Sybil napadi u društvenim mrežama i zaštita od njih

Antun Razum
Voditelj: prof. dr. sc. Siniša Srbljić

6. lipnja 2016.

Fakultet elektrotehnike i računarstva

Sadržaj

Uvod

Sybil napadi

- *Sybil* – prema istoimenoj knjizi o ženi s disocijativnim poremećajem osobnosti
- napadi na distribuiranim sustavima poput senzorskih i *peer-to-peer* mreža
- napadač stvara velik broj lažnih identiteta preko kojih utječe na ponašanje sustava
- danas veoma aktualno na društvenim mrežama

Primjeri

- širenje spam sadržaja na društvenim mrežama – često maliciozni sadržaj
- korištenje velikog broja lažnih identiteta za postizanje nekih "ciljeva", npr. glasanje, podizanje reputacije, lažno prijavljivanje sadržaja
- prosječno 20% zahtjeva za prijateljstvo od lažnih profila bude prihvaćeno

Povijest i motivacija

Središnji autoritet

- izdaje i provjerava podatke jedinstvene stvarnom čovjeku
- zahtijevanje osobnih podataka (npr. broj osobne iskaznice) ili plaćanje registracije
- nepoželjno jer odbija velik broj korisnika
- problem oko odabira središnjeg autoriteta
- može biti *signle point of failure*

Decentralizirani pristupi

- povezivanje korisnika s IP adresom – lagano se može ukrasti i iskoristiti veći broj različitih IP adresa
- zagonetke koje zahtijevaju ljudski napor (npr. *CAPTCHA*) – predstavljanje na vlastitoj stranici ili plaćanje jeftinih servisa za rješavanje zagonetki

Motivacija

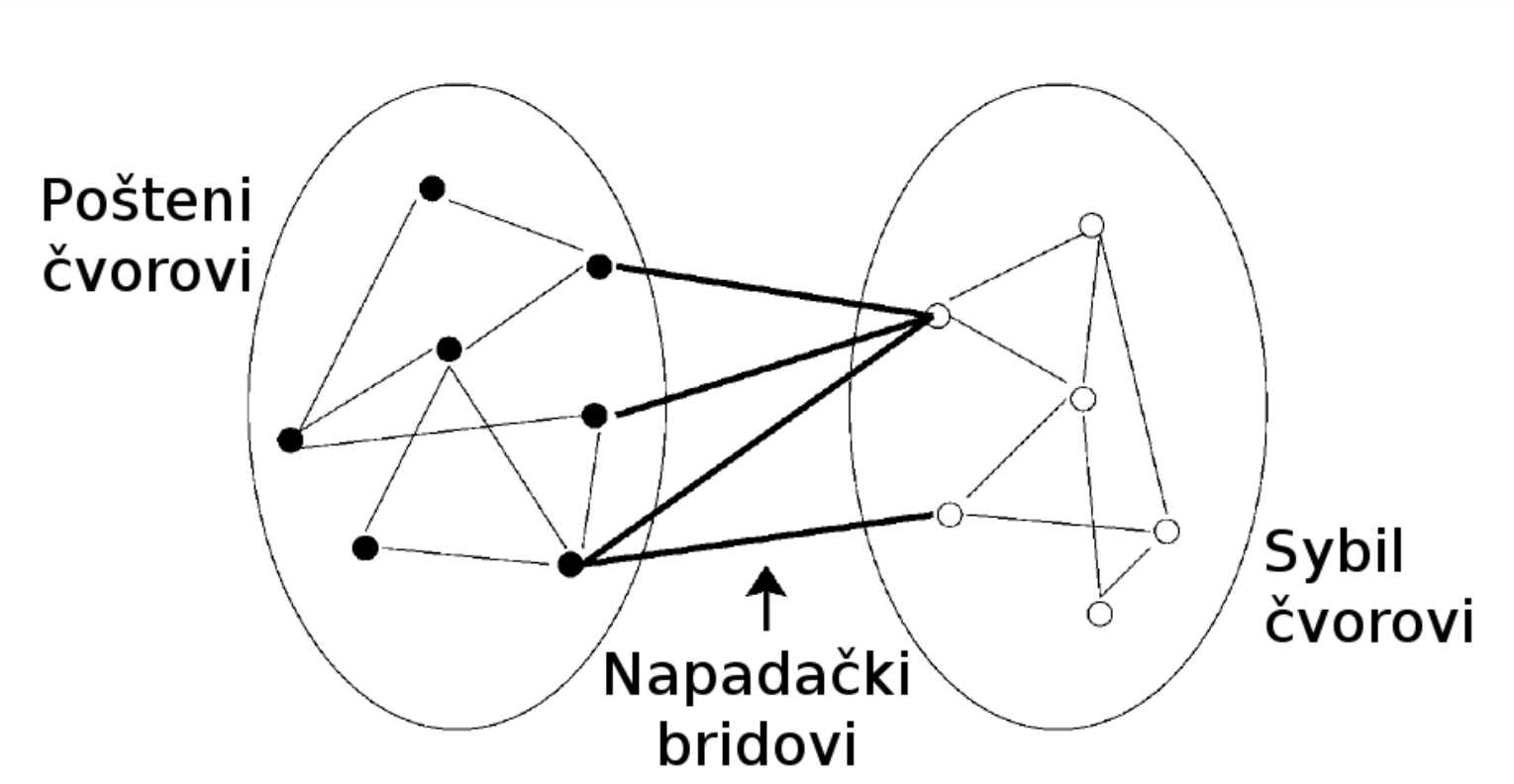
- predložene metode su ograničene
- omogućuju smanjenje, ali ne i eliminaciju sybil napada
- potrebna obrana temeljena na analizi grafa društvene mreže

Pojmovi i definicije

Model društvene mreže

- neusmjereni bezteˇzinski graf – cˇvorovi su korisnici, a bridovi odnosi među njima, npr. prijateljstva
- *pošteni cˇvorovi* – predstavljaju stvarne korisnike mreže
- *sybil cˇvorovi* – lažni identiteti stvoreni od strane napadača
- *napadački bridovi* – bridovi između sybil cˇvorova i poštenih cˇvorova
- *sybil regija* sastoji se od svih sybil cˇvorova, a *poštena regija* od svih poštenih cˇvorova

Model društvene mreže



Slučajne šetnje

- *slučajna šetnja* – šetnja u grafu s nasumično odabranim prijelazima
- slučajne šetnje su *ergodične* – konvergiraju prema *stacionarnoj distribuciji* kada im duljina teži u beskonačnost

Vrijeme miješanja

- definira se kao najmanja duljina slučajne šetnje kojom se postiže stacionarna distribucija do neke mjere ϵ:

$$T(\epsilon) = \max_i \min\{t : |\pi - \pi^{(i)} P^t|_1 < \epsilon\}$$

- graf s n čvorova je *brzo miješajući* ako je:

$$T(\epsilon) = O(\log n)$$

- dobro povezani grafovi su brzo miješajući

Obrana od sybil napada

Pretpostavke algoritma

- poštena regija je brzo miješajuća
- jedan poznat pošteni čvor
- administratoru je poznata topologija društvene mreže
- veličina sybil regije nije usporediva s veličinom poštene regije
- broj napadačkih bridova je ograničen

Identifikacija sybil čvorova: prva faza

- rade se slučajne šetnje od poznatog poštenog čvora
- konačni čvorovi šetnji podvrgavaju se stacionarnoj distribuciji i s visokom su vjerojatnošću pošteni
- početni pošteni čvor i dobiveni konačni čvorovi su *čvorovi sudci*
- iz svakog od njih napravi se veći broj šetnji različitih duljina
- za svaku duljinu pamti se broj čvorova s frekvencijom većom od nekog praga

Identifikacija sybil čvorova: druga faza

- napravi se veći broj šetnji određene duljine iz *osumljičenog čvora*
- izračuna se broj čvorova čija je frekvencija veća od praga
- ako je dobiveni broj dovoljno manji od broja izračunatog u prvom koraku za odgovarajuću duljinu, čvor je sybil čvor
- u suprotnom postupak, povećava se duljina šetnje i postpuak se ponavlja
- ako se dođe do gornje granice za duljinu šetnje, čvor je pošten

Pronalazak sybil grupa: prva faza

- *mrtva šetnja* je ona koja ponovo prolazi već prijeđenim čvorom
- rez između sybil i poštene regije je mali – ako je duljina šetnje dovoljno velika, omjer mrtvih šetnji biti će blizak 1
- u prvoj se fazi određuje duljina šetnje kako bi skup šetnji pokrio barem sybil grupu

Pronalazak sybil grupa: druga faza

- zadaća druge faze je uklanjanje poštenih čvorova iz nađene sybil grupe
- *provodljivost* podgrafa – omjer broja bridova koji ga spajaju s ostatkom grafa i bridova u njemu
- čvorovi se u dobivenoj grupi sortiraju silazno po frekvenciji i višestrukim se iteracijama dodaju sve dok se provodljivost smanjuje

Rezultati

Korištene metode i skupovi podataka

- stvarni skupovi podataka iz društvenih mreža Facebook i Orkut s preko 3 milijuna čvorova
- dva modela stvaranja sybil regija – *preferencijalno vezivanje* (PA) i *Erdös-Rényi* (ER)
- PA je model s "prirodnom" zastupljenošću stupnjeva čvorova, a ER je potpuno nasumičan
- stvorene sybil regije imale su 10,000 čvorova i 1,000 napadačkih bridova

Rezultati

R	Orkut				Facebook			
	PA model		ER model		PA model		ER model	
	F^+	F^-	F^+	F^-	F^+	F^-	F^+	F^-
1000	0	0.02%	0	0.28%	0	0.22%	0.1%	0.54%
1500	0	0.02%	0	0.32%	0.3%	0.12%	0.2%	0.44%
2000	0	0	0	0.22%	0.5%	0.04%	0.5%	0.4%

Zaključak

Zaključak

- algoritam za obranu od sybil napada temeljen na slučajnim šetnjama i algoritamskim svojstvima grafova
- identifikacija sybil čvorova i pronalazak grupa koje ih okružuju
- algoritam se pokazao veoma učinkovitim i brzim prilikom testiranja na skupovima podataka iz stvarnog svijeta

Pitanja?

Hvala na pažnji!