaljnjega posredovanja zajetih podatkov agenciji, ki izvaja legalni nadzor. Za razliko od aktivnega prestrezanja, v tem primeru prestrezamo ves promet v točki zajema podatkov. Novejše naprave (tako imenovane sonde) so sicer sposobne do določene mere izločit nepomemben promet, vendar ne v tej meri, kot to omogočajo naprave, ki sodelujejo pri aktivnem prestrezanju. To prinese na strani agencije, ki sprejme zajet promet veliko dodatnega procesiranja.

Prednosti:

- zavedanje ISP.

Slabosti:

- visoki stroški,
- dodatna strojna oprema (TAP-i).



Slika 10: Pasivno prestrezanje

Poleg aktivnega in pasivnega nadzora obstaja še t.i. hibridni nadzor, ki predstavlja kombinacijo aktivnega in pasivnega nadzora. To pomeni, da imamo v omrežju operaterja tako elemente, ki omogočajo aktivni nadzor, kot elemente, ki omogočajo pasivni nadzor.

5 Zaključek

Ugotovili smo, da je vpogled v vsebino elektronskih komunikacij pomembna tehnologija pri izvajanju legalnega nadzora. Uporabimo jo lahko tudi pri drugih operacijah v omrežju. DPI tehnologija nam omogoča operacije kot so nadziranje prometa (angl. traffic throttling), izboljšuje varnost, saj zaznava vdore, pomaga pri detekciji nelegalnih vsebin, analizira obnašanje uporabnikov s strani ponudnikov internetnih storitev, ki jim posledično ponujajo ustrezne storitve, kakor tudi omogoča zniževanje stroškov pri le-teh.

Ugotovili smo, da je legalni nadzor potreben za uveljavljanje zakonodaje v elektronskih komunikacijah, saj na ta način lahko prestrežemo relevantne informacije, ki kršijo zakone v določeni državi. S pomočjo legalnega nadzora lahko tudi izboljšamo varnost v omrežju in preprečimo napade.

6 Literatura

[1]. http://en.wikipedia.org/wiki/Deep_packet_inspection

[2]. http://www.thetanetworks.com/resources/shallow_packet_inspection.html

[3]. http://www.lightreading.com/document.asp?doc_id=139773

[4]. http://www.dpacket.org

[5]. http://www.aktivno.si/ai/sl/688-legalni-nadzor-v-omrezjih-ip

[6]. http://lms.uni-mb.si/vitel/14delavnica/predstavitve/ppt-naim_maloku.pdf