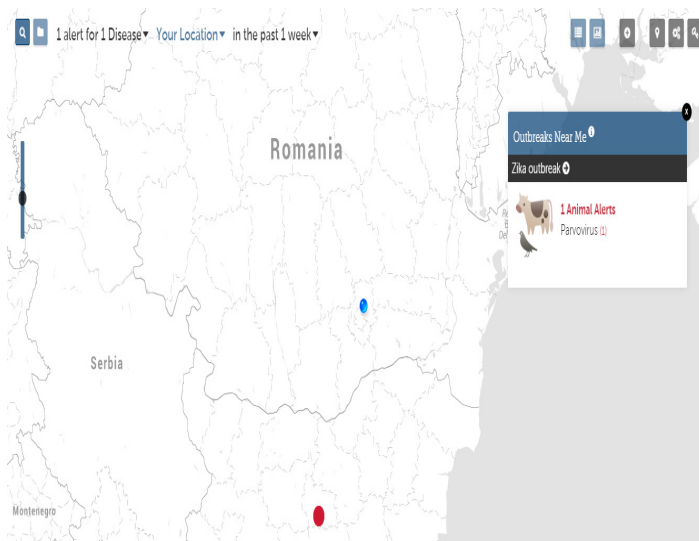


ACADEMIA DE STUDII ECONOMICE

CIBERNETICĂ, STATISTICĂ ȘI INFORMATICĂ ECONOMICĂ



**Proiect Modelarea
proceselor economice**

Popa Ionela Maria

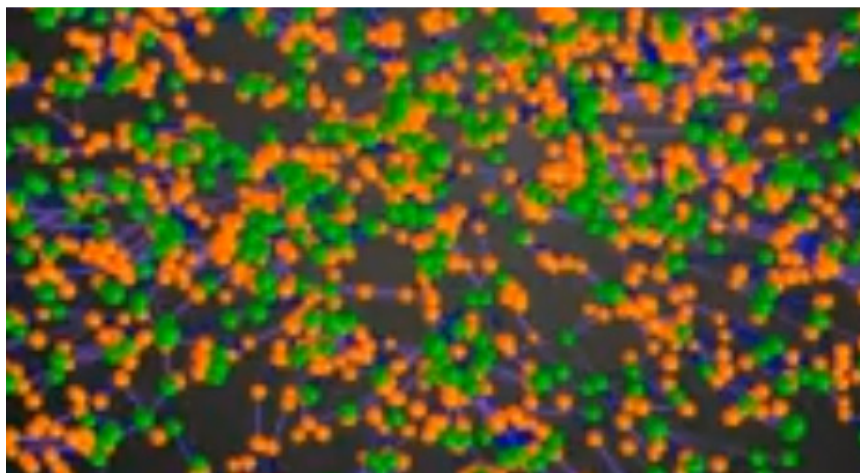
Cibernetica si Economie Cantitativa, master An II

grupa 1093

Cum prezic rețelele sociale epidemiile?

Intr-un studiu realizat de Nicholas Christakis acesta vorbește despre cum și de ce oamenii se organizează în rețele sociale. Prin rețelele sociale nu înțelegem recenta varietate online ci, mai degrabă, de tipul de rețele sociale în care oamenii se organizează de sute de mii de ani, încă de când au apărut în savana africană.

Deci o persoană formează relații de prietenie, de colegialitate, relații între frați sau de rudenie, cu alți oameni, care, la rândul lor, stabilesc aceleași tipuri de relații cu alți oameni. Și acest proces se extinde la nesfârșit. Ce obținem este o rețea care arată astfel:



Fiecare punct reprezintă o persoană. Fiecare linie care le unește reprezintă relația dintre acele persoane și astfel se stabilesc diferite tipuri de relații. Și putem obține o reprezentare a acestei vaste țesături a umanității, în care toți suntem încorporați.

Nicholas Christakis este un sociolog și medic american, cunoscut pentru cercetarile sale asupra rețelelor sociale și factorii socio-economici și biosociale ce determină comportamentul, sănătatea și longevitatea.

Acesta, împreună cu un coleg de-al său James Fowler, au studiat care sunt regulile matematice, sociale, biologice și psihologice care guvernează modul în care aceste rețele sunt formate și care sunt regulile, similare, care guvernează modul lor de operare și modul în care ne afectează viața. Și, și-au pus întrebarea dacă ar putea să profite de pe urma acestor intuiții, în așa fel încât să facă lumea un loc mai bun, să facă ceva mai bine, să remedieze ceva, nu doar să înțeleagă. Astfel, unul din primele lucruri pe care le-au avut în vedere a fost cum am putea prezice epidemii.

Stadiul actual în acest domeniu, de exemplu dacă vorbim de o instituție națională o persoană poate să colecteze date de la medici și laboratoare specializate care raportează prevalența sau incidența anumitor afecțiuni.

Atâți pacienți au fost diagnosticați cu o boală [aici], sau atâți pacienți au fost diagnosticați [acolo], toate aceste date sunt introduse într-un repertoriu central, cu o anumită

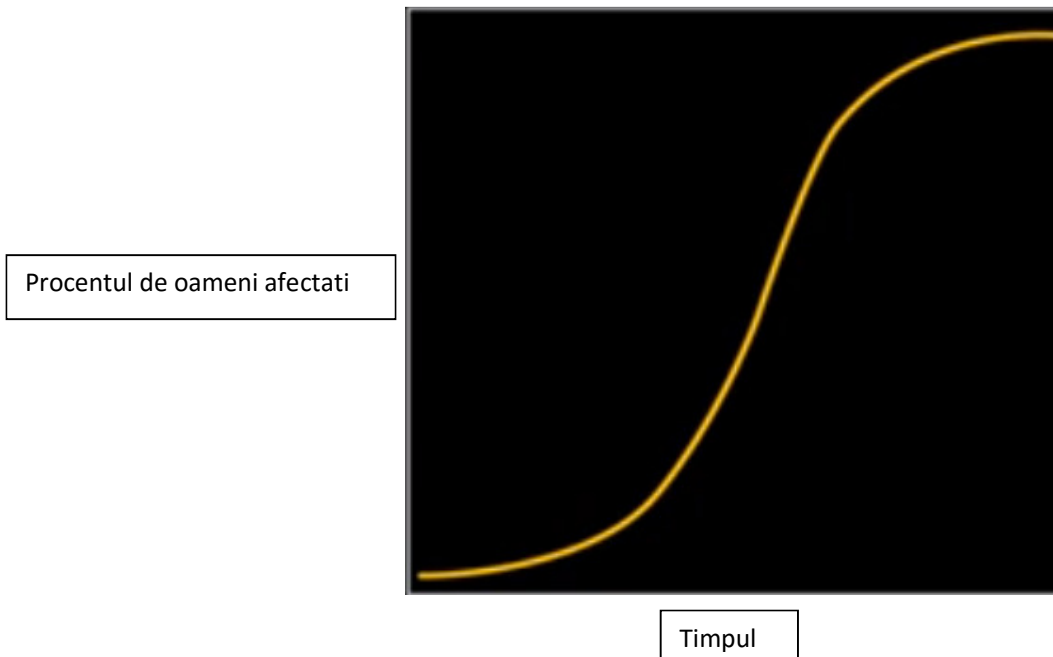
întârziere. Și, dacă totul funcționează cum trebuie, o săptămână sau două mai târziu, se știe unde s-a aflat epidemia azi.

De fapt, a apărut această idee la Google Flu Trends, în ceea ce privește gripa, în care, dacă studiem tiparele căutărilor oamenilor azi, am putea ști unde se afla gripa, care a fost starea epidemiei azi, unde a dominat aceasta epidemiei azi.

Acesta prezintă o metodă prin care putem obține nu doar alertare rapidă în cazul unei epidemii, dar de asemenea și detectarea din timp a acesteia. Și, de fapt, această idee poate fi folosită nu doar pentru predicția epidemiilor microbiene ci pentru orice tip de epidemie.

De exemplu, orice se răspândește printr-o formă de contagiune socială poate fi înțeles în acest fel, de la idei abstracte, la stânga, precum patriotismul, altruismul sau religia, până la practici precum dietele, sau achiziționarea de cărți, consumul de alcool, practicile de siguranță sau produsele pe care oamenii le-ar cumpăra, achizițiile de bunuri electronice, orice are vreo înclinație către o răspândire interpersonală. Un tip de difuzie a inovațiilor poate fi înțeles și prezis de mecanismul pe care acesta îl prezintă.

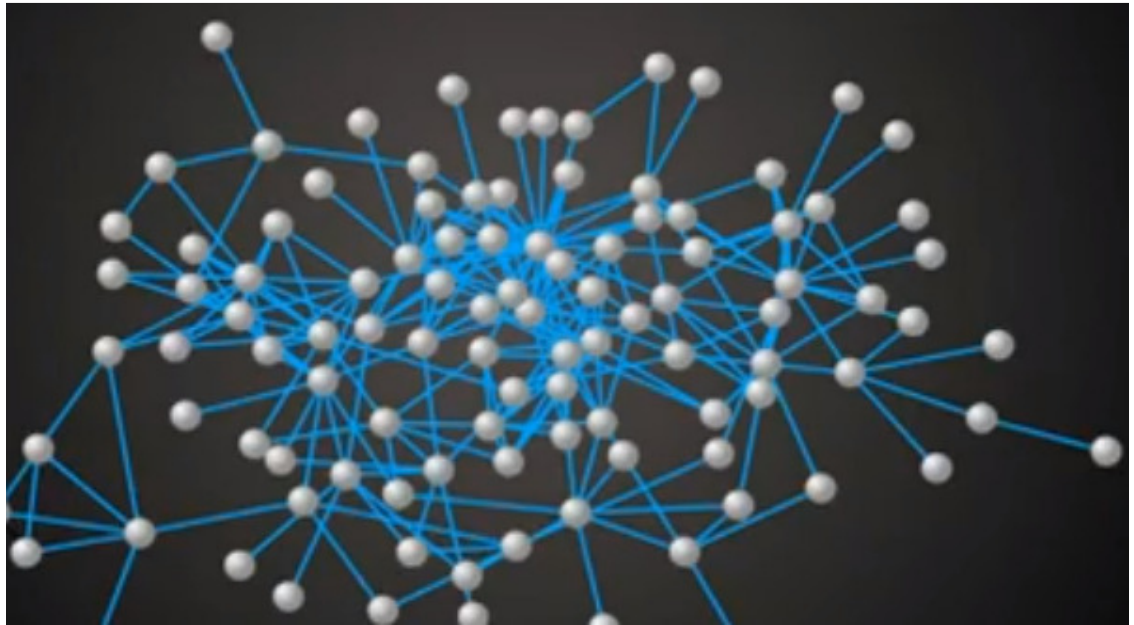
Modul clasic de a privi problema este "răspândirea-inovației" sau curba de adoptare.



Aici, pe axa Y, avem procentul de oameni afectați și pe axa X, avem timpul. La început, nu foarte multe persoane sunt afectate și obținem această curbă clasică sigmoidă sau de forma literei S. Motivul pentru care obținem această formă este că, la început, să spunem, unul sau doi oameni sunt afectați sau infectați și apoi afectează sau infectează doi oameni, care, la rândul lor, infectează patru, opt, 16 și așa mai departe, astfel obținând porțiunea din curbă care indică creșterea.

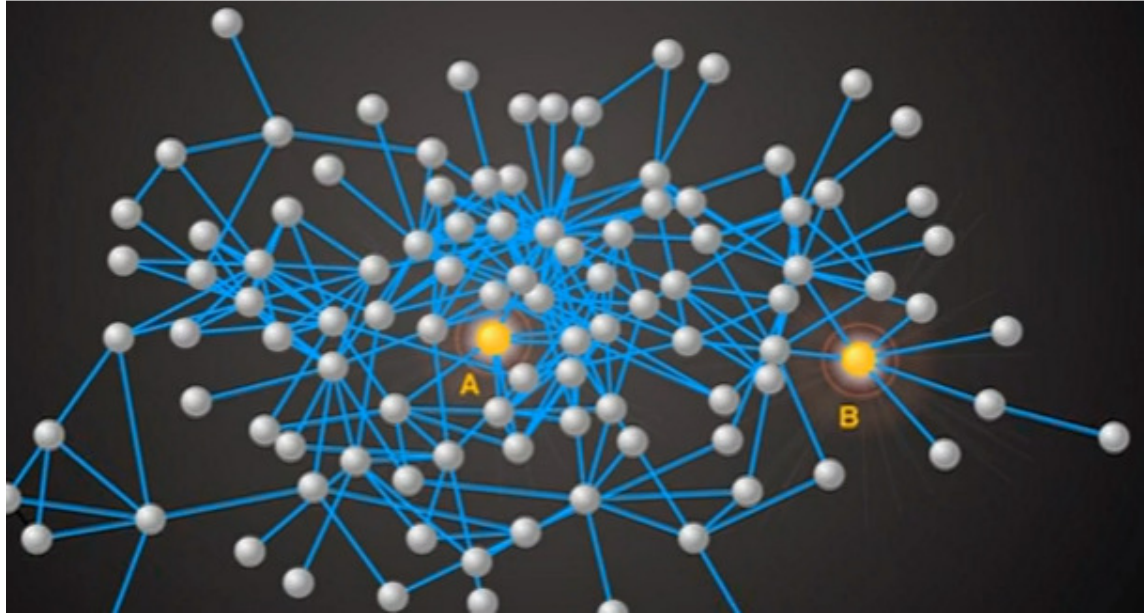
Până la urmă, saturezi populația. Din ce în ce mai puțini oameni mai pot fi infectați și așa obții porțiunea de platou a curbei, rezultatul fiind această formă clasică sigmoidă. Același

lucru este valabil pentru microbi, idei, produse noi, comportamente și altele. Dar aceste lucruri nu se propagă în rândurile populației aleatoriu. Ele, de fapt, se propagă prin rețele. Pentru că, așa cum am spus, noi înșine trăim în rețele și aceste rețele au un anumit tip de structură.



Dacă ne uităm la o rețea ca aceasta. Aici sunt 105 oameni, punctele sunt oamenii, iar liniile reprezintă relațiile de prietenie. Putem vedea că oamenii ocupă locații diferite în cadrul rețelei și sunt diverse tipuri de relații între oameni. Am putea avea relații de prietenie, relații între frați, relații între soți sau între colegi de serviciu, relații între vecini și altele. Și multe alte lucruri se propagă de-a lungul multor altor tipuri de legături. De exemplu, bolile cu transmitere sexuală se vor propaga de-a lungul legăturilor sexuale. Sau, de exemplu, obiceiul de a fuma poate fi generat de anturaj sau comportamentul altruist sau caritabil poate fi influențat de colegii de muncă sau de către vecini. Dar nu toate pozițiile în rețea sunt identice.

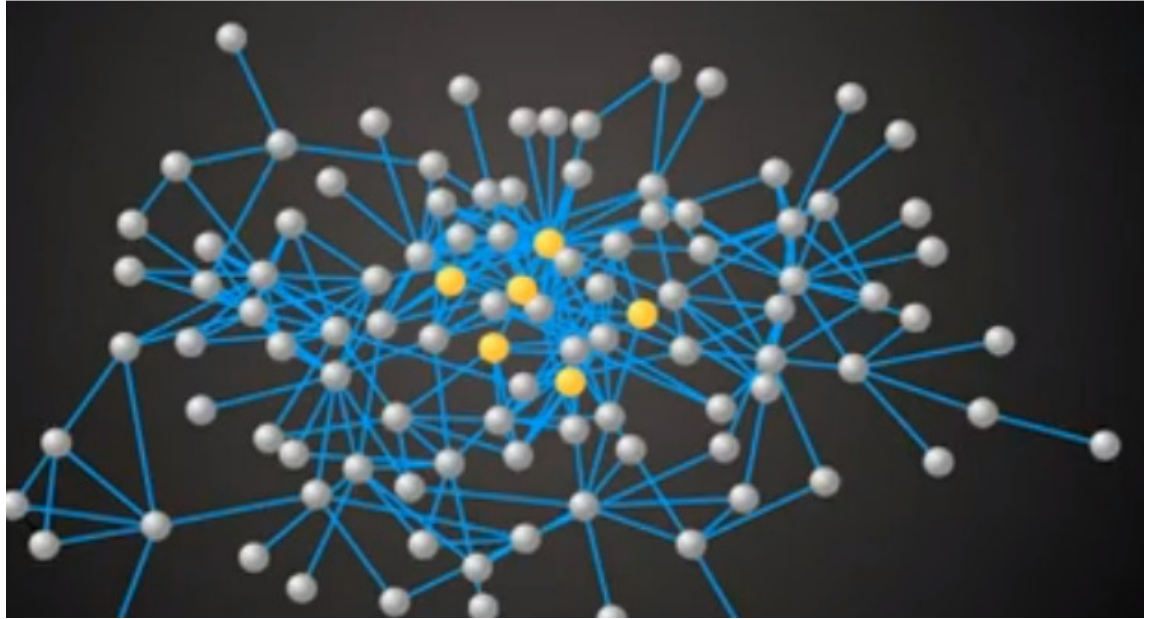
Dacă privim aici, ne putem da seama imediat că persoane diferite au numărul de conexiuni diferit. Unii oameni au o conexiune, alții au două, unii au șase, iar alții au 10. Această caracteristică se numește "gradul" unui nod, sau numărul de conexiuni pe care le are un nod. Dar, în plus, mai e ceva.



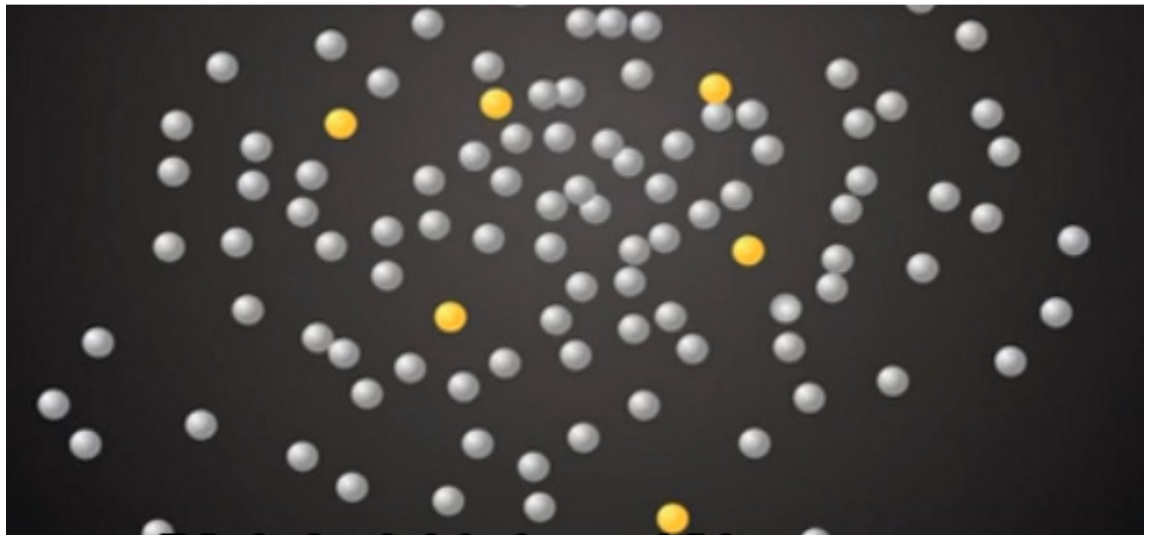
Dacă ne uităm la nodurile A și B, ambele au câte șase conexiuni. Dar dacă am vedea această imagine de undeva de deasupra, am observa că ceva diferă într-o măsură foarte mare în ce privește nodurile A și B. Așa că, a întrebat audiența este cine ar prefera să fie, dacă un microb letal s-ar răspândi prin rețea, A sau B? (Audiența a răspuns: B.) Nicholas Christakis: B, este evident.

B e așezat la marginea rețelei. Apoi, cine ar prefera să fie dacă o bârfă interesantă s-ar răspândi prin rețea? A. Și înțelegem că e mai probabil ca A să intre în contact cu orice se propagă în rețea și să o facă mai devreme în virtutea poziției sale în structura rețelei. A, de fapt, este mai central, și asta poate fi exprimată matematic.

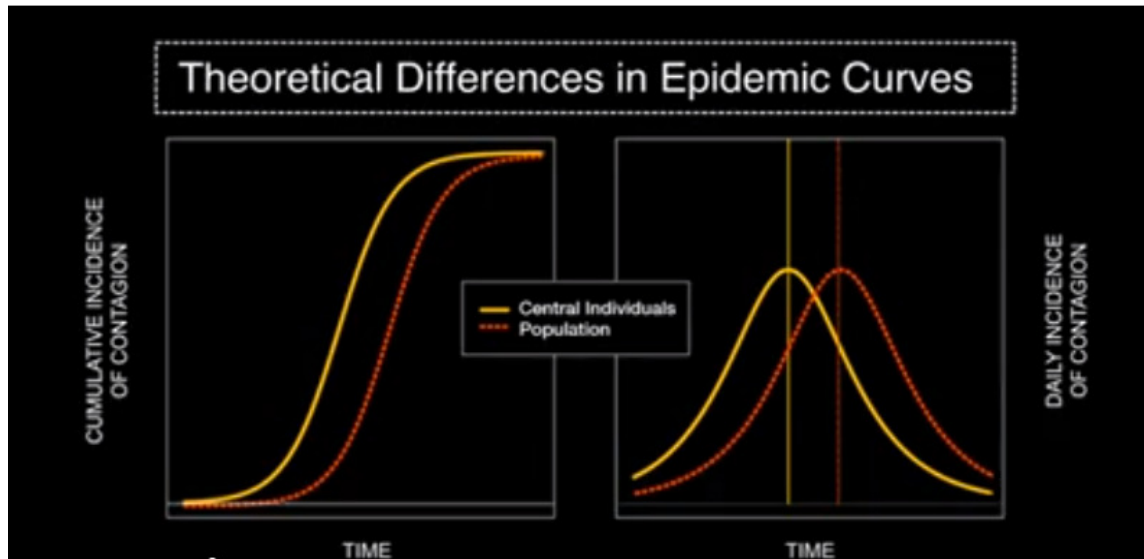
Așa că, dacă vrem să dăm de urma a ceva ce se propagă prin rețea, ce ar trebui să facem este să folosim senzori conectați la indivizii centrali din cadrul rețelei, incluzând nodul A, monitorizând acei oameni care sunt acolo, în centrul rețelei, și, cumva, obținând o detecție precoce a orice s-ar propaga prin rețea.



Adică, dacă ei ar intra în contact cu vreun microb sau o informație, am ști că, ulterior, toți ceilalți vor intra în contact cu acel microb sau acea informație. Această metodă este mult mai bună decât cea bazată pe șase oameni aleși aleatoriu, fără a lua în calcul structura populației.



Și, de fapt, dacă am putea face așa ceva, am obține ceva de genul acesta.



În partea stângă, din nou, avem curba adoptării, în forma de S. Linia punctată roșie indică cum ar arăta adoptarea într-un eșantion aleatoriu, iar linia din stânga, trecută la stânga, indică cum ar arăta adoptarea luând în calcul indivizii centrali ai rețelei. Pe axa Y se află stadiile contagiunii, luate cumulativ, iar pe axa X se află timpul. În partea dreaptă avem aceleași date, dar, de data asta, în funcție de incidența zilnică. Și ce vrem să arătăm aici este ca sunt foarte puțini oameni afectați, și din ce în ce mai mulți și până sus, unde avem vârful epidemiei.

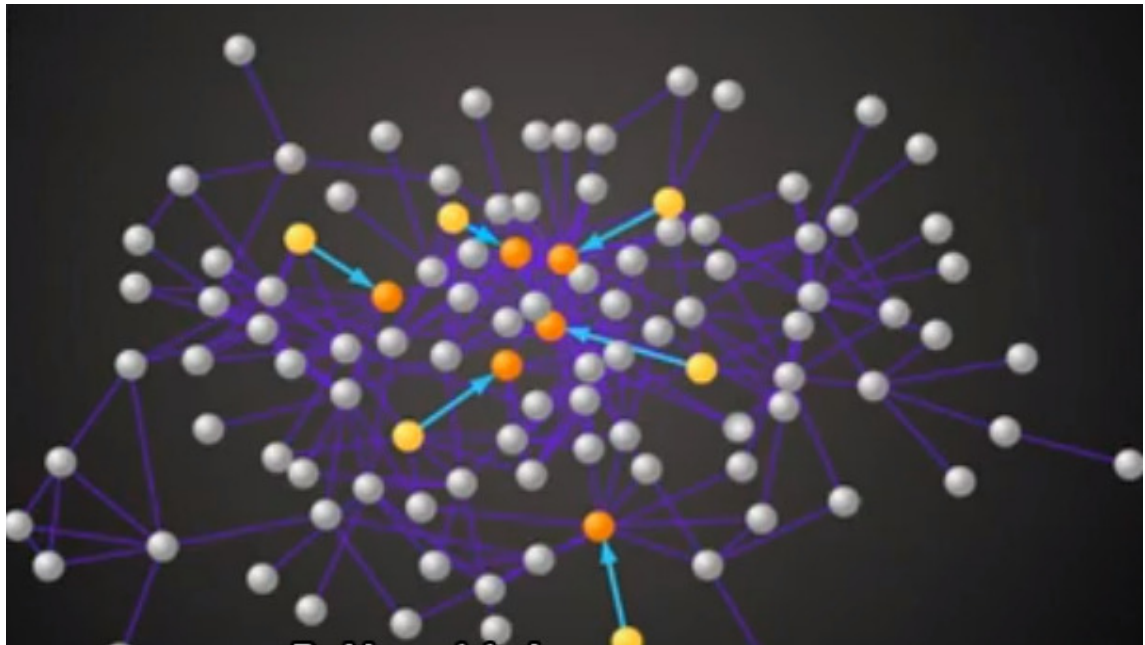
În stânga, avem o reprezentare a ce se întâmplă cu indivizii centrali. Și această diferență în timp dintre cele două e detecția precoce, acel preaviz pe care îl putem obține, despre o epidemie iminentă într-o anumită populație.

Problema, totuși, e că o cartografiere precisă a rețelelor sociale nu este întotdeauna posibilă. Poate fi prea scumpă, [prea dificilă], poate să nu fie etică sau, chiar, poate să nu fie posibilă. Deci cum ne putem da seama cine sunt indivizii centrali într-o rețea fără a cartografia rețeaua?

Acest individ prezintă ideea de a exploata un fapt vechi, cunoscut, despre rețelele sociale, care sună așa: Știi că prietenii tăi au mai mulți prieteni decât tine? Prietenii tăi au mai mulți prieteni decât tine. E cunoscut ca paradoxul prieteniei.

Să ne imaginăm o persoană foarte populară în rețeaua socială, cum ar fi gazda unei petreceri cu sute de prieteni și un tip care are un singur prieten, și alegi pe cineva aleatoriu din populație; acel cineva e mult mai probabil să fie conectat cu gazda petrecerii. Și dacă acea persoană nominalizează gazda ca prieten, atunci acea gazdă are o sută de prieteni, deci are mai mulți prieteni decât persoana aleasă aleatoriu. Și acesta, în esență, este paradoxul prieteniei. Prietenii unor indivizi aleși aleatoriu au grade mai mari și sunt mai centrali, decât indivizii pe care îi alegem aleatoriu.

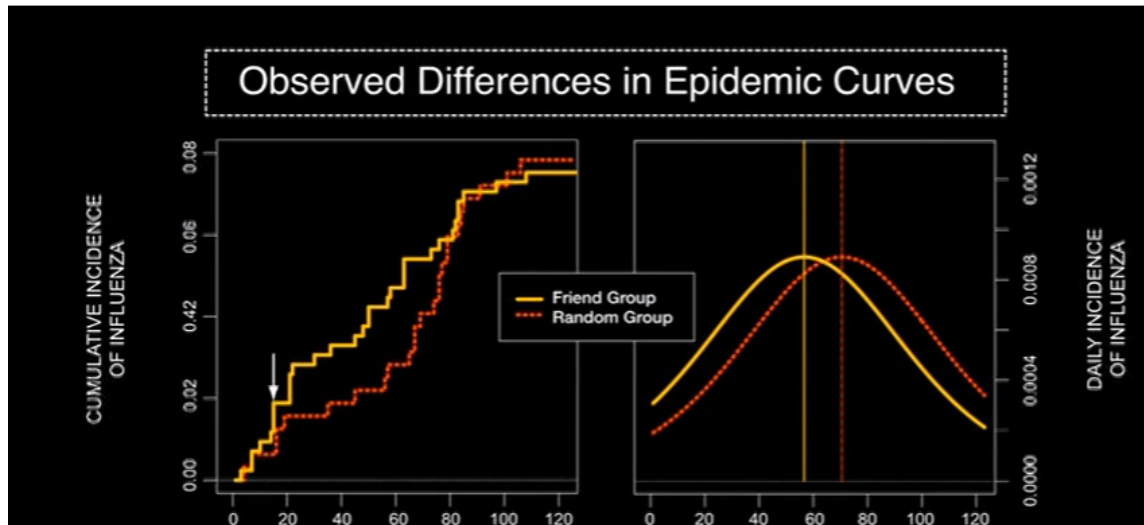
Ne imaginăm doar oamenii de la periferia rețelei. Dacă alegem o persoană, singurul prieten pe care aceasta îl poate nominaliza, este unul care prin definiție are cel puțin doi prieteni, și, de obicei, mai mulți prieteni. Acest fapt apare la orice nod periferic.



De fapt, același lucru se întâmplă pe măsură ce înaintăm înspre interiorul rețelei, pe oricine am alege, atunci când nominalizăm aleatoriu... când o persoană la întâmplare își nominalizează un prieten, ajungem mai aproape de centrul rețelei. S-au gândit să exploateze această idee pentru a studia dacă putem prezice fenomene în interiorul rețelelor.

Pentru că acum, cu această idee, putem extrage un eșantion aleatoriu de oameni, să îi rugăm să își nominalizeze prietenii, și acei prieteni vor fi automat mai centrali, și putem face asta fără să trebuiască să cartografiem rețeaua.

Nicholas Christakis a testat această idee la apariția gripei H1N1 la Colegiul Harvard în toamna și iarna lui 2009. A ales un eșantion aleatoriu de 1.300 de studenți, i-a rugat să își nominalizeze prietenii și i-a urmărit atât pe studenții din eșantion, cât și pe prietenii lor zilnic să vedă dacă contractează sau nu gripa. A făcut acest lucru pasiv, observând dacă au apelat la serviciile medicale ale universității. De asemenea, i-a rugat să le scrie prin e-mail de câteva ori pe săptămână. S-a întâmplat exact ce a prezis.

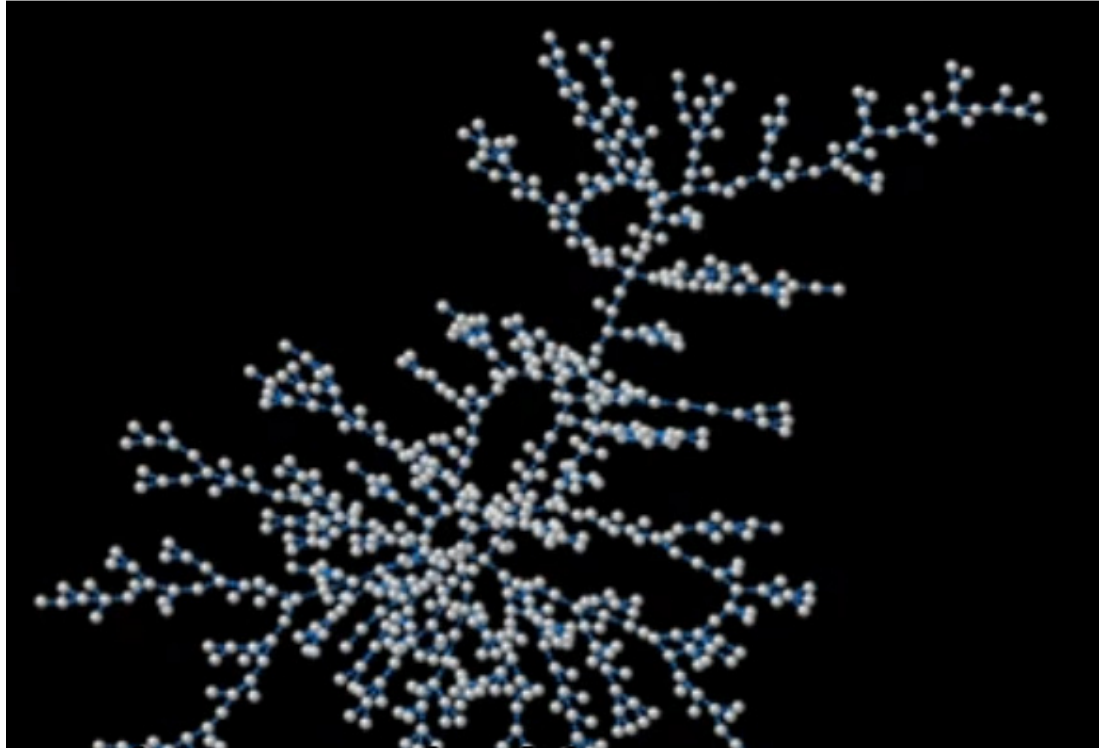


Grupul aleatoriu este reprezentat de linia roșie. Epidemia în grupul prietenilor apare la stânga. Diferența între cele două este de 16 zile. Monitorizând grupul prietenilor, putem fi preveniți cu 16 zile înainte despre o epidemie iminentă în populație.

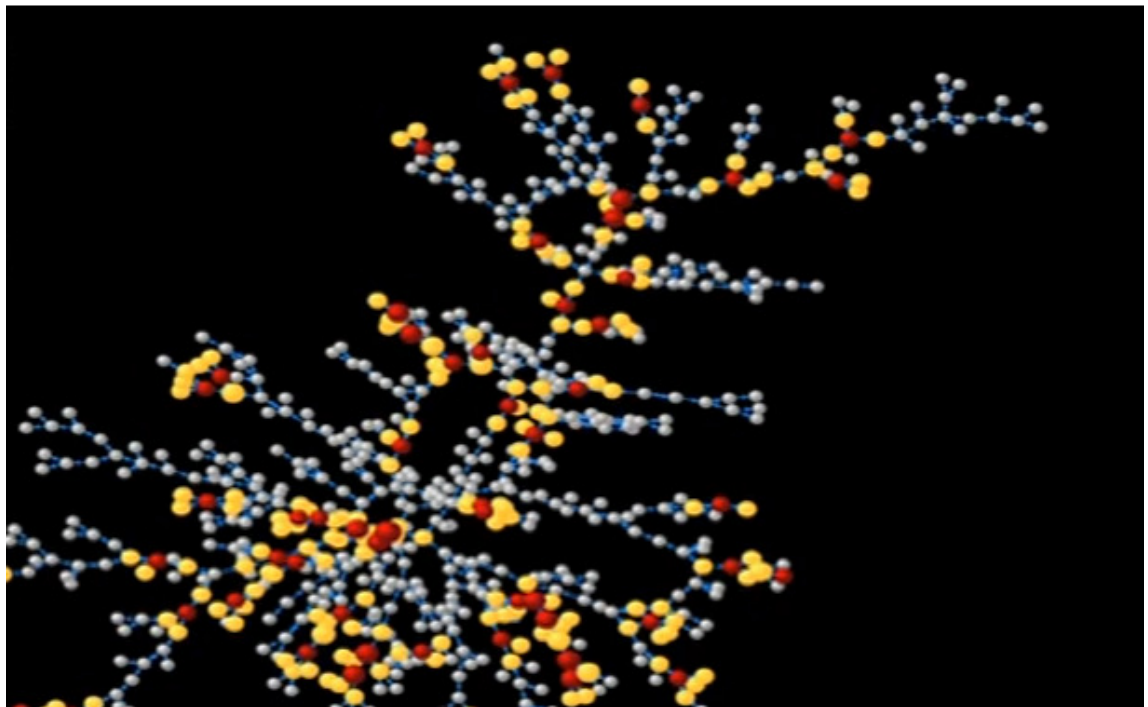
În plus, în calitate de analisti care încearcă să studieze o epidemie sau să prezică adoptarea unui produs, de exemplu, ce am putea face este să alegem un eșantion aleatoriu din populație, și, în același mod, să îi rugăm să-și nominalizeze prietenii, iar noi să-i urmărești, atât pe cei din eșantion, cât și pe prietenii lor.

Între prieteni, primul semn al unei variații peste 0 în adoptarea unei inovații, de exemplu, dovedește iminența unei epidemii. Sau am putea observa prima divergență între cele două curbe, după cum apare în stânga. Când au început cei aleși aleatoriu, când au început prietenii să se despartă de cei aleși aleatoriu, și când a început curba să se desprindă? Și aceasta, așa cum indică linia albă, s-a întâmplat cu 46 de zile înainte de punctul maxim al epidemiei. Deci aceasta ar fi o tehnică prin care am putea obține o avertizare cu o lună și jumătate înainte despre o epidemie într-o anumită populație.

Avansul cu care se da un astfel de avertisment depinde de o multitudine de factori. Ar putea depinde de natura patogenului, patogeni diferiți, folosind această tehnică, permit avertismente diferite sau alte fenomene care se răspândesc, sau, pur și simplu, de structura rețelei umane. Acum, în cazul nostru, deși nu a fost necesar, a putut cartografia rețeaua de studenți. Așa, a obținut o hartă cu 714 studenți și legăturile de prietenie dintre ei.

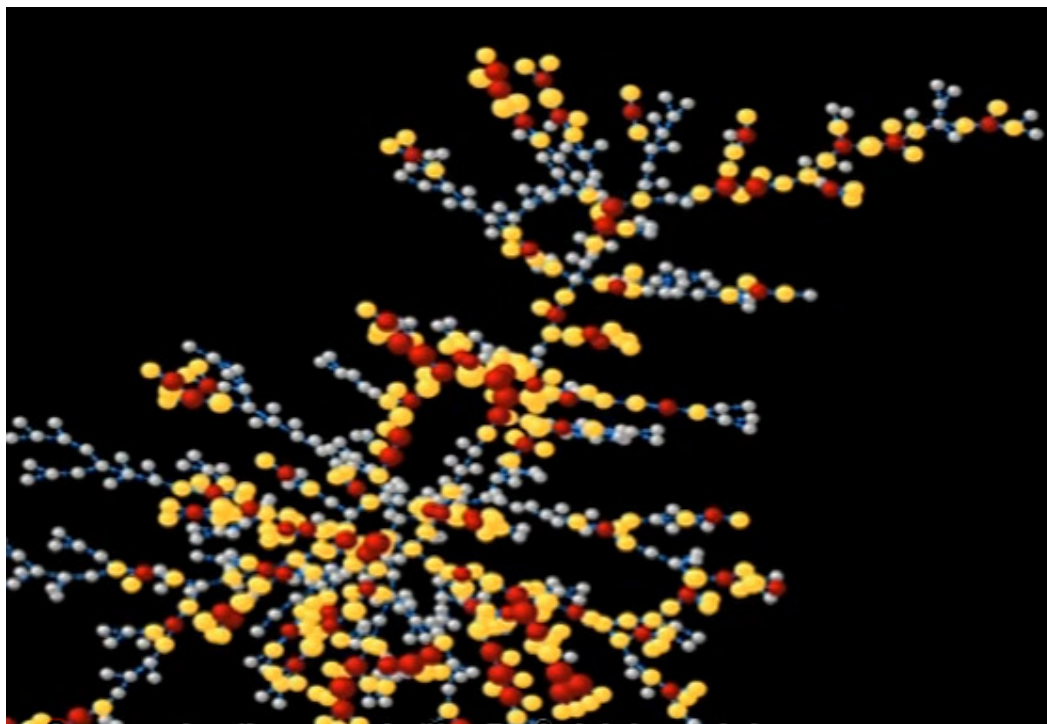


Vom avea în vedere secțiuni zilnice prin rețea timp de 120 zile. Punctele roșii sunt cazurile de gripă, iar punctele galbene sunt prietenii persoanelor cu gripă. Și mărimea punctului este proporțională cu numărul prietenilor care au gripă. Deci, cu cât punctul este mai mare, cu atât mai mulți prieteni au gripă.



Și dacă ne uităm la această imagine, aici suntem în data de 13 septembrie vedem că apar câteva cazuri și o înflorire a gripei în mijloc.

Aici suntem în data de 19 octombrie. Panta curbei epidemiei se apropie de acest moment, în noiembrie apoi un fel de nivelare, din ce în ce mai puține cazuri către finele lui decembrie.



Și acest tip de vizualizare poate indica faptul că epidemii de genul acesta își au originea și afectează mai întâi indivizi centrali, înainte de a-i afecta pe alții.

Această metodă nu este aplicabilă doar microbilor, ci, de fapt, oricărui lucru care se poate răspândi într-o populație. Informația se răspândește în populație. Normele se pot răspândi în populație. Comportamentele se pot răspândi în populație. Și, prin comportament, înțelegem lucruri precum comportamentul infracțional sau comportamentul de vot sau comportamentul legat de sănătate, precum fumatul sau vaccinarea sau adoptarea de produse, sau alte tipuri de comportamente care se leagă de influența interpersonală.

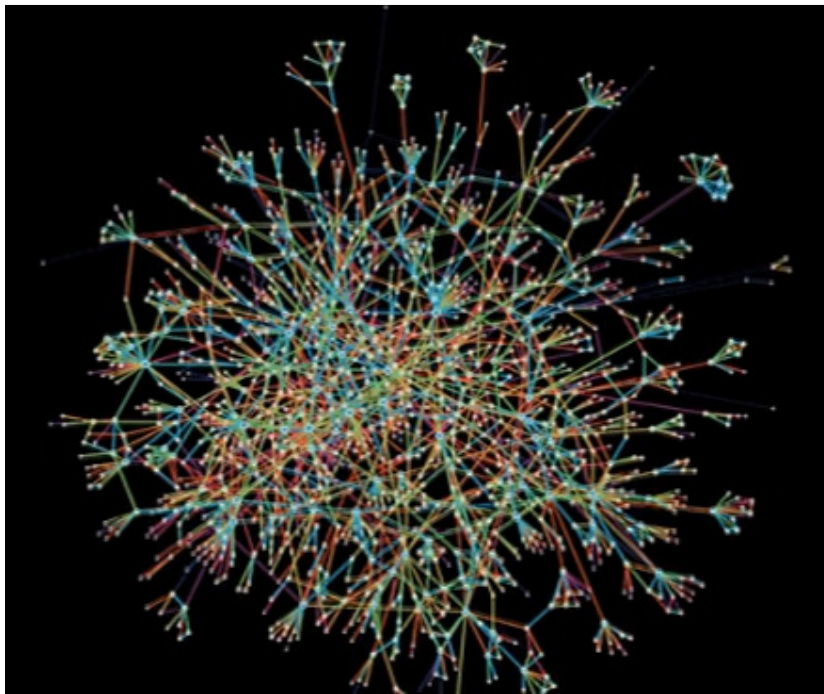
Dacă suntem înclinați să facem ceva care îi afectează pe cei din jur, această tehnică poate avertiza sau poate detecta din timp, adoptarea comportamentului în masa populației. Cheia este că, pentru a funcționa, trebuie să existe influență interpersonală. Nu se poate lega de vreo metodă de propagare care îi afectează pe toți uniform.

În ceea ce privește rețelele, aceleași idei pot fi exploatate, pot fi exploatate în alte moduri, de exemplu, în vizarea anumitor persoane pentru intervenții. De exemplu, cei mai mulți dintre voi suntem familiarizați cu noțiunea de "imunitate de turmă". Deci, dacă avem o populație de o mie de persoane și vrem să imunizăm populația la un anumit patogen, nu e nevoie să imunizăm toate persoanele.

Dacă imunizăm 960 dintre ei, e ca și cum i-am fi imunizat în proporție de 100%. Pentru că, deși dacă unul sau doi dintre cei ne-imunizați e infectat, nu au cum să-i infecteze pe alții. Ei sunt înconjurați de persoane imune. Așa că un procentaj de 96 este la fel de bun ca unul de 100.

Unii oameni de știință au estimat ce s-ar întâmpla dacă am lua un eșantion aleatoriu de 30 de procente din acești 1.000 de oameni, 300, și i-am imuniza. Am obține vreo imunitate la nivelul populației? Răspunsul este nu. Dar, dacă am lua acest procentaj, acești 300 de oameni, și i-am pune să își nominalizeze prietenii și am folosi același număr de doze de vaccin și am vaccina prietenii celor 300, cei 300 de prieteni, pot obține același nivel de imunitate de turmă ca atunci când am vaccina un procentaj de 96 din populație cu o eficiență mult mai mare, cu o constrângere bugetară strictă.

Idei similare pot fi folosite, de exemplu, pentru țintirea distribuției unor lucruri precum plasele împotriva fânțarilor în statele în curs de dezvoltare. Dacă am putea înțelege structura rețelelor în sate, am putea ținti beneficiarii unor astfel de intervenții pentru a obține răspândirea dorită. Sau, pentru publicitatea diferitelor tipuri de produse. Dacă am putea înțelege cum să țintim, asta ar putea afecta eficiența a ceea ce încercăm să obținem. Și, de fapt, putem folosi date dintr-o multitudine de surse în ziua de azi.



Aceasta este o hartă cu opt milioane de utilizatori de telefonie într-o țară europeană. Fiecare punct este o persoană, iar fiecare linie reprezintă volumul convorbirilor telefonice dintre persoane. Putem folosi astfel de date, obținute pasiv, să cartografiem aceste țări în întregime și să înțelegem cine este așezat unde în cadrul rețelei. Fără a fi nevoie să îi contactăm, putem obține astfel de perspective structurale. Și alte surse de informație sunt disponibile pentru caracteristici precum interacțiuni pe e-mail, interacțiuni online, rețele sociale online, și așa mai departe.

Sunt multe moduri în care putem folosi date obținute astfel pentru a crea rețele de senzori pentru a urmări populația, a înțelege ce se întâmplă în populație și a interveni la nivelul populației pentru a face bine. Pentru că aceste noi tehnologii ne spun nu doar cine vorbește cu cine, dar și unde se află, și ce gândește fiecare, bazându-se pe conținutul încărcat de aceștia pe Internet, și ce cumpără fiecare, bazându-se pe achizițiile lor. Toate aceste date administrative pot fi agregate și procesate pentru a înțelege comportamentul uman într-un mod imposibil până acum.

Deci, de exemplu, putem folosi achizițiile de combustibil ale conducătorilor de camioane. Deci camionagii își văd de treabă și cumpără combustibil. Iar noi vedem o creștere a achizițiilor lor de combustibil și știm că recesiunea e aproape de sfârșit.

Sau putem monitoriza viteza cu care oamenii se deplasează pe autostradă folosind telefoanele și compania de telefonie poate vedea, atunci când viteza e în descreștere, că există un ambuteiaj. Iar ei pot trimite informația înapoi la abonați, dar doar abonaților aflați pe acea autostradă aflați în spatele ambuteiajului! Sau putem monitoriza pasiv comportamentul medicilor de a prescrie medicamente și putem vedea cum arată răspândirea inovațiilor din domeniul farmaceutic în interiorul rețelei de medici. Sau, din nou, putem monitoriza comportamentul de achiziție al oamenilor, și putem observa cum acest gen de fenomene se pot propaga în populație.

Sunt trei moduri în care aceste date masiv-pasive pot fi folosite. Unul dintre ele este în totalitate pasiv, așa cum este de exemplu situația camionagiilor, unde nu intervenim de fapt la nivelul populației în niciun fel.

Unul este cvasi-activ, ca în exemplul gripei pe care l-am dat, unde putem ruga câteva persoane să își nominalizeze prietenii și apoi, pasiv, să îi monitorizam pe aceștia: au sau nu au gripă? și astfel să fim avertizați.

Un alt exemplu ar fi, dacă suntem o companie telefonică, ne dam seama cine se află în centrul rețelei și îl întrebăm "Uite, vrei să ne trimiți zilnic un mesaj în care să ne spui dacă ai sau nu febră? Doar trimite-ne un mesaj cu temperatura ta." Și colectăm cantități mari de informație despre temperatura oamenilor, dar de la indivizi localizați central. Și putem, pe scară largă, să monitorizam o epidemie iminentă cu o contribuție minimă din partea oamenilor.

Sau, în cele din urmă, poate fi vorba despre ceva activ, unde oamenii pot participa global la platforme de tip "wiki" sau prin fotografii, sau monitorizând alegeri și încărcând informația într-un mod care să ne permită să o agregăm pentru a înțelege procese sociale și fenomene sociale.

De fapt, disponibilitatea unor astfel de date prevestește o nouă epocă a ceea ce Christakis, dar și alții, numeste "știință socială computațională". E ca atunci când Galileo a inventat -- sau, nu a inventat -- a ajuns să folosească telescopul și a putut privi cerul într-un nou mod sau Leeuwenhoek a aflat despre microscop sau, de fapt, l-a inventat și a putut vedea biologia într-un nou mod.

Dar acum avem acces la aceste tipuri de date care ne permit să înțelegem procesele sociale și fenomenele sociale într-un mod cu totul nou, imposibil până acum. Și cu acest tip de știință, putem înțelege exact cum întregul ajunge să fie mai mult decât suma părților sale. Și, de fapt, putem folosi aceste informații pentru a îmbunătăți societatea și a crește bunăstarea omului.

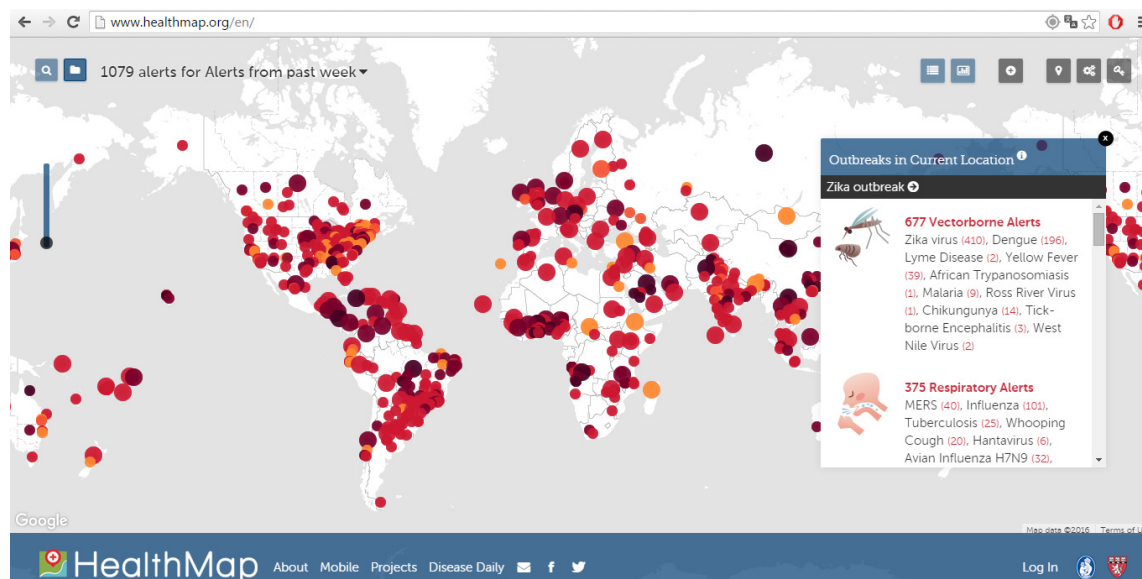
Un alt studiu, de aceasta data, al unui grup de cercetători și specialiști în informatică din Boston. Cu nouă zile înainte ca propagarea Ebola să fie declarată oficial o epidemie, acești au remarcat deja răspândirea virusului febrei hemoragice în Guinea. Monitorizând rețelele de socializare, buletinele de știri locale și alte baze de date, algoritmul dezvoltat de compania HealthMap a detectat foarte repede această propagare a virusului în Africa de Vest.

Compania HealthMap analizează date în 15 limbi și a monitorizat anterior virusul gripei H1N1. Compania poate să supravegheze astfel evoluțiile avute și de alte maladii.

HealthMap este un instrument inovator care utilizează algoritmi care centralizează zeci de mii de informații de pe site-uri de media, știri locale, site-uri guvernamentale, rețele sociale ale biologilor specializați în boli infecțioase, clasifică informația relevantă, identifică virusurile și locațiile și îi înștiințează pe experți.

Cum funcționează HealthMap: localizează o epidemie pe harta mondială și creează un sistem de codare cu ajutorul culorilor care indică severitatea epidemiei, bazându-se pe cele mai recente știri relaționate cu acest subiect. Utilizatorii site-ului HealthMap pot vizualiza și analiza informațiile disponibile despre virusuri.

Site-ul healthmap.org, interfața sistemului, centralizează informații indexate de Google News, comunicate al Organizației Mondiale a Sănătății (OMS), rapoarte ale altor institute medicale de cercetare, informații de pe rețele de socializare, precum și din multe alte surse.



Când virusul Ebola a făcut primele victime în Africa de Vest, declarațiile martorilor, rețelele de socializare și alte informații de la spitalele din regiune și de la angajații lor au fost primele informații disponibile. Mai târziu, rapoartele au arătat că cel mai probabil primul caz de infecție a fost în Guinea, pe 9 februarie, însă la acea vreme nu se știa identitatea virusului.

Pe 14 martie, HealthMap a centralizat rapoarte despre o “misterioasă febră hemoragică” care ucisese deja 8 oameni în Guinea. Pe 19 martie, algoritmul a detectat cel mai probabil prima știre locală despre o posibilă epidemie de Ebola, sistemul alertând utilizatorii. Nouă zile mai târziu, în data de 23 martie, Organizația Mondială a Sănătății a făcut prima declarație despre epidemia de Ebola din Africa de Vest.

Cum a reușit un program informatic să depisteze epidemia de Ebola înaintea Organizației Mondiale a Sănătății? Câțiva dintre primii doctori care au tratat pacienți infectați cu Ebola în Guinea au împărtășit experiența lor pe blogurile personale. Pe măsură ce doctorii scriau despre tratamentele oferite pacienților cu simptome ale virusului Ebola, câțiva oameni au preluat articolele pe rețele de socializare. Informațiile au ajuns în scurt timp și la HealthMap.

Bilanțul victimelor febrei hemoragice Ebola a depășit pragul de 1.000 de morți, provocând decesul a 1.013 persoane și 1.848 de cazuri înregistrate, potrivit OMS.

Asemenea acestui studiu pe site-ul lor mai sunt și alte exemple, actualizate zilnic.