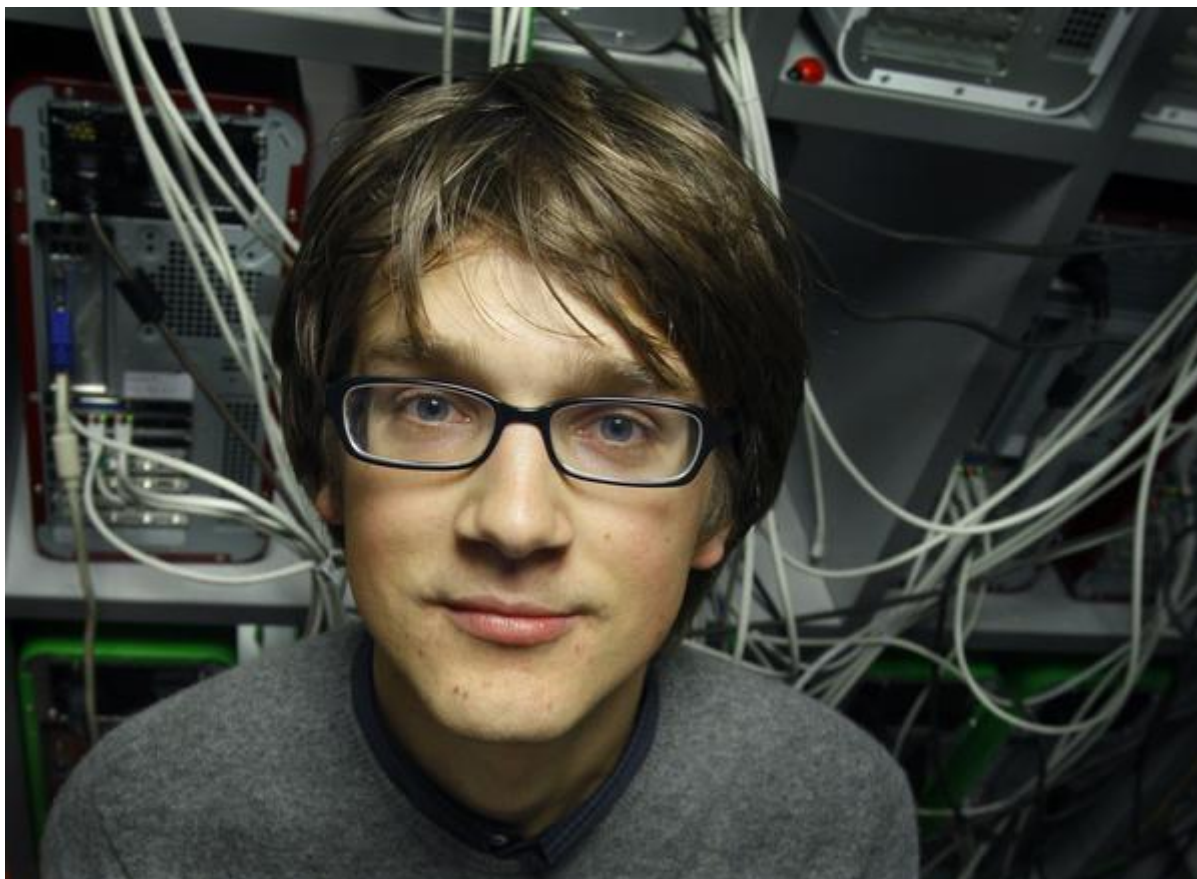


Iskanje orodja, s katerim bomo izluščili modrost množic

Ko uporabnik računalniškega družabnega omrežja Facebook slovesno potrdi svojega tisočega elektronskega prijatelja in se prepusti kratkemu trenutku zadovoljstva, preden začne zbirati novo stotino elektronskih stikov, niti ne pomisli, da njegovo ravnanje ves čas skrbno spremljajo analitiki.

Lenart J. Kučič, sob, 09.01.2010 20:28



Jure Leskovec, profesor na Stanfordu

foto: Jure Eržen/DELO

Ko uporabnik računalniškega družabnega omrežja Facebook slovesno potrdi svojega tisočega elektronskega prijatelja in se prepusti kratkemu trenutku zadovoljstva, je povedal devetindvajsetletni računalniški strokovnjak Jure Leskovec, ki na kalifornijski univerzi Stanford raziskuje velika računalniška, informacijska in družbena omrežja. Ugotoviti znamo, kam je treba namestiti senzorce v vodovodnem omrežju, da pravočasno zaznamo zastrupitev vodnih virov. Določimo lahko, komu

Če pridobimo dovolj podatkov in znamo uporabiti današnja računska orodja, je mogoče z analizo omrežij uspešno analizirati številna področja človeškega delovanja, je povedal devetindvajsetletni računalniški strokovnjak Jure Leskovec, ki na kalifornijski univerzi Stanford raziskuje velika računalniška, informacijska in družbena omrežja. Ugotoviti znamo, kam je treba namestiti senzorce v vodovodnem omrežju, da pravočasno zaznamo zastrupitev vodnih virov. Določimo lahko, komu

moramo prišepniti govorico, da bo dosegla kar največji učinek, ali pokazati, kdo ima največ vpliva na medijskem ali političnem trgu. Naučili smo se izračunati, kam postaviti oglas, da bo dosegel kar največ potencialnih kupcev, in predvideti, kateri člani omrežja bodo med sabo prijatelji ali sovražniki. Le prihodnosti trenutno še ne znamo napovedati, je priznal. Zato si je prav to področje izbral za svoj največji raziskovalni izziv.

Pričakovanje, da je mogoče s številkami razložiti družbo ali celo napovedovati prihodnost, že stoletja zaposluje politike, igralce na srečo, ekonomiste, vojskovodje in znanstvenike. Zdaj ste na to področje z analizo elektronskih omrežij vstopili še računalnikarji. Zakaj? Ker se je v zadnjih dvajsetih letih toliko družbe preselilo v elektronske svetove?

Ko opazuješ dovolj velik vzorec ljudi, ugotoviš, da se začnejo človeške družbene mreže obnašati v skladu z določenimi splošnimi pravili, čeprav ima vsak posameznik svojo osebnost in identiteto. Če si izposodim primer iz fizike: ljudi je mogoče včasih modelirati zelo podobno kot pline (nasmešek). Na področju analize omrežij se namreč stikajo tri znanosti: fizika, družbene vede in računalništvo. Fiziki na omrežja gledajo kot na kompleksne sisteme in poskušajo ugotoviti, ali je iz obnašanja enega delca ali manjšega sistema mogoče razumeti ali celo napovedati obnašanja večjih fizikalnih in bioloških sistemov. Podobno razmišljajo tudi družboslovci, ki raziskujejo odnose med posamezniki in večjimi družbenimi sistemi. Računalnikarji smo poskušali z analizami omrežij sprva odgovoriti predvsem na tehnična vprašanja - kako pohitriti prenos podatkov prek računalniških mrež, ustaviti širjenje računalniških virusov in podobno. Z razmahom interneta in razvojem elektronskih družabnih omrežij pa se je družba dejansko začela seliti na naš teren, saj so se naše tehnologije in storitve razširile v vse pore življenja več milijard posameznikov.

Ti posamezniki pa ustvarjajo neskončne količine podatkov ...

... ki jih je mogoče zelo natančno meriti, jim slediti, jih zbirati in analizirati. Poleg tega ti podatki niso zgolj pasivne številke, ampak nam povedo veliko o aktivnem obnašanju uporabnikov. Z analizo blogov, uporabe spletnih medijev in družabnih omrežij izvemo, kaj ljudi zanima, kaj iščejo, o čem pišejo, s čim se izražajo, kdo komunicira s kom, kakšni so njihovi odnosi in na kakšna sporočila se odzivajo. Ti podatki omogočajo zelo natančne posnetke njihovega posamičnega delovanja, kažejo pa tudi, kako ta posamična dejanja vplivajo na širše omrežje.

Primerjava je privlačna: ljudje kot posamezne molekule, ki oblikujejo »plin« - družbo z določenimi lastnostmi. Pa je res tako preprosto?

Analize, ki jih znamo danes izdelati analitiki omrežij, so vsekakor bistveno bolj površinske od tistih, ki jih izvajajo sociologi, antropologi ali psihologi. Tudi zato so družboslovci do takega pristopa še nekoliko zadržani, saj jih zanimajo predvsem posamezniki kot taki in ne kot nekakšna abstraktna mrežna vozlišča, povezave in uporabniške šifre. Vendar se tudi zavedajo, da lahko z analizo omrežij raziskujejo stvari, ki prej niso bile mogoče, saj so bile klasične družboslovne raziskave omejene z razmeroma majhnimi vzorci - nekaj sto do morda tisoč ljudi. Take raziskave nam povedo marsikaj o delovanju manjših človeških omrežij, vendar teh pravil ni mogoče preprosto posplošiti na večja omrežja. Ko omrežje doseže določeno velikost, se začne obnašati bistveno drugače kot prej. To pa lahko ugotovimo šele, ko analiziramo omrežja, v katerih je povezanih na milijone uporabnikov.

Katere nove ugotovitve je prinesla analiza velikih omrežij?

Lahko smo, denimo, potrdili ali ovrgli nekatere stare teorije. Pred dobrim letom smo na 250 milijonih uporabnikov Microsoftovega omrežja MSN ugotovili, da Milgramova in Traversova teza o »šestih stopnjah ločenosti«, po kateri vsakega poljubnega človeka na planetu do drugega loči v povprečju le šest stikov rok oziroma poznanstev, dejansko drži, saj so bili uporabniki omrežja med sabo v povprečju oddaljeni le za nekaj več kot šest povezav. Prav tako smo potrdili tezo britanskega antropologa Robina Dunbarja, ki je na podlagi povprečnih velikosti skupin, v katerih živijo posamezne sesalske vrste, in velikosti njihovih možganskih neokorteksov postavil hipotezo, da ljudje lahko sklenemo in vzdržujemo do približno 150 stikov, preden dosežemo svojo kognitivno mejo. Naše analize so pokazale, da je prav 150 povezav tista vrednost, pri katerem se začnejo mreže drugače obnašati, kar pa se vidi šele na zelo velikih vzorcih.

Kljub temu imajo nekateri uporabniki Facebooka in drugih družabnih omrežij tudi po tisoč »prijateljev« in več ...

To načtenja zelo zanimivo raziskovalno vprašanje. Po nekaterih teorijah so se naši možgani prilagodili na dejstvo, da smo velik del svoje evolucije preživeli v skupnostih, ki so imele v povprečju po 150 pripadnikov, zato večjih skupnosti pač nismo sposobni trajno ohranjati. Po drugi strani pa je mogoče, da so nam elektronska orodja omogočila vzdrževanje bistveno večjega števila stabilnih povezav. Ne vem, saj so Facebook in druga orodja z nami še premalo časa, da bi lahko spremljali evoliucijske spremembe. Prav tako je še neodgovorjeno vprašanje, kako izmeriti vplive tisočih elektronskih prijateljev na dejansko družabno življenje posameznika.

Katera stara prepričanja pa so analize velikih omrežij ovrgle?

Izvedeli smo marsikaj novega o tem, kako ljudje ustvarjamo pozitivne in negativne povezave, komu zaupamo ali kako izbiramo naše prijatelje in sovražnike. Trije ljudje so lahko povezani različno (ozre se po mizi, poišče papirnati prtiček in začne skicirati trikotnike s povezavami). Ena mogoča razlaga človeških povezav je razmeroma preprosta: prijatelj mojega prijatelja je tudi moj prijatelj, sovražnik mojega prijatelja je moj sovražnik, sovražnik mojega sovražnika pa je tudi moj

sovražnik. Vendar se je izkazalo, da ljudje na internetu ustvarjajo povezave drugače. Pozitivne povezave ustvarjajo do ljudi, ki imajo višji status ali ugled, medtem ko negativne povezave ustvarjajo do tistih, ki imajo nižji status od njih samih. Še bolj presenetljivo pa je, da je mogoče le iz lokalne strukture omrežja v več kot 90 odstotkih primerov pravilno napovedati, ali je določena povezava pozitivna ali negativna.

Na podobne modele stavita tudi zavarovalniška in igralniška dejavnost - 90 odstotkov ljudi naj bi se v 90 odstotkih primerov obnašalo 90-odstotno predvidljivo.

Tako nekako. Če vem, v kakšnih odnosih so si posamezni stiki - kar lahko precej natančno izluščimo iz pridobljenih podatkov -, lahko z veliko gotovostjo napovem, kakšnega predznaka bodo nove povezave v omrežju in kako bodo določena dejanja vplivala na delovanje omrežja. Najbolj zanimivo pa je to, da se ta napoved ne spreminja ne glede na tip omrežja. Ljudje se začnejo v velikih omrežjih obnašati predvidljivo, ne glede na to, ali so ta omrežja javna ali anonimna, zasebna ali komercialna ... Zato lahko modele iz enega omrežja natančno uporabim v drugih omrežjih. Tudi zato se trenutno veliko ljudi ukvarja z analizo zelo različnih omrežij, saj je mogoče zelo podobno modelirati marsikaj - kje in komu pokazati oglas, da bo imel kar največ učinka in prodal največ izdelkov, komu posredovati govorico, da se bo najhitreje razširila ... Teh uporab je neskončno dosti - od znanstvenih do takih, ki ustvarjajo dolarje. Trenutno se dogovarjamo o analizi poslovnega družabnega omrežja LinkedIn, v katerem je zbranih več kot 30 milijonov visokokvalificiranih delavcev in ponuja odličen vpogled v trg dela v ZDA. Za vsakega člana omrežja vemo, kakšna je njegova zaposlitvena pot in kje dela, zato bi lahko z analizo omrežja na ogromnem vzorcu natančno proučevali, kako se je na ameriškem trgu dela odrazila gospodarska kriza, kdo je izgubljal delo, katere panoge so zaposlovale in kakšna je bila smer teh zaposlitvenih migracij. Na neki točki bi bilo morda mogoče celo napovedovati, kdo bo verjetno izgubil službo in kdo razmišlja o menjavi, čeprav se tega sam še niti ne zaveda.

Analiza omrežij lahko danes zelo natančno opiše, kakšne so lastnosti teh mrež, kako dolge so povezave med uporabniki, kako se širijo, kje so ključna vozlišča, katerih tipov povezav je največ in kakšne povezave bodo najverjetneje nastale v določenih okoliščinah. Kakšne pa so omejitve te metode? Česa nam analiza omrežij ne pove?

Analitiki omrežij smo trenutno predvsem opazovalci, ki znamo razmeroma dobro opisati preteklo dogajanje. Sam pa priznam, da me bolj zanima, kako aktivno oblikovati prihodnost. Zato bi se zelo rad lotil družbenih poskusov na ljudeh, denimo izgradil dva vzporedna virtualna svetova in ugotavljal, kako se ljudje obnašajo v njih in kako na to obnašanje vplivajo razlike, ki bi bile vgrajene v ta dva sistema. Industrija se s tovrstnimi poskusi ukvarja že veliko let. Amazoni, google in druga spletna podjetja redno preizkušajo, kako uporabniki sprejmejo določeno spremembo - ali kupijo več knjig, če je gumb za dokončanje nakupa postavljen višje ali nižje na strani ...

... ali števec prijateljev na Facebooku vpliva na to, da se bodo uporabniki bolj pridno mrežili, ker jim več prijateljev dviguje status ...

Vsekakor. Facebook se trudi, da bi uporabniki pri njih preživel kar največ časa, zato uporabljajo različne prijeme, kako jih obdržati - dodajajo števec prijateljev, možnosti komentiranja, klepeta, pošiljanja darilc in podobno. Ta orodja niso naključna. Enemu milijonu uporabnikov pokažeš števec prijateljev, drugemu milijonu ga skriješ, nato pa analiziraš razlike in ugotoviš, kaj bolj ustreza tvoji strategiji in poslovnim interesom. Te analize so izjemno natančne, saj nekateri ponudniki preizkušajo skoraj vse - od barve določenih gumbov do pisav. Vendar so to že podrobnosti. Mene bolj mika pravi socialni inženiring - kako in ali je možno ustvariti ter kontrolirano vzgajati spletno skupnost v določeno smer. Ali je možno postaviti sistem, da se bo skupnost razvijala tako, kot hočem, in ne naključno. Ljudi kot posameznikov verjetno nikoli ne bomo mogli modelirati. Mislim pa, da lahko precej modeliramo družbo kot celoto.

Ali ni to nevarna želja? Ljudje si včasih predstavljajo, da je mogoče socialni inženiring uporabljati za »sistemsko« urejanje družbe, kar se je v preteklosti že nekajkrat slabo končalo ...

Mene zanima, kako zasnovati orodja, s katerimi bomo ljudem pomagali in znali nevsiljivo izluščiti neko pasivno globalno znanje, modrost množic, ne pa to, kako nadzorovati ljudi, jim slediti, brati pošto in kukati v spalnice. Prepričan sem, da je mogoče z boljšim razumevanjem omrežij izgraditi marsikaj dobrega. To me žene in motivira.

In upravičuje tudi prijeme, s katerimi bi kot prihodnji družbeni inženir poskušali vplivati na posameznike, da sledijo vašemu cilju?

Moram se popraviti, če je tako zvenelo, ampak moč družbenega inženiringa nikakor ni neskončna. Seveda obstajajo načini, kako spodbujati določeno uporabniško ali potrošniško obnašanje, vendar so zelo omejeni. Poleg tega raziskave omrežij jasno kažejo, da sistema in posameznika ni mogoče obravnavati kot dveh ločenih enot, saj vsak posameznik s svojimi dejanji vpliva na dejanja drugih okoli sebe in soustvarja celoten sistem. Predvsem pa nikoli ne moremo predvideti rojstva izjemnih posameznikov, ki bodo nepredvidljivo radikalno spremenili obstoječe sisteme in pravila. Izjemnih dogodkov verjetno nikoli ne bo mogoče napovedati, sprožiti ali preprečiti.

Za večjo natančnost svojih prihodnjih raziskav boste potrebovali dostop do vse več podatkov o uporabnikih, kar bo na neki točki nujno začelo posegati v njihovo zasebnost. Hkrati se bodo povečevale zahteve držav in komercialnih naročnikov po natančnih analizah obnašanja državljanov, potrošniških navad in podobno. Kako se bi kot raziskovalec odzvali na take ponudbe in projekte?

Podatke in orodja, s katerimi delamo, bi načeloma lahko zlorabili. Tega se zelo dobro zavedam, a lahko le ponovim, da me kot znanstvenika to ne zanima in da pri svojem delu zelo pazim, da ne posegam v zasebnost uporabnikov. Prav tako nimam nobene želje po preganjanju teroristov, čeprav jih tajne službe iščejo tudi z mrežno analizo elektronskih komunikacij. Računalnikarji smo pogosto v nevhvaležni vlogi, saj smo se zaradi naših orodij znašli v prvih bojnih vrstah. Meje možnega pomikamo naprej in se med prvimi srečujemo z dilemami, kaj se na spletu sme in ne sme početi, katere podatke o uporabnikih shranjevati in kako jih obdelovati. Zelo dobro se zavedamo, kako veliko moč ima Google, ki mu z vsakim iskanjem ali elektronskim pismom povemo, kaj počnemo in kaj nas zanima. Prav tako pa smo se naučili, da gre pri posegih v zasebnost pogosto za predvsem psihološko vprašanje, kaj se nam zdi v nekem obdobju sprejemljivo in kaj ne. Davčna uprava pozna naše finance bolje od nas, banke nas izjemno podrobno preverijo, preden nam odobrijo kredite, a smo se na to navadili. Na internetu se ta pravila šele oblikujejo.

Več v Sobotni prilogi Dela