Connection Science and Engineering

Interdisciplinary Workshop on Information and Decision in Social Networks

November 8 - 9, 2012
MIT

Editors:
Vincent Blondel, Costis Daskalakis, David Gamarnik, Asu Ozdaglar, Alex 'Sandy' Pentland, Devavrat Shah, John Tsitsiklis

Sponsored by:

Los Alamos
NATIONAL LABORATORY
EST. 1943

→MISTI
MIT International Science &
Technology Initiatives

Interdisciplinary Workshop on Information and Decision in Social Networks
November 8 & 9, 2012

This workshop is being organized by the **Connection Science and Engineering Center at MIT.**

Organizers:
Vincent Blondel, UCLouvain (Belgium) and LIDS, MIT
Costis Daskalakis, CSAIL, MIT
David Gamarnik, Sloan School, MIT
Asu Ozdaglar, LIDS, MIT
Alex 'Sandy' Pentland, Media Lab, MIT
Devavrat Shah, LIDS, MIT
John Tsitsiklis, LIDS, MIT

Support Staff:
For administrative requests and general questions about the conference, please contact Lynne or Nicole.
Lynne Dell – Administrative Support – 617-452-3679
Nicole Freedman – Administrative Support – 617-253-3818

For problems with the website, or technical questions regarding the conference, please contact Brian.
Brian Jones – Technical Support – 617-253-4070

Location:
The workshop will take place on the 6th floor of the MIT Media Lab (building E14).
75 Amherst Street, E14 room 674, Cambridge, MA

Travel and accommodation:
Workshop participants travelling to Cambridge should arrange their travel and accommodation; no support will be provided by the workshop organizers. There are several hotels close to MIT campus, including the historic Kendall Hotel and the Marriott Cambridge (for both hotels, please ask for the MIT rate).

Plenary Speakers:

- Prof. Vincent Blondel, Dept. of Applied Mathematics, Université catholique de Louvain LIDS, MIT
- Prof. Michael Kearns, Dept. of Computer and Information Science and Market and Social Systems Engineering, University of Pennsylvania
- Prof. Alvin Roth, Dept. of Economics, Harvard University and Stanford University

Panel:
The program will also feature a panel entitled "What problems does network research need to solve?" with leading thinkers from the communication, mobile, banking, and information industries. The participants are:

- Steve Whittaker, BT
- Ken Gabriel, Motorola Mobility
- David Zafrilla-Gonzales, BBVA
- Martin Wattenberg, Google

The panel will be moderated by Professor Alex Pentland.


E14
E14 E14
E60
E56
E53
E52
E55
E51
E48
E42
E40
E39
E38
E34
E33
E32
E23
E25
E28
E2
E1
E15
E19
E18
E17
E70
NE18
NE20
NE25
NE43
NE46
NE47
NE48
NE49
NE80
NE83
51
50
64
62
68
66
54
18
14N
14E
14W
14S
56
57
32
16
8
6
6A
6B
4
2
1
3
5
7A
9
10
11
12
13
17
24
26
31
33
34
35
36
37
38
39
41
42
43
44
46
48
N4
N9
N10
N16
N16A
N42
N51
N52
NN57
NW10
NW12
NW13
NW14
NW15
NW16
NW17
NW21
NW22
NW30
NW61
NW62
NW86
NW95
W1
W2
W4
W5
W7
W8
W11
W13
W15
W16
W20
W31
W32
W33
W34
W35
W45
W51
W53
W53A
W59
W61
W70
W71
W79
W84
W85
W85FG
W85DE
W85HJK
W89
W91
W92
MEMORIAL DR
MASSACHUSETTS AVE
HARVARD BRIDGE
Harvard Bridge
MAIN ST
BROADWAY
AMES ST
CARLETON ST
HAYWARD ST
WADSWORTH ST
AMHERST ST
AMHERST ALLEY
VASSAR ST
ALBANY ST
PORTLAND ST
WINDSOR ST
OSBORN ST
SMART ST
FRONT ST
STATE ST
CROSS ST
LANDSDOWNE ST
PURRINGTON ST
FRANKLIN ST
SIDNEY ST
PACIFIC ST
WAVERLY ST
TUDOR ST
EMILY ST
MERRIAM ST
ERIE ST
TALBOT ST
AUDREY ST
AMESBURY ST
DANFORTH
ENDICOTT ST
FOWLER
ANGLIM
BROOKLINE ST
PILGRIM ST
BROOKLINE PL
DE DECATUR ST
LOPEZ ST
WATSON ST
SALEM ST
AUBURN ST
WILLIAM ST
MAGAZINE ST
RIVER ST
PEARL ST
GREEN ST
TEMPLE ST
VAIL CT
AUSTIN PARK
MASSACHUSETTS
BISHOP ALLEN DR
ASHBURTON PL
ESSEX ST
LAMSON PL
NORFOLK ST
DOUGLASS ST
COOLIDGE PL
WORCESTER
SUFFOLK ST
COLUMBIA ST
MARVIN PL
VOLTON CT
PINE ST
WASHINGTON ST
CHERRY ST
SCHOOL ST
HAROLD PL
BOARDMAN ST
MARKET ST
ROSSLAND ST
CLARK ST
HARVARD ST
MOORE ST
DAVIS ST
HAMPSHIRE ST
CARDINAL MEDEIROS AVE
TECHNOLOGY SQ
2ND ST
BINNEY ST
WESTERN CONNECTOR
EASTERN CONNECTOR
FULKERSON ST
ROGERS ST
BENT ST
3RD ST
5TH ST
6TH ST
MONROE ST
POTTER ST
LINSKY WAY
Kendall Square
Marriott Hotel
Marriott/Residence Inn
Whitehead Institute
University Park Hotel @ M.I.T.
77 Massachusetts Ave.
84 Massachusetts Ave.
Hyatt Regency
Hyatt

Interdisciplinary Workshop on Information and Decision in Social Networks

November 8 & 9, 2012

Scientific Committee

Daron Acemoglu,Economics, MIT
Itai Ashlagi. Sloan School, MIT
Albert-Laszlo Barabasi, Northeastern University
Kostas Bimpikis, Stanford University
Larry Blume, Cornell University
Damon Centola, Sloan School, MIT
Munther Dahleh, EECS, MIT
Vivek Farias, Sloan School, MIT
Ben Golub, Harvard University and MIT
Marta Gonzalez, Civil and Environmental Engineering, MIT
Sanjeev Goyal, University of Cambridge (UK)
Julien Hendrickx, EPLouvain and UCLouvain (Belgium)
Cesar Hidalgo, Media Lab, MIT and Harvard University
Ali Jadbabaie, University of Pennsylvania
Patrick Jaillet, LIDS, MIT
Sep Kamvar. Media Lab, MIT
Gary King, Harvard University
David Lazer, Harvard University
Jure Leskovec, Stanford University
Andrew W. Lo, Sloan School, MIT
Mihai Manea, Economics, MIT
Andrew McAfee, Economics, MIT
Christos Papadimitriou, EECS, MIT
Pablo Parrilo, EECS, MIT
R. Ravi, Carnegie Mellon University
Whitman Richards, CSAIL, MIT
Alessandro Vespignani, Northeastern University
Martin Wainwright, University of California at Berkeley
John Williams, Civil and Environmental Engineering, MIT

Interdisciplinary Workshop on Information and Decision in Social Networks
November 8 - 9, 2012
MIT

Thursday, November 8, 2012	
8:30	Registration
8:45-9:00	Opening Remarks
9:00-10:00	Industry Panel – “What problems does network research need to solve?”
10:00-10:30	Coffee Break
10:30-11:30	Brief Talks – Session 1: “Structure and Communities in Networks”
11:30-11:45	Break
11:45 – 12:45	Brief Talks – Session 2: “Dynamics and Networks”
12:45 – 14:00	Lunch - maps to local restaurants available at the registration desk
14:00-14:30	Invited Plenary Presentation: Professor Vincent Blondel, Université catholique de Louvain and MIT "Social communities and mobile phone communication networks”
14:30-15:30	Poster Session – 3rd Floor Atrium
15:30-16:00	Coffee Break
16:00-17:30	Brief Talks – Session 3: “Information Propagation, Influence and Control”
Friday, November, 9, 2012	
8:30	Registration
9:00-9:30	Invited Plenary Presentation: Professor Alvin Roth, Harvard University and Stanford University “Compatibility Networks in Kidney Exchange”
9:30-9:45	Break
9:45-10:45	Brief Talks – Session 4: “Information Propagation, Influence and Control”
10:45-11:15	Coffee Break
11:15-12:30	Brief Talks – Session 5: “Social Data Collection and Analysis”
12:30-14:00	Lunch - maps to local restaurants available at the registration desk
14:00-14:30	Invited Plenary Presentation: Professor Michael Kearns, University of Pennsylvania “Competitive Contagion, Anarchy, and Budgets”
14:30-15:30	Poster Session - 3rd Floor Atrium
15:30-16:00	Coffee Break
16:00-17:00	Brief Talks –Session 6: “Structure and Communities in Networks”
17:00-17:15	Break
17:15-18:15	Brief Talks –Session 7: “Opinion Dynamics and Learning”

Program Outline
Thursday
November 8

8:30 Registration

8:45-9:00 Opening Remarks

9:00-10:00 Industry Panel
What problems does network research need to solve?

The world is moving to distributed, network systems for health, finance, work, government, and community. But there are big questions about how well these systems really work...do they make the right decisions...and we know little about their stability, optimality or how they will change society. We have four leading thinkers from the communication, mobile, banking, and information industries who will help us understand where the research community needs to go.

Participants:
Steve Whittaker, BT
Ken Gabriel, Motorola Mobility
David Zafrilla-Gonzales, BBVA
Martin Wattenberg, Google
Moderator:
Alex Pentland

10:00-10:30 Coffee Break

10:30-11:30 Brief Talks Session 1
Structure and Communities in Networks

Network structure and the aggregation of information: theory and evidence from Indonesia
V. Alatas, *World Bank*
A. Banerjee, *MIT*
A.G. Chandrasekhar, *Microsoft Research*
R. Hanna, *Harvard University*
B.A. Olken, *MIT*

Processing Power Limits Social Group Size: Computational Evidence for the Social
Brain Hypothesis
T. Dávid-Barrett, *University of London*
R.I.M. Dunbar, *University of Oxford*

Scaling Theory of Human Mobility and Spatial Network
P. Deville, *Université Catholique de Louvain*
D. Wang, *Northeastern University, Dana Farber Cancer Institute*
C. Song, *Northeastern University, Dana Farber Cancer Institute*
N. Eagle, *Northeastern University*
V. Blondel, *Université Catholique de Louvain*
A. László Barabási, *Northeastern University, Dana Farber Cancer Institute, Harvard Medical School*

Local computation of network centrality
C. Lee, *MIT*
A. Ozdaglar, *MIT*
D. Shah, *MIT*

11:30-11:45: Coffee Break

11:45-12:45: Brief Talks Session 2
Dynamics and Networks

Social Learning and Network Uncertainty
I. Lobel, *New York University*
E. Sadler, *New York University*

Intermediation and Exclusive Representation in Financial Networks
I. Fainmesser, *Brown Univ.*

Competitive Marketing Strategies over Social Networks
K. Bimpikis, *Stanford University*
A. Ozdaglar, *MIT*
E. Yildiz, *Accenture Technology Labs*

Nine Facts, Three Parameters and One Theory: A Precise Analytical Model for the Dynamics of Industry Location Networks
R. Charan, *MIT*
C. Hidalgo, *MIT*

12:45 – 14:00: Lunch

14:00-14:30 Invited Plenary Presentation
Professor Vincent Blondel

"Social communities and mobile phone communication networks"

Abstract:
We describe several recent results on large network analysis with a special emphasis on community detection and on the analysis of mobile phone datasets. In particular, we describe two simple and efficient methods - the "Louvain method" and the recent Partition-Merge method - for the detection of communities. Partition-Merge is a general and versatile method that allows to turn any existing centralized algorithm for graph computation distributed while keeping accuracy guarantees. The Louvain method - now used by LinkedIn for its visualization application InMaps - has sub-quadratic computational complexity and can be routinely used for analyzing networks with billions of nodes or links.

We analyze communities obtained on a nationwide dataset of criminal records, as well as on a social network constructed from mobile phone communications in Belgium and in France on periods covering several months. In the later we observe a gravitational law of social interaction as well as spatially distributed social structures that have potential political implication.
We finally describe applications of mobile phone dataset analysis for a range of applications such as urban planning, traffic optimization, monitoring of development policy, crisis management, and control of epidemics. With these applications in mind we overview the ongoing "Data for Development" (D4D) challenge (http://www.d4d.orange.com/) organized jointly with Orange on the analysis of mobile phone datasets from an African country and for development purposes. So far the D4D challenge has attracted projects by more than 200 research groups.

Program Outline
Thursday
November 8

14:00-14:30 Invited Plenary Presentation Cont.

Bio:

Vincent D. Blondel is professor of applied mathematics at the University of Louvain (Belgium) were he has been department head for close to a decade. He is also affiliated with the Massachusetts Institute of Technology (Cambridge, USA) where he was a visiting professor and Fulbright scholar in 2004-2005 and again in 2010-2011. Vincent has held various positions with the University of Oxford, the Australian National University, the Santa Fe Institute. In 2012-2013 he is the Kokotovic Distinguished Visiting Professor at the University of California (Santa Barbara). He has directed more than thirty PhD and Master thesis and is the recipient of several international prizes for research done in mathematical control theory, theoretical computer science and network science. Vincent is the coordinator a national research network between 200+ researchers in systems and control in Belgium, Princeton and Stanford and he is the Belgian coordinator of the mammoth one billion dollars EU research Flagship proposal FuturICT (under review!).

14:30-15:30 Poster Session 1

IRIE: Scalable and Robust Influence Maximization in Social Networks
K.Jung, W. Heo, W. Chen

Exploring the effect of the duration and period of Face-to-Face interactions on Close Friendships
R. Oloritun, A. Pentland, A. Madan, I. Khayal

Change in BMI Accurately Predicted via Social Exposure
R. Oloritun, T. Ouarda, S. Moturu, A. Madan, A. Pentland, I. Khayal

The Rewarding Nature of Matchmaking
L. Anik, M. Norton

Graphlet decomposition of a weighted network
H.Soufiani, E. Airoldi

Optimal Induction of a new contact in a social network
V. Borkar

Why are your Facebook “friends” not optimal?
G. Ducoffe, D. Mazauric, A. Chaintreau

Modeling balance in social networks
P.De Leenheer, V. Traag, P. Van Dooren

Understanding User Triads on Facebook
D. Doran, A. De la Rosa Algarin, S. Gokhale

The value of feedback on containment of epidemics on social networks
K. Drakopoulos, A. Ozdaglar, J. Tsitsiklis

Sociodynamic Discrete Choice: Equilibrium Behavior of the Nested Logit Model with Social Interactions
E. Dugundji
The effect of leader centrality on team performance
K. Ehrlich, H. Tong
Game Theoretic Formation of a Centrality Based Network
R. Tatko, C. Griffin

14:30-15:30 Poster Session 1 Cont.

Causal Strategic Inference in Networks
M. Irfan, L. Ortiz

Analyzing Scrip Systems: Selection Rules and Optimality
K. Johnson, D. Simchi-Levi, P. Sun

Strategic Disciplining Behavior in Socialized KingPawn
J. Judd

An internet experiment on bargaining
C. Chang, Y. Kanoria, A. Montanari, M. White

Disrupting the Spread of Information
A. Khanafer, B.Touri, T.Başar

More Contentious Issues Lead to More Factions: Bounded Confidence Opinion Dynamics of Bayesian Decision Makers
K. Varshney

15:30-16:00 Coffee Break

16:00-17:30 Brief Talks Session 3

Information Propagation, Influence and Control

Trends Prediction in Social Networks Using Influence Diffusion Models
Y. Altshuler, *MIT*
W. Pan, *MIT*
A. Pentland, *MIT*

The Fallacy of Pigeonholing Global Users into Local Models
C. Budak, *UCSB*
D. Agrawal, *UCSB*
A. El Abbadi, *UCSB*

Apple vs. Android: a comparison of social effects in adoption
J. Bjelland, *Telenor ASA*
G. Canright, *Telenor ASA*
K. Engø-Monsen, *Telenor ASA*
P. Roe Sundsøy, *Telenor ASA*
R.S. Ling, *IT University / Telenor ASA*

Connectivity & Collective Action in Social Networks
N. Hassanpour, *Yale University*
J. Liu, *Yale University*
S. Tatikonda, *Yale University*

An Analysis of One-Dimensional Schelling Segregation
C. Brandt, *Stanford University*
N. Immorlica, *Northwestern University*
G.Kamath, *MIT*
R. Kleinberg, *Cornell University*

On Bitcoin and Red Balloons
M. Babaioff, *Microsoft Research*
S. Dobzinski, *Cornell University*
S. Oren, *Cornell University*
A. Zohar, *Microsoft Research*

Program Outline
Friday
November 9

8:30 Registration

9:00-9:30 Invited Plenary Presentation:

Professor Alvin Roth
"Compatibility Networks in Kidney Exchange"

Bio:
Al Roth received his B.S. in OR from Columbia University in 1971, and his Ph.D. in OR from Stanford in 1974. He has taught at the U of I, Pitt, Harvard, and now Stanford. Several weeks ago he was awakened from a sound sleep to learn that he had won the 2012 Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel.

9:30 – 9:45 Break

9:45-10:45 Brief Talks Session 4

Information Propagation, Influence and Control

On Threshold Models over Finite Networks
E.M. Adam, *MIT*
M.A. Dahleh, *MIT*
A. Ozdaglar, *MIT*

A Nonparametric Method for Early Detection of Trending Topics
S. Nikolov, *MIT, Twitter, Inc.*
D. Shah, *MIT*

Active Influence in Dynamical Models of Structural Balance in Social Networks
T.H. Summers, ETH Zürich
I. Shames, University of Melbourne

SODEXO: A System Framework for Deployment and Exploitation of Deceptive Honeybots in Social Networks
Q. Zhu, *University of Illinois at Urbana-Champaign*
A. Clark, *University of Washington, Seattle*
R. Poovendran, *University of Washington, Seattle*
T. Başar, *University of Illinois at Urbana-Champaign*

10:45-11:15 Coffee Break

11:15-12:30: Brief Talks Session 5

Social Data Collection and Analysis

Game in the Newsroom: Greedy Bloggers for Picky Audience
A. May, *Columbia University*
A. Chaintreau, *Columbia University*
N. Korula, *Google*
S. Lattanzi, *Google*

A Comparative Study of Geographic Routing in Social Network Based on Mobile Phone Data
C. Herrera, *MIT*
T. Couronne, *Orange Labs*
Z. Smoreda, *Orange Labs*
C. M. Schneider, *MIT*
R. M. Benito, *MIT*
M. C. González, *MIT*

11:15-12:30: Brief Talks Session 5 Cont.

Homophily does not inflate contagion estimates in a social dilemma laboratory experiment
J.J. Jordan, *Harvard University*
D. G. Rand, *Harvard University*
S. Arbesman, *Harvard University & Ewing Marion Kauffman Foundation*
J. H. Fowler, *Harvard University & UCSD*
N. A. Christakis, *Harvard University & Harvard Medical School*

Topic-Specific Communication Patterns in Email Data
P. Krafft, *University of Massachusetts Amherst*
J. Moore, *University of Massachusetts Amherst*
H. Wallach, *University of Massachusetts Amherst*
B. Desmarais, *University of Massachusetts Amherst*
J. ben-Aaron, *University of Massachusetts Amherst*

Directed Acyclic Motifs for Conversation Analytics
L. R. Varshney, *IBM Thomas J. Watson Research Center*

12:30-14:00 Lunch

14:00-14:30 Invited Plenary Presentation:

Professor Michael Kearns

"Competitive Contagion, Anarchy, and Budgets"

Abstract:
We examine a game-theoretic model of "competitive contagion" in networks, where two competing companies or other entities have limited budgets to seed initial infections in an underlying social network, which then mediates stochastic contagion. The payoff to each party is their final number of infections, which at equilibrium may come at the expense of the other party. In this model, we provide characterizations of the Price of Anarchy and a new notion called the Budget Multiplier. These characterizations depend on properties of the local stochastic dynamics of contagion and network structure, and in some cases exhibit sharp threshold behavior.

Joint research with Sanjeev Goyal of the University of Cambridge.

Bio:
Michael Kearns is the National Center professor of Computer and Information Science at the University of Pennsylvania, where he is the director of the new Penn program in Market and Social Systems Engineering. His research interests include topics in machine learning, algorithmic game theory, social networks, computational finance and artificial intelligence.

Program Outline
Friday
November 9

14:30-15:30 Poster Session 2

Phase Transitions for Information Spreading in Random Clustered Networks with Arbitrary Contact Rates
S. Lim, N. Kwak, K. Jung

Using Social Influence to Predict Subscriber Churn
D. Doran, V. Mendiratta, C. Phadke, D. Kushnir, H. Uzunalioglu

Eulerian Opinion Dynamics with Bounded Confidence and Exogenous Inputs
A. Mirtabatabaei, P. Jia, F. Bullo

State-Dependent Opinion Dynamics
D. Acemoglu, M. Mostagir, A. Ozdaglar

Sparsifying Defaults: Optimal bailout policies for financial networks in distress
Z. Li, I. Pollak

Social Mobilization in Context
A. Rutherford, M. Cebrian, S. d'Souza, E. Moro, A. Pentland, I. Rahwan

Baiting Strategies for Viral Marketing
L. Seeman, Y. Singer

Search Friction and the Stationarity of Networked Markets
R. Berry , T. Nguyen, V. Subramanian

Social Behavior and Human Mobility
J. Toole, C. Herrera-Yagüe, C. Schneider, M. González

On Language Formation in a Noisy Environment
B.Touri, C. Langbort

Focal Vocabularies vs. Shared Vocabularies in Social Networks: Balancing Individual Concerns and Social Exchange
A. Mani, L. Varshney, A. Pentland

How many people you have to know to talk to most people?
T. Wang, P. Hui

Collaborative Human Decision Making with Imperfect Information
T.Wimalajeewa, P. Varshney

A Bayesian Approach for Predicting the Popularity of Tweets
T. Zaman, E. Fox, E. Bradlow

Coordination with Local Information
M. Dahleh, A.Tahbaz-Salehi, J. Tsitsiklis, S. Zoumpoulis

Dynamic Pricing for Revenue Maximization in the Presence of Social Influences
B. T. Swapna, A. Eryilmaz, N. Shroff

Boosting the Public Health against Infectious Diseases through Information Dissemination in Social Networks
F. Darabi Sahneh, C. Scoglio

14:30-15:30 Poster Session 2 Cont.

Exploring the role of duration of interaction in ad-hoc mobile face-to-face networks
R. Oloritun, A. Pentland, A. Madan, I. Khayal

Language, Knowledge, and Power in the International System – A Linguistic Network Analysis of Published Political Science Research by Nation-State from 1991 to 2008.
C. Gomez, D. McFarland

15:30-16:00 Coffee Break

16:00-17:00 Brief Talks

Session 6: Structure and Communities in Networks

Artificial Social Scientist: an Application to Marriage Networks
T. Menezes, CAMS, CNRS/EHESS, France
C. Roth, CMB, CNRS/HU/MAEE, Germany

Urban characteristics attributable to density-driven tie formation
W. Pan, *MIT*
G. Ghoshal, *Harvard University*
C. Krumme, *MIT*
A. Pentland, *MIT*
M. Cebrian, *UCSD*

Structural Analysis of Viral Spreading Processes in Social and Communication Networks Using Egonets
V. M. Preciado, Univ. of Pennsylvania

On the structure of communities in networks
B. Abrahao, *Cornell University*
S. Soundarajan, *Cornell University*
J. Hopcroft, *Cornell University*
R. Kleinberg, *Cornell University*

17:00 – 17:15: Break

17:15-18:15 Brief Talks

Session 7: Opinion Dynamics and Learning

On the Convergence of the Hegselmann-Krause System
A. Bhattacharyya, *Princeton University*
M. Braverman, *Princeton University*
B. Chazelle, *Princeton University*
H. L. Nguyen, *Princeton University*

Variational Inference for Label Aggregation in Crowdsourcing
Q. Liu, *Univ. of Irvine*
J. Peng, *TTI-C, MIT*
A. Ihler, *Univ. of Irvine*

Generalized Mean-Field Approximation for Opinion Spreading in Social Networks
S. Lim, *KAIST*
K. Jung, *KAIST*

NETWORK STRUCTURE AND THE AGGREGATION OF INFORMATION: THEORY AND EVIDENCE FROM INDONESIA

VIVI ALATAS*, ABHIJIT BANERJEE†, ARUN G. CHANDRASEKHAR‡, REMA HANNA*,
AND BENJAMIN A. OLKEN♯

ABSTRACT. We use a unique data-set from Indonesia on what individuals know about the income distribution in their village to test theories such as Jackson and Rogers (2007) that link information aggregation in networks to the structure of the network. The observed patterns are consistent with a basic diffusion model: more central individuals are better informed and individuals are able to better evaluate the poverty status of those to whom they are more socially proximate. To understand what the theory predicts for cross-village patterns, we estimate a simple diffusion model using within-village variation, simulate network-level diffusion under this model for the over 600 different networks in our data, and use this simulated data to gauge what the simple diffusion model predicts for the cross-village relationship between information diffusion and network characteristics (e.g. clustering, density). The coefficients in these simulated regressions are generally consistent with relationships suggested in previous theoretical work, even though in our setting formal analytical predictions have not been derived. We then show that the qualitative predictions from the simulated model largely match the actual data in the sense that we obtain similar results both when the dependent variable is an empirical measure of the accuracy of a village's aggregate information and when it is the simulation outcome. Finally, we consider a real-world application to community based targeting, where villagers chose which households should receive an anti-poverty program, and show that networks with better diffusive properties (as predicted by our model) differentially benefit from community based targeting policies.

JEL Classification Codes: D83, D85

Keywords: Networks, Diffusion of information, Targeting, Development

---

Date: August 2012.

We thank Ritwik Sarkar, Prani Sastiono, Ririn Purnamasari, Hendratno Tuhiman, Matthew Wai-Poi, and Chaeruddin Kodir for outstanding research assistance and thank Mitra Samya, SurveyMeter the Indonesian Central Bureau of Statistics for their cooperation implementing the project. Most of all we thank Lina Marliani for her exceptional work leading the field implementation teams. Funding for this project came from a World Bank – Royal Netherlands Embassy trust fund. All views expressed are those of the authors, and do not necessarily reflect the views of the World Bank, the Royal Netherlands Embassy, Mitra Samya, SurveyMeter or the Indonesian Central Bureau of Statistics. Chandrasekhar is grateful for support from the National Science Foundation GRFP. We thank Matthew Elliott, Ben Golub, Matthew O. Jackson, Chris Udry, participants at the Calvó-Armengol Workshop and NEUDC 2010 for helpful comments. All errors are our own.

*World Bank. Email: valatas@worldbank.org.

†Massachusetts Institute of Technology, Department of Economics. Email: banerjee@mit.edu.

‡Microsoft Research New England. Email: arc@microsoft.com.

*Harvard University, Kennedy School of Government. Email: rema_hanna@ksg.harvard.edu.

♯Massachusetts Institute of Technology, Department of Economics. Email: bolken@mit.edu.

Processing Power Limits Social Group Size: Computational Evidence for the Social Brain Hypothesis

T. Dávid-Barrett[1,2] and R.I.M. Dunbar[2]

[1] Dept of Economics, Birkbeck College, University of London
[2] Dept of Experimental Psychology, University of Oxford

The social brain hypothesis[1-6] was originally proposed as an explanation for the evolution of large brains in primates, suggesting that cognitive abilities limited the size of social group that a species could maintain, and that this imposed limits on the size of human communities[7]. Although there is now considerable neuroanatomical evidence to support the social brain hypothesis in both primates[8] and humans[9-12], the claim that communication competences determined by underlying cognitive abilities might limit the size of social communities remains untested. Here we present the first mathematical formulation and computational evidence to support for this hypothesis.

We assume a group of n agents that face a coordination problem which requires behavioural synchrony. We use synchronisation on a dial to capture this problem, a simple but widely used[13,14] device. Each agent is first assigned an initial information value, a vector between 0 and 360 degrees. One of these vectors is defined as the 'true information' and is the property of just one agent, while all the other agents are assigned a randomly distributed value. The agents are arranged in a random n-node network in which each agent is linked to k others. Synchronisation takes place via a set of random dyadic meetings, in which the agents exchange their information, and then each of them calculates a weighted average of three bits of information: her original information, the partner's information, and the information the partner received in his previous meeting. These meetings are repeated until, on average, each agent has taken part in τ meetings. At this point, the average distance between the agents' individual information and the 'true information' is:

$$d(n,w,\omega)=\frac{1}{n}\sum_{i=1}^{n}\left|\phi_{T,i}-\phi_{TI}\right|$$

where w and ω are weight matrices for the weights the agent uses for the partners' and the third party information respectively, and $\phi_{T,i}$ is the information held by agent i at the end of the synchronisation episode, and ϕ_{TI} is the 'true information'.

Agents are trying to get as close to the 'true information' as possible by estimating an optimal set of weights using a memory of past information exchanges and a simple least squares optimisation function. The agents have three possible ways of estimating the optimal weights. Model F1: agents ignore both the third party information and the differences among their partners. Model F2: agents ignore third party information, but recognise the differences among their partners. Model F3: agents use both types of information. In calculating the weights, we varied the size of memory sample that the agents could use in their optimisation. We used the processor time associated with each optimisation act as an index of the cognitive demand of a strategy, and use this as a proxy for the amount of brain tissue needed in managing a strategy. This allowed us to introduce an implicit distance function:

$$\delta(F,n,\tilde{c})\big|_{\tau,k} = d(n,\hat{w},\hat{\omega})\big|_{\tau,k}$$

where $\tilde{c}$ is the measured processor time, F is the index of the method used by the agents, $\hat{w}$ and $\hat{\omega}$ are the 'optimal' weight matrices as estimated by the agents.

We assume that there is a threshold of synchronization efficiency, which has to be reached for the group to be in sufficient behavioural synchrony to be able to act as one. Then the largest group that can be in synchrony is:

$$n^*(F,\tilde{c})\big|_{\tau,k,\lambda} = \max n \text{ s.t. } \delta(F,n,\tilde{c})\big|_{\tau,k} \le \lambda$$

where λ is the synchrony threshold.

Our simulations show that the maximum group size increases as calculation capacity increases for all the three models (Fig. 1), and that for both models F=2 and 3 there is a $c^*(F)$ such that $n^*(F-1,c) > n^*(F,c)$ for $c < c^*$ and $n^*(F-1,c) < n^*(F,c)$ for $c > c^*$.

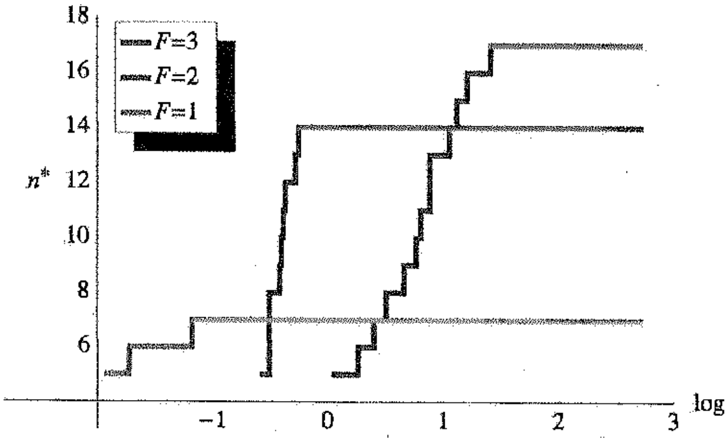


Fig. 1. Limiting values for social group size n^* as a function of the processor time costs $\tilde{c}$ required to achieve synchrony in an ecological objective, for three different cognitive strategies in an agent based model with a structured network. F=1: agents rely only on current information about agents they interact with; F=2: agents take notice of individual differences in the quality of information other agents offer; F=3: agents take note of individual differences between other agents and rely on third party information about each other received from other agents. Increasing calculation capacity allows larger groups in each of the three strategies for evaluating the quality of potential collaborators. However, each strategy has an upper limit (glass ceiling), and if groups need larger group size they have to switch to more sophisticated methods of information acquisition and that necessitates an increase in the computational costs (i.e. brain size). (Parameters used: k=4, τ=20, λ=11. The results are robust to these parameters.)

Note that the more complex strategies are highly disadvantageous when group size is small, and the more complex the strategy the more disadvantageous it is. Thus, the evolution of communicative and cognitive complexity is explicitly dependent on an ecological demand for large social groups. It is only when there is a need for large groups that the selection pressure will be sufficient to warrant the costs involved.

Our analysis provides direct support for both the social brain hypothesis[1] and the social (or communicative) complexity hypothesis[15] by demonstrating that greater information processing demands are reflected in greater cognitive (computational) costs, but that bringing these on stream allows organisms to significantly increase social group size. If size matters (large groups offer greater protection against predators[16] or are more efficient for foraging[17]) or win more territorial fights[18,19]), then there will be selection pressure to pay these costs. But, the significant finding is that these costs are considerable, and, when optimal group size is small, make the costs prohibitive. In these circumstances, simpler cognitive strategies are more profitable. Broadly speaking, this is what we see in the natural world.

REFERENCES

1 Dunbar, R. I. M. Neocortex Size as a Constraint on Group-Size in Primates. *J Hum Evol* **22**, 469-493 (1992).

2 Dunbar, R. I. M. The social brain hypothesis. *Evol Anthropol* **6**, 178-190 (1998).

3 Barton, R. A. Neocortex size and behavioural ecology in primates. *P Roy Soc Lond B Bio* **263**, 173-177 (1996).

4 Shultz, S. & Dunbar, R. Encephalization is not a universal macroevolutionary phenomenon in mammals but is associated with sociality. *P Natl Acad Sci USA* **107**, 21582-21586, doi:Doi 10.1073/Pnas.1005246107 (2010).

5 Byrne, R. B. & Whiten, A. (Oxford University Press, Oxord, 1988).

6 Barton, R. A. & Dunbar, R. I. M. in *Machiavellian Intelligence II. Extensions and evaluations* (eds A. Whiten & R.B. Byrne) (Cambridge University Press, 1997).

7 Dunbar, R. I. M. Coevolution of Neocortical Size, Group-Size and Language in Humans. *Behav Brain Sci* **16**, 681-694 (1993).

8 Sallet, J. *et al.* Social Network Size Affects Neural Circuits in Macaques. *Science* **334**, 697-700, doi:Doi 10.1126/Science.1210027 (2011).

9 Lewis, P. A., Rezaie, R., Brown, R., Roberts, N. & Dunbar, R. I. Ventromedial prefrontal volume predicts understanding of others and social network size. *Neuroimage* **57**, 1624-1629, doi:S1053-8119(11)00518-0 (2011)

10 Powell, J., Lewis, P. A., Roberts, N., Garcia-Finana, M. & Dunbar, R. I. M. Orbital prefrontal cortex volume predicts social network size: an imaging study of individual differences in humans. *P Roy Soc B-Biol Sci* **279**, 2157-2162, doi:Doi 10.1098/Rspb.2011.2574 (2012).

11 Bickart, K. C., Wright, C. I., Dautoff, R. J., Dickerson, B. C. & Barrett, L. F. Amygdala volume and social network size in humans. *Nat Neurosci* **14**, 163-164, doi:Doi 10.1038/Nn.2724 (2011).

12 Kanai, R., Bahrami, B., Roylance, R. & Rees, G. Online social network size is reflected in human brain structure. *P Roy Soc B-Biol Sci* **279**, 1327-1334, doi:Doi 10.1098/Rspb.2011.1959 (2012).

13 David-Barrett, T. & Dunbar, R. I. Cooperation, behavioural synchrony and status in social networks. *J Theor Biol*, doi:S0022-5193(12)00233-0 (2012).

14 Couzin, I. D., Krause, J., Franks, N. R. & Levin, S. A. Effective Leadership and Decision-making in Animal Groups on the Move. *Nature* **433**, 513-516, doi:nature03236 (2005).

15 Freeberg, T. M., Dunbar, R. I. & Ord, T. J. Social complexity as a proximate and ultimate factor in communicative complexity. *Philos Trans R Soc Lond B Biol Sci* **367**, 1785-1801, doi:rstb.2011.0213 (2012).

16 Shultz, S., Noe, R., McGraw, W. S. & Dunbar, R. I. M. A community-level evaluation of the impact of prey behavioural and ecological characteristics on predator diet composition. *P Roy Soc Lond B Bio* **271**, 725-732, doi:Doi 10.1098/Rspb.2003.2626 (2004).

17 Reader, S. M., Hager, Y. & Laland, K. N. The evolution of primate general and cultural intelligence. *Philos Trans R Soc Lond B Biol Sci* **366**, 1017-1027, doi:366/1567/1017 (2011).

18 Adams, E. S. Boundary Disputes in the Territorial Ant Azteca-Trigona - Effects of Asymmetries in Colony Size. *Anim Behav* **39**, 321-328 (1990).

19 Wilson, M. L. & Wrangham, R. W. Intergroup relations in chimpanzees. *Annu Rev Anthropol* **32**, 363-392 (2003).

Scaling Theory of Human Mobility and Spatial Network[a]

Pierre Deville,[1] Dashun Wang,[2,3] Chaoming Song,[2,3] Nathan Eagle,[4] Vincent Blondel,[1] and Albert-László Barabási[2,3,5]

[1] *Department of Applied Mathematics, Université Catholique de Louvain, Belgium*

[2] *Center for Complex Network Research, Department of Physics, Biology and Computer Science, Northeastern University, Boston, Massachusetts 02115, USA*

[3] *Center for Cancer Systems Biology, Dana Farber Cancer Institute, Boston, Massachusetts 02115, USA*

[4] *College of Computer Science, Northeastern University, Boston, Massachusetts 02115, USA*

[5] *Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA*

Abstract

With the increasing availability of large-scale datasets that simultaneously capture human movements and social interactions, advances in human mobility and spatial networks have rapidly proliferated during the past years, impacting in a meaningful fashion a wide range of areas, from epidemic prevention and emergency response to urban planning and traffic forecasting. As human mobility and spatial networks have developed in parallel, being pursued as separate lines of inquiry, we lack any known relationships between the quantities explored by them, despite the fact that they often study the same systems and datasets. Here, by exploiting three different cell phone datasets, we find a set of scaling relationships, mediated by a universal flux distribution, that link the quantities characterizing human mobility and spatial networks, showing that the widely studied scaling laws uncovered in the two areas represent two facets of the same underlying phenomena.

---

[a] Corresponding authors: pierre.deville@uclouvain.be, dashunwang@gmail.com

LOCAL COMPUTATION OF NETWORK CENTRALITY

BY CHRISTINA LEE ASUMAN OZDAGLAR DEVAVRAT SHAH*

LIDS, Dept. of EECS
Massachusetts Institute of Technology

Network centrality is a graph-based function that assigns scores to the nodes of a network. Many network centralities can be represented in terms of stationary distributions of random walks on the underlying graph – PageRank is an example of such a network centrality. The classical power-iteration method provides a simple iterative algorithm for computing such centralities. However, each iteration of this method requires global network-wide computation which becomes extremely challenging to perform over large graphs such as World Wide Web. In this paper, we propose a finitely terminated algorithm for approximately computing network centrality of a node which requires information about a local neighborhood of that node. We provide performance bounds on our estimates which highlight the dependence on the termination time and network properties. Our bounds yield tighter estimates for nodes with high centrality. Our analysis relies on the classical theory of positive-recurrence for countable state space Markov chain; as a byprodnct it provides meaningful implications towards the *mixing time* of such Markov chains which could be of independent interest.

Background. Networks have become ubiquitous representations for capturing interactions and relationships between entities across disciplines: social interactions between individuals, interdependence between financial institutions, hyper-link structure between web-pages or more generally correlations between distinct events. Many decision problems over networks rely on information about the importance of different nodes as quantified by network centrality measures. The celebrated PageRank, which is commonly used in Internet search algorithms, and the Bonacich centrality and eigencentrality measures, encountered in the analysis of social network are just two examples. Network centralities are often represented as stationary distributions of properly designed random walks (or Markov chains) over the network graph, a representation that is also useful for computation. Two classical approaches for computing network centralities are the power-iteration method, which involves iterative multiplication of the transition probability matrix of the random walk, and the Monte Carlo methods, which involve sampling

---

*Email addresses of authors are {celee, asuman, devavrat}@mit.edu.

states from a long random walk over the entire state space. The algorithm can take possibly unbounded time depending on the mixing properties of the underlying network. Both of these methods effectively require global information about the network, which would be prohibitive for large networks.

In this paper, we develop a *truncated Monte Carlo type method*, which ensures finite termination time and exploits local information. We provide performance guarantees as a function of the truncation time and network properties. Our analysis relies on the classical theory of positive-recurrence for countable state space Markov chains and uses a Lyapunov function over the state space with a negative drift outside a finite set of states to be able to bound the impact of truncation on the stationary probabilities.

Our work is most related to Monte Carlo methods for computing PageRank vector (see Jeh-Widom [4], Fogaras et al [3], Avrachenkov et al [1] which showed that PageRank is given by the distribution over the end nodes of a random walk with geometric length and Bahmani et al [2] which addressed how to incrementally update the PageRank vector for dynamically evolving graphs). The analyses in these papers rely on the specific structure of the random walk describing the PageRank, which involves at each step uniformly jumping to any other node with constant probability $\epsilon > 0$. This implies that the transition probability matrix P of the random walk can be decomposed as $P = (1-\epsilon)Q + \frac{\epsilon}{N}\mathbf{1}\mathbf{1}^T$. This decomposition is used to show that sampling with geometric length random walks obtains samples from the exact distribution. Therefore, concentration results follow naturally from Chernoff bounds for binomial distributions. In contrast, in this paper we focus on general network centralities, which can be represented as the stationary distribution of an irreducible, aperiodic and positive recurrent Markov chain and present an algorithm with performance guarantees.

Model and Notation. We are given a network represented by a graph $G = (\Sigma, E)$, where Σ is the (possibly countably infinite) set of nodes and E is the set of edges. Our goal is to compute a network centrality vector, which is given by the stationary distribution of an irreducible, aperiodic and positive-recurrent Markov chain on Σ with transition matrix $P : \Sigma \times \Sigma \to [0,1]$. We provide a simple finitely terminated local algorithm (defined below) that provides estimates for the stationary probabilities. We present performance bounds for our estimates as a function of the termination time and network properties. Our bounds are tighter for nodes with higher stationary probability.

Algorithm. To estimate the stationary probability π_i of node $i \in \Sigma$, we sample N independent paths of the Markov chain starting in state i for up to θ steps. Here N and θ are parameters of the algorithm (which determine the amount of local information used by the algorithm). Each of the sample-path of the Markov chain is either stopped if the Markov chain revisits the node i within first θ steps or it has taken θ steps without revisiting starting node i. Let t_k be the number of steps after which Markov chain was stopped in the kth run, $1 \leq k \leq N$. Then the estimate of the stationary probability of node i is given by

$$(1) \qquad \hat{\pi}_i = \frac{N}{\sum_{k=1}^{N} t_k} := \frac{1}{\hat{T}}.$$

Observe that the algorithm only needs to keep track of the location of the random walk and the number of steps taken. Therefore it can be distributed and run in parallel. In addition, since each sample path is only at most length θ, the algorithm may not need to store the entire network.

Main result. We start by recalling a sufficient condition for positive recurrent Markov chains.

Sufficient condition for positive-recurrence. An irreducible, aperiodic Markov chain $X_t, t \geq 0$, on countable state space Σ is positive recurrent if there exists a Lyapunov function $V : \Sigma \to \mathbb{R}_+$, a finite set $B \subset \Sigma$ and constants $\nu_{max}, \gamma > 0$ such that

1. $|V(y) - V(x)| \leq \nu_{max}$ for all $x, y \in \Sigma$ with $\mathbb{P}(X_{t+1} = y | X_t = x) > 0$,
2. $\mathbb{E}[V(X_{t+1}) - V(X_t) | X_t = x] \leq -\gamma, \ \forall\, x \in B^c$.

Due to the finiteness of the set B, the negative drift over B^c and bounded increment in the Lyapunov function V, it follows that the return time to a node $i \in B$ is exponentially concentrated, which is formally stated next.

Proposition 1 *For any $i \in B$ and for all $k \in \mathbb{Z}_+$, let T_i be return time of Markov chain to node i starting from $X_0 = i$. Then,*

$$(2) \quad \mathbb{P}\big(T_i > R_i \ln(2) + k\big) \leq e^{-\frac{k}{R_i}}, \quad \textit{where} \ \ R_i \approx O\Big(H_i \cdot e^{\frac{\gamma}{\nu_{max}}} \Big(\frac{\nu_{max}}{\gamma}\Big)^2\Big).$$

Here, the parameter $H_i = \max_{j \in B} \mathbb{E}[T_B | X_0 = j]$ where T_B is the return time to i for the chain watched only on the subset B.

We can use the preceding proposition to obtain for all θ,

$$\mathbb{E}[T_i] - \mathbb{E}[\min(T_i, \theta)] = \sum_{g=0}^{\infty} \mathbb{P}(T_i > g) - \sum_{g=0}^{\theta-1} \mathbb{P}(T_i > g) \tag{3}$$

$$\leq \frac{2e^{-\theta/R_i}}{1 - e^{-1/R_i}} := \epsilon_\theta \tag{4}$$

Theorem 1 *For any* $i \in B$ *and for all* $\epsilon > 0$ *and* $\lambda > 0$

$$\mathbb{P}\left(\left|\mathbb{E}[T_i] - \hat{T}\right| \geq \epsilon_\theta + \epsilon\right) \leq e^{-(\phi' + \lambda\epsilon)N} + e^{-(\phi + \lambda\epsilon)N} \tag{5}$$

Thus,

$$\mathbb{P}\left(|\pi_i - \hat{\pi}_i| \geq \frac{(\epsilon_\theta + \epsilon)\pi^2}{1 - (\epsilon_\theta + \epsilon)\pi}\right) \leq e^{-(\phi' + \lambda\epsilon)N} + e^{-(\phi + \lambda\epsilon)N} \tag{6}$$

where ϕ', ϕ *are fixed values that depend on the graph properties, the chosen node* i*, and* λ*.*

This theorem highlights the two main sources of error in the algorithm's estimation. The first source of error, ϵ_θ, is a result of the trunction. It is deterministic and depends on the chosen threshold θ and the graph and target node properties $\frac{\nu_{max}}{\gamma}$ and H_i. The second source of error ϵ is the sampling error, a probabilistic quantity. The parameter ϕ and ϕ' which dictate rates of convergence also depend on the graph properties and the chosen target node.

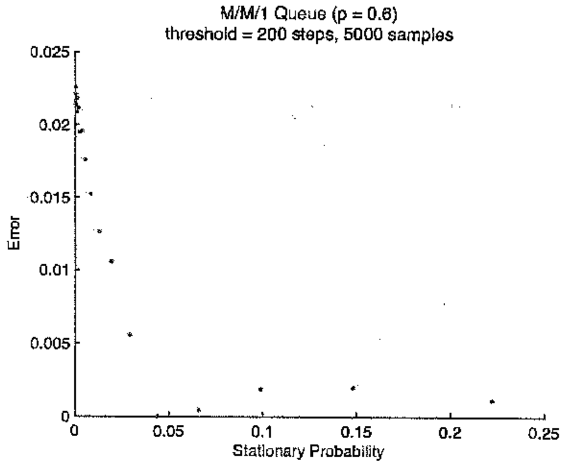
Simulations. We tested our algorithm on two types of markov chains. The first example uses a M/M/1 queue, where

$$p = \mathbb{P}(\text{departure occurs before an arrival})$$

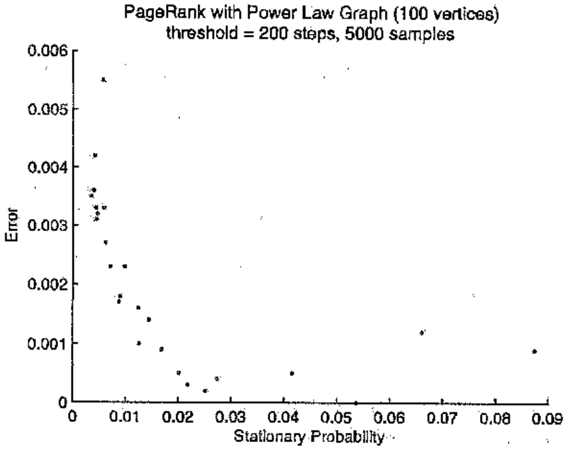
The second example uses the PageRank markov chain over a randomly generated graph. The random graph is generated with the configuration model, where the degree distribution is chosen according to the power law with parameter $\alpha = 1.5$.

$$\mathbb{P}(d) = cd^{-\alpha}$$

The plot shows the stationary probability of the target node and the error of the estimate. As expected, for nodes with high stationary distribution as compared to $\frac{1}{\theta}$, the algorithm performs well with low error.



M/M/1 Queue (ρ = 0.6)
threshold = 200 steps, 5000 samples



PageRank with Power Law Graph (100 vertices)
threshold = 200 steps, 5000 samples

REFERENCES

[1] K. Avrachenkov, N. Litvak, D. Nemirovsky, and N. Osipova. Monte carlo methods in pagerank computation: When one iteration is sufficient. *SIAM Journal on Numerical Analysis*, 45(2):890–904, 2007.

[2] B. Bahmani, A. Chowdhury, and A. Goel. Fast incremental and personalized pagerank. *Proc. VLDB Endow.*, 4(3):173–184, Dec. 2010.

[3] D. Fogaras, B. Racz, K. Csalogany, and T. Sarlos. Towards scaling fully personalized page rank: Algorithms, lower bounds, and experiments. *Internet Mathematics*, 2(3):333–358, 2005.

[4] G. Jeh and J. Widom. Scaling personalized web search. In *Proceedings of the 12th international conference on World Wide Web*, WWW '03, pages 271–279, New York, NY, USA, 2003. ACM.

Social Learning and Network Uncertainty

Ilan Lobel and Evan Sadler*

July 25, 2012

EXTENDED ABSTRACT

We consider the perfect Bayesian equilibria of a model of social learning in networks where agents have imperfect knowledge of the social network topology. Each agent receives a signal about an underlying state of the world, observes the actions of her neighbors and subsequently chooses an action herself. The topology of the social network is drawn from a commonly known distribution, but each agent only observes her own neighborhood. We characterize properties of social networks that lead to or preclude the successful aggregation of information in society.

Our model is a generalization of that introduced by Acemoglu et al. [2011]. While Acemoglu et al. [2011] assume that neighborhoods of different individuals are realized independently from one another, we allow for arbitrary distributions over the space of social networks. This extension allows us to capture real-world network phenomena such as clustering and assortativity and to consider the performance of social learning in widely used models of social networks such as preferential attachment models.

When a network topology exhibits these phenomena, agents in the network may disagree over who is well-informed or well-connected, and we say that the topology features *network uncertainty*. Network uncertainty forces us to reevaluate our understanding of social learning dynamics and whether networks successfully aggregate dispersed information. According to the last two decades of economics scholarship, the key inefficiency in the social learning process is the phenomenon of herding, as first described by Banerjee [1992] and Bikhchandani et al. [1992]. While our model still captures herding behavior, we find that network uncertainty can cause more significant failures of information aggregation to occur.

We argue that previous papers have set the bar too high for social learning. The traditional metric for successful information aggregation is whether, as society grows large, later agents approach certainty about the underlying state of the world. When this occurs, we say that learning is *absolutely successful.* Prior work has shown that absolute success can be difficult to achieve when agents only observe the discrete actions taken by their neighbors, instead of actual signals or beliefs. However, in many networks, agents benefit substantially from observing their peers without approaching full knowledge of the state of the world.

---

*Stern School of Business, New York University – {ilobel, esadler}@stern.nyu.edu

To better understand the extent to which learning occurs, we introduce a second metric for social learning motivated by the notion of an expert—an individual with an exceptionally strong private signal. We say that learning is *successful* if all agents perform at least as well as an expert in the limit as society grows. We study conditions leading to both success and absolute success. If neither metric is attained, we say that social learning has *failed* to aggregate information.

One reason prior literature has focused on absolute success, rather than success, is that the complete network studied in early papers (e.g. Banerjee [1992], Bikhchandani et al. [1992], and Smith and Sorensen [2000]) always successfully aggregates dispersed information. In fact, we show that herding can only occur once social learning has been successful. At worst, herding prevents absolute success. More recent papers have studied other deterministic network topologies (Celen and Kariv [2004]) as well as stochastic topologies where agents have independently drawn neighborhoods (Banerjee and Fudenberg [2004], Smith and Sorensen [2008], Acemoglu et al. [2011]). One may plausibly conjecture that many such networks fail to aggregate information, but we show this is essentially false. In the absence of network uncertainty, learning is always successful so long as the network features expanding observations. This is a mild condition introduced by Acemoglu et al. [2011] that serves only to eliminate insufficiently connected networks.

Without network uncertainty, learning fails only if the network is too disconnected, but with network uncertainty, we find that learning can fail via a multitude of complex mechanisms. Learning might fail because agents cannot identify an information path even though it exists. Learning may also fail because agents are either underconfident or overconfident in their neighbors, so they incorrectly weigh their private beliefs against their social beliefs. Finally, learning can fail because network uncertainty creates correlations between the actions of different neighbors that render them uninformative. These modes of failure by no means constitute an exhaustive list, but we have catalogued several of the mechanisms through which network uncertainty can disrupt the learning process.

Our main positive result provides sufficient conditions for successful information aggregation. We introduce the notion of a *neighbor choice function*—a rule that selects one especially trustworthy person from each agent's neighborhood. Given a network topology and a neighbor choice function for each agent, we obtain a *trusted network topology.* This topology represents a network where each agent considers at most one neighbor's action when formulating a decision, discarding the information provided by all other neighbors. We show that successful social learning occurs if there exists a trusted network topology that is sufficiently connected and has *low (belief) distortion.* Low distortion provides the key new constraint: if the act of observing another agent does not excessively affect this person's probability of being correct, then the network has low distortion. We also demonstrate that in special cases these conditions are both necessary and sufficient for successful learning.

These conditions grant us substantial freedom to select neighbor choice functions, and via appropriate selection we can use this result to demonstrate successful information aggregation in many types of networks. For instance, any network with long deterministic information paths automatically satisfies these conditions. Since low belief distortion is often challenging to verify directly in a network, we also introduce the notion of *network distortion*. Network distortion captures how being observed by someone else affects the composition of an agent's neighborhood. We show that network distortion provides an upper bound on belief distortion; thus, the existence of a trusted network topology that is sufficiently connected and has low network distortion leads to successful learning. Models based on preferential attachment—a popular generative model used to simulate real-world social networks—satisfy these properties, thereby leading to successful information aggregation.

References

D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar. Bayesian Learning in Social Networks. *The Review of Economic Studies*, 78(4):1201–1236, 2011.

A. Banerjee. A Simple Model of Herd Behavior. *The Quarterly Journal of Economics*, 107: 797–817, 1992.

A. Banerjee and D. Fudenberg. Word-of-Mouth Learning. *Games and Economic Behavior*, 46:1–22, 2004.

S. Bikhchandani, D. Hirshleifer, and I. Welch. A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades. *The Journal of Political Economy*, 100:992–1026, 1992.

B. Celen and S. Kariv. Observational Learning under Imperfect Information. *Games and Economic Behavior*, 47(1):72–86, 2004.

L. Smith and P. Sorensen. Pathological Outcomes of Observational Learning. *Econometrica*, 68:371–398, 2000.

L. Smith and P. Sorensen. Rational Social Learning with Random Sampling. Working Paper, 2008.

INTERMEDIATION AND EXCLUSIVE REPRESENTATION IN FINANCIAL NETWORKS

ITAY P. FAINMESSER

Extended abstract. This paper develops a theory of financial intermediation based on the resolution of incentives problems via repeated interactions. Consider a potential lender (e.g. investor) and a borrower (e.g. entrepreneur) who needs to raise capital for a risky project. Even if both parties know the expected returns of the project and observe its outcome, limited liability combined with lack of verifiability of the realization of the risky investment leave open the possibility of strategic default. If the frequency with which the lender has liquidity (when the borrower needs liquidity) is low, then strategic default by the borrower cannot be deterred by a threat of losing access to future funds from the lender. However, a financial intermediary (e.g. an investment bank) who exclusively represents a large pool of lenders in their transactions with a borrower can still enforce repayment by the borrower. The intermediary can do that by threatening to eliminate the borrower's access to future funds from many lenders.

To be specific, we show that if interactions between any lender and borrower are infrequent and if market participants have incomplete knowledge of the patterns of interactions in the market, then all investments must be intermediated. Moreover, each intermediary must exclusively represent many lenders in their interactions with a given borrower, so that she can "punish" the borrower severely for any strategic default by eliminating the borrower's access to many future loans. To this end, we develop a model of financial networks that are shaped by exogenous forces as well as by lenders' decisions and new tools to study the ability of market participants to learn about the structure of the financial network. We then characterize networks that are robust – networks that can be sustained in equilibrium given (almost) any belief that is consistent with agents' knowledge of the network structure. Our characterization sheds light on the complementarity and substitutability of self-finance clauses, the use of collateral, and intermediation; and suggests also that the riskiest assets will be traded by intermediaries without full collateral. The effect of macroeconomic conditions and the presence of credit bureaus on the patterns of intermediation are also studied.

Clearly, we are not the first to study game theoretic foundations for the enforcement of informal contracts. The literature on community enforcement offers two enforcement mechanisms to explain the prevalence of informal contracts in the presence of incentives problems. One mechanism is *ostracism.* Ostracizing a borrower requires coordination. In some markets coordination is achieved by tight social groups, i.e. family or an ethnic group. When a market is not dominated by social groups, coordination requires common observations and common knowledge of

---

Fainmesser: Economics Department, Brown University, Providence, RI 02906. E-mail: *itay_fainmesser@brown.edu.*

the patterns of interactions between agents. A second mechanism suggested in the literature is *contagion* – any agent who observes a default reacts by defaulting (if a borrower) or by avoiding the provision of liquidity (if lender), independent of the identity of their trading partner. Contagion is hard to motivate in large markets, and requires implicit coordination between agents in order to provide the incentives to spread 'bad behavior' to the entire population.

In this paper we propose a third mechanism – *intermediation*. If each of a group of lenders agrees to invest with a given borrower only via a given intermediary, the intermediary can single-handedly 'cut off' a defaulting borrower's access to liquidity from a large group of lenders. Our analysis suggests that well positioned intermediaries can enforce repayment in environments in which ostracism and contagion cannot.

A novel feature of our model is that agents are not assumed to observe the network structure directly; agents observe their own financial interactions, and their knowledge of the network is derived as an upper bound on what they would be able to learn about the network structure based on their observations in many such interactions. The idea that agents do not observe the network structure directly, but rather infer the network structure from their observations during their own interactions is reasonable given that the network is not a physical object, but rather a collection of relationships that generate the activity in the economy.

Given that agents receive information only on parts of the network that affect their own financial interactions, some forms of community enforcement are infeasible. If a borrower b strategically defaults, his link with the lender or intermediary, say k, who provided him with the liquidity is lost. However, for any additional intermediary or lender, say j, to disconnect her link to the borrower, two conditions need be fulfilled: [1] j observes the default (or the elimination of the link between k and b); and [2] given her observations and beliefs, j has the incentives to eliminate her link to b rather then "pretend" not to have observed the default (or the elimination of the link between k and b). For example, j may prefer to "cover-up" b's default if j believes that other lenders or intermediaries did not observe the default (or the elimination of the link between k and b) and that b has sufficient incentives not to default on j as long as no additional links are eliminated.

The main result of the paper offers a complete characterization of the set of networks that are *robust* – networks that can be sustained in pure strategy perfect Bayesian equilibria of the infinitely repeated game given any belief from a large set of beliefs that we consider. We show that there exists a mapping from the parameters of the model to a positive integer m such that in robust networks any active intermediary is an *m-local monopoly* – for every borrower who the intermediary is connected to, she is also connected to at least m lenders who are not connected to the borrower in any other way, either directly or via another intermediary. That is, any intermediary exclusively represents at least m lenders in transactions with any borrower that she is connected to. Figure 0.1 demonstrates the notion of local monopolism.

We also show that if the parameters of the model are such that $m > 1$, then in all robust networks *all* investments are intermediated. This explains the presence of intermediaries even

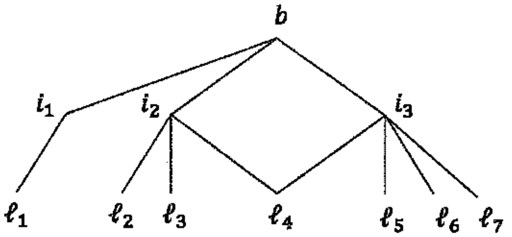


FIGURE 0.1. Local monopolism. Intermediary i_1 is a 1-local monopoly, i_2 is k-local monopoly for any $k \leq 2$, and i_3 is k-local monopoly for any $k \leq 3$.

in markets in which there is no exogenous cost advantage to intermediation. Our result also highlights that the important factor is not the absolute size of an intermediary, or the overall number of investments that she intermediates, but rather the exclusivity over a sufficient number of *investment paths*. Such exclusivity can be achieved by a large intermediary, but it can also be achieved by an intermediary who specializes and focuses on a small (but not too small) number of borrowers and lenders that cannot transact otherwise. E.g., an intermediary can focus on local businesses, or provide a connection between otherwise disconnected communities.

By relating m to the parameters of the model, we are able to show that the minimal level of monopoly power that an intermediary is required to hold (as captured by m) decreases in the frequency of arrival of investment opportunities and the expected return on investment, and increases in the borrowers discount rate and in the return on capital demanded by lenders.

From a macroeconomic perspective, the model predicts that in times of economic booms (high frequency of arrival of investment opportunities and high expected returns on investments) there is room for a large number of intermediaries and more competitive markets in which no single intermediary has significant market power (as captured by m). On the other hand, in times of economic downturns, especially ones that are triggered by liquidity crunches, the model predicts a smaller number of intermediaries, each with significant market power.

We also find that requiring that a borrower self-finance a positive fraction of the investment opportunity reduces the monopoly power required by intermediaries in order to enforce repayment, and that the same is true for partial collateral and for the presence of bankruptcy laws. On the other hand, we find that introducing the possibility of pledging full collateral increases the monopoly power that intermediaries are required to have in order to enforce repayment of uncollateralized investments. Therefore, if the cost of pledging full collateral is sufficiently low, the market may revert to simple debt contracts even when equity-like contracts are more efficient. In particular, full collateral contracts are likely to undermine the role of intermediaries in markets for less risky assets, and intermediaries are likely to continue trading the riskiest assets without collateral.

Finally, we show that central credit information agencies (a generalized version of credit rating agencies) may relax the requirement of local monopolism, but do not eliminate the need for intermediaries.

Competitive Marketing Strategies over Social Networks

Kostas Bimpikis [*] Asuman Ozdaglar [†] Ercan Yildiz [‡]

It is widely accepted that word-of-mouth plays a central role in the propagation of brand or product information, and, thus, it is a first order consideration in the design and implementation of a marketing strategy. Moreover, marketers, nowadays, have access to and can take advantage of vast amounts of data on the pattern and intensity of social interactions between consumers. The advent of the Internet as a prominent communication and advertising platform has enabled firms to implement targeted marketing campaigns and direct their efforts to certain subsets of the population. The natural question that arises in this setting is whether a firm can use the wealth of available information along with the targeting technologies to increase the awareness about its products. The recent acquisitions of Buddy Media, Vitrue, and Wildfire Interactive by Salesforce, Oracle, and Google respectively point to the indisputable fact that social media marketing emerges as a viable alternative to traditional advertising and tech giants are striving to obtain a competitive advantage in the new landscape.

The focus in this paper is to prescribe the best way a firm can exploit word-of-mouth and its knowledge over the social network structure of consumers when devising a targeted marketing campaign. At the core of the model lies an *information externality*: information obtained by an agent in the network can be passed along to her peer group and, thus, word of mouth communication among agents may amplify the effect of a firm's marketing efforts. As a consequence, an optimal targeting strategy may involve allocating disproportionate fraction of the advertising budget to certain agents that play an influential role in the word of mouth process expecting that they will pass the relevant information to the rest. Indeed, strategies of similar flavor have been applied in practice, however in an ad hoc and heuristic way.[1] Our goal is to provide a systematic characterization of optimal targeting strategies and identify qualitative insights that lead to their success.

We study a dynamic model of information exchange among agents embedded in a given social network. Agents communicate with their peers over time and obtain information about a number of products. Specifically, when two agents communicate, then one suggests a product to the other in proportion to her belief that the product is better than the competition. Individual beliefs evolve over time incorporating information that agents obtain. Finally, firms can influence this process by targeting advertising funds to specific individuals, thus, making them more prone to pass information regarding their product. As a motivating example consider a group of consumers acquiring information about competing firms in a new consumer market (e.g., smartphones, tablets). There is a priori uncertainty about the relative quality of the products and firms can assist consumers through

---

[*] Graduate School of Business, Stanford University.

[†] Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.

[‡] MultiChannel Interactions Research and Development Group, Accenture Technology Labs.

[1] For example, Klout (www.klout.com), a San Francisco-based company, provides social media analytics to measure a user's influence across his or her social network. The service scrapes social network data and assigns individuals a "Klout" score, which presumably reflects their influence. It, then, connects businesses with individuals of high score with the intention of influencing the latter to spread good publicity for the former in exchange for free merchandise and other perks.

informative advertising.

We are interested in the limiting behavior of this dynamical system and how it relates to the underlying network structure and the marketing strategies of the competing firms. First, we show that the limiting behavior of the system is well defined, as the beliefs of agents converge to a fixed vector. Next, we provide a sharp characterization of the average belief at the limit as a function of the underlying network structure and the advertising efforts of the firms. It turns out that the average limiting belief is equal to a weighted sum of the advertising funds that firms allocate to individuals, where the weights are given by a novel notion of *centrality* of the agents in the underlying network structure.

Our main objective is to devise optimal marketing strategies for firms over the existing social network structure. Armed with a complete characterization of the limiting behavior of the dynamical system that represents the information exchange process among agents, we proceed to define the following optimization problem for a firm: given the underlying network structure and the marketing strategies of the competing firms, what is the best allocation of advertising funds to individuals, so as the firm maximizes the expected average belief about its product over the population of agents. Interestingly, we provide a closed-form expression for the optimal budget allocation to an agent: a firm should optimally spend advertising funds to an individual in proportion to the latter's centrality in the network. As a corollary of this characterization, we obtain the following intuitive result: if the network structure is complete, i.e., the level of interaction between any pair of agents is the same, then targeting does not offer any competitive advantage and it is best for the firm to allocate its advertising budget uniformly over the population of agents.

In the second part of the paper, we model the competition between two firms that offer substitutable products as a two-stage game over the network of agents. In particular, in the first stage, firms simultaneously choose their marketing strategies, i.e., how to allocate their advertising budgets over the population of agents. Then, in the second stage, agents obtain information both from their peers and the firms over time until their beliefs about the relative quality of the products converge to a limiting vector. We are interested in characterizing the equilibrium strategies for the firms and derive qualitative insights for the relation of the network structure with the competition between them. Specifically, firms compete the fiercest for agents that exhibit high centrality and allocate a disproportionate fraction of their budgets on them. We provide conditions on the dynamics of the information exchange process and the network structure that lead to *asymmetric* allocations at equilibrium: the two firms target different sets of consumers and effectively divide the markets into submarkets, where they act as local monopolies.

As a way of further illustrating the effectiveness of targeting technologies, we compare the profits of a firm in the following two settings: first, when the firm has complete knowledge over the social network structure and can fully exploit it by targeting its advertising efforts and, second, when the firm simply allocates its advertising budget uniformly over the entire population of agents. We provide a characterization of the difference in the limiting average belief for the quality of the product in the two settings as a function of the underlying network structure and highlight the relation of the success of a targeted marketing strategy with the level of heterogeneity in the network interactions among agents: the more homogeneous these interactions are, the less effective a targeted marketing strategy is compared to one that is agnostic of the underlying network. To further build intuition, the two extremes are, on the one hand, the complete network, where an agent interacts with all other agents in the same way, and the star network, where the only interactions present are between the "star" node and the rest of the agents. In the former, targeting does not offer any advantage compared to uniform advertising, whereas in the latter the difference in the limiting behavior of the agents is maximized.

EXTENDED ABSTRACT FOR SUBMISSION TO LIDS CONFERENCE @MIT

Nine Facts, Three Parameters and One Theory:
A Precise Analytical Model for the Dynamics of Industry Location Networks

Ravi Charan and Cesar A. Hidalgo
Macro Connections, The MIT Media Lab
rcharan@mit.edu; hidalgo@mit.edu

INTRODUCTION:

Understanding the process by which countries and locations diversify their industrial structures has been historically difficult. In fact, many theories have been constructed to explain the process by which industrial structures evolve.

On the one hand, there are the approaches followed by classical theories of structural transformation and development, where the focus is on the ability of a country to move its economy away from an aggregate sector, such as "agriculture", and into another, such as "manufacturing". On the other hand, we have stylized formal models of location, like those in the new economic geography of Krugman, Fujita and Venables. Moreover, we have empirical approaches, like as those followed by Glaeser, Boschma, Frenken and Saviotti, where the focus has been on measuring the different types of externalities that affect the location of an industry.

All of these approaches, however, have two things in common. On the one hand, they assume that industries require the existence of a set of factors, capabilities or characteristics. These could range from the distance between an industry and private and public inputs, such as workers, resources and institutions, to the availability of a local market where the industry can sell its specific output. On the other hand, these theories assume that there is something special about each place, in the sense that some of these characteristics are presents in some place but not others.

In this paper we present an analytical model that uses networks to formalize this two observations and use empirical data to show that the model can reproduce both, the structure and the dynamics of the networks connecting countries to products, and connecting industries to locations. We find the evidence in favor of the model to be strong, since we show that the model can reproduce the dynamic patterns observed in the data after having calibrated all the fitting parameters in the cross section. This shows that the model makes accurate out of sample predictions, with no fitting parameters, regarding the dynamics of industry location networks.

MODEL DESCRIPTION:

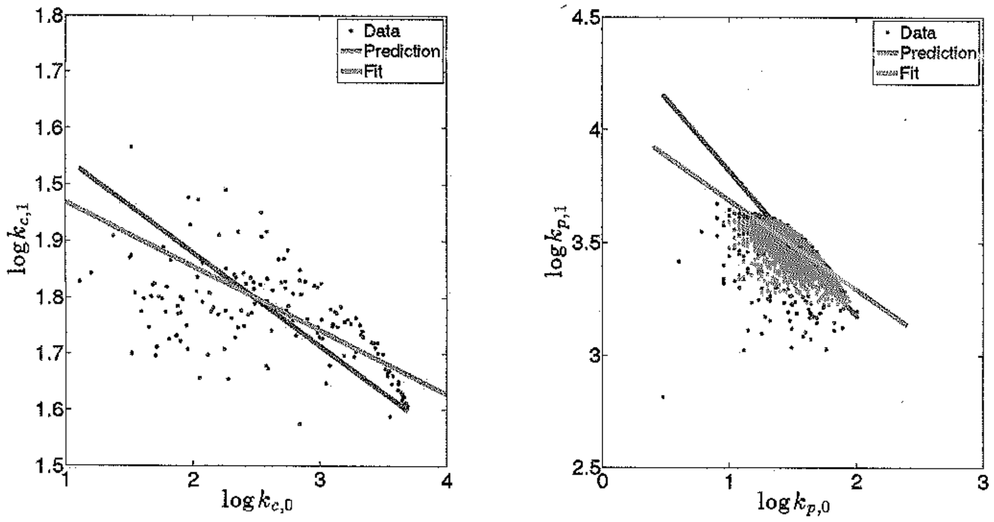
Our model is based on three bipartite networks. One connecting locations (c) to the characteristics (a) that are present in them ($\boldsymbol{C}_{ca}$), another connecting industries (p) to the characteristics (a) they require ($\boldsymbol{P}_{pa}$), and a third one, connecting each location c to the industries p present in it ($\boldsymbol{M}_{cp}$). We assume that industry p will be present at location c ($\boldsymbol{M}_{cp}=1$) only when the characteristics required by the industry are a subset of those present at the location. We solve a particular case of this model in which locations have characteristics with probability r, industries require characteristics with probability q, and there are N_a characteristics in the world.

RESULTS

To compare the data and the model we calculate a series of observables of $\boldsymbol{M_{cp}}$ -the adjacency matrix connecting industries to locations- using data connecting both, countries to the products they export (BACI), and cities to the industries present in them (U.S. M.S.A. Census). Here, we distinguish between static, or cross sectional observables, and dynamic observables. The latter of these observables involve changes in $\boldsymbol{M_{cp}}$ over time.

We show that based on the aforementioned assumptions it is possible to derive theoretical predictions about (i) the average diversity of locations, (ii) the average ubiquity of industries, (iii) the fill of $\boldsymbol{M_{cp}}$, (iv) the average ubiquity of the industries present at a location and (v) the average diversity of the locations where an industry is present. We show that the latter of these three can be used to calibrate all of the model parameters (r, q and N_a).

Also, we derive theoretical predictions for four dynamic observables. These connect the diversity of a location, and the average ubiquity of its industries in the present, with the ubiquity of a location's new industries (vi - vii), and the average diversity of the locations where a location's new industries used to be present (viii - ix). We show that the model predicts that three of these relationships will follow a power, and one of them to be linear, and find these predictions to hold empirically. Moreover, we find that the predicted exponents of the power-laws are functions only of the parameters calibrated in the cross sections (r, q and N_a). We compare the empirically observed exponents with those predicted with the theory and find these to agree within measurement error. These shows that the model makes predictions for the dynamic properties of the network connecting industries to locations that are able to reproduce the network properties in absence of fitting parameters.

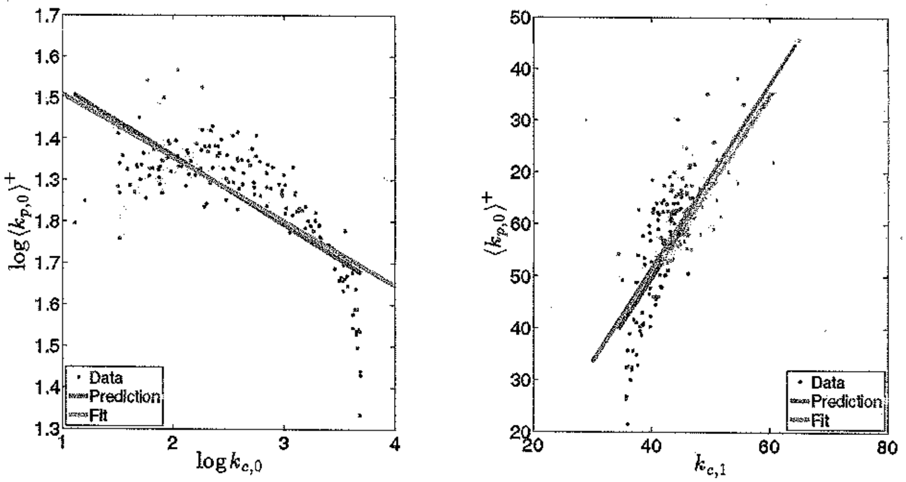


(a) $\log k_{c,1}$ vs $\log k_{c,0}$

(b) $\log k_{p,1}$ vs $\log k_{p,0}$

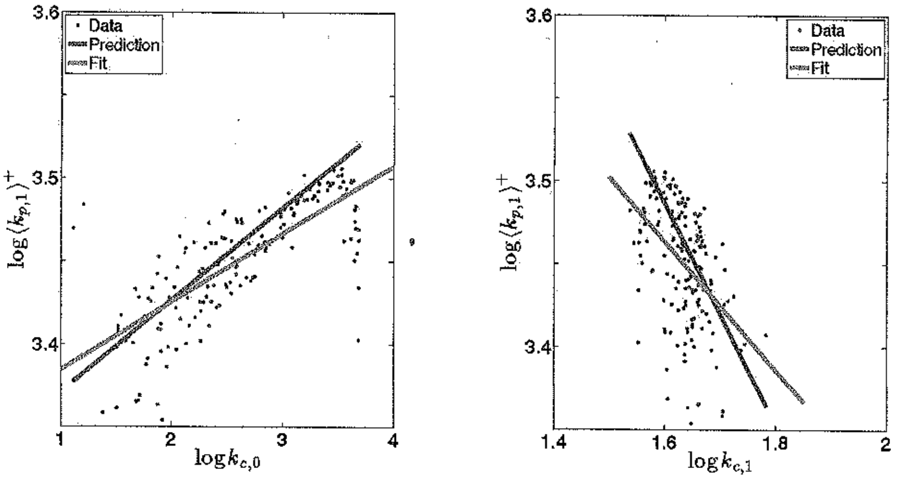
				Prediction			Best Fit			
Data	Relationship			R^2	R^2 (tail)	Coeff.	R^2 (tail)	Coeff.	95% CI	
HS6	$k_{c,0}$	vs	$k_{c,1}$	-0.159	0.299	-0.083	0.461	-0.057	-0.068	-0.046
HS6	$k_{p,0}$	vs	$k_{p,1}$	-0.112	0.275	-0.642	0.471	-0.393	-0.406	-0.381

Figure 1. Predicted relationship (red), best-fit line (green), and data (blue) between cross-sectional observables for the HS6 dataset. (a) average ubiquity of products produced ($k_{c,1}$) and diversity ($k_{c,0}$), and (b) average diversity of countries product a product ($k_{p,1}$) versus ubiquity ($k_{p,0}$). Both relationships are predicted to be power laws. The predictions are made based on calibrations produced from (iii), (iv), and (v), above. The table gives the R^2 values as well as predicted exponents (Coeff.) and the 95% confidence interval. The best-fit is performed to the tail; a cutoff of 100 for $k_{c,0}$, and a cutoff of 20 for $k_{p,0}$.



(a) $\log\langle k_{p,0}\rangle^+$ vs $\log k_{c,0}$

(b) $\langle k_{p,0}\rangle^+$ vs $k_{c,0}$



(c) $\log\langle k_{p,1}\rangle^+$ vs $\log k_{c,0}$

(d) $\log\langle k_{p,1}\rangle^+$ vs $\log k_{c,1}$

		Prediction		Best Fit			
Data	Relationship	R^2	Fit Coefficient	R^2	Fit Coefficient	95% CI	
HS6	$\langle k_{p,0}\rangle^+$ vs $k_{c,0}$	0.531	-0.083	0.535	-0.077	-0.088	-0.066
HS6	$\langle k_{p,0}\rangle^+$ vs $k_{c,1}$	0.541	0.870	0.569	0.892	0.878	0.906
HS6	$\langle k_{p,1}\rangle^+$ vs $k_{c,0}$	0.430	0.055	0.578	0.041	0.035	0.046
HS6	$\langle k_{p,1}\rangle^+$ vs $k_{c,1}$	-0.040	-0.668	0.215	-0.387	-0.503	-0.271

Figure 2. Predicted relationship (red), best-fit line (green), and data (blue) between cross-sectional observables for the HS6 dataset. (a) ubiquity of new industries versus diversity. (b) ubiquity of new industries versus mean ubiquity of products. (c) diversity of co-exporters of new products versus diversity. (d) diversity mean ubiquity of products. Three relationships are predicted to be power laws, while (b) is a proportional relationship. The predictions are made based on calibrations produced from (iii), (iv), and (v), above, alongside the fill of M_{cp}. The table gives the R^2 values as well as predicted exponents (Coeff.) and the 95% confidence interval.

CONCLUSIONS

In this presentation we show that it is possible to construct a model that predicts the structure and dynamics of the networks connecting industries to locations. We find that the model connects with the theory with a great degree of precision, since it makes predictions about non-trivial dynamics exponents using only parameters calibrated in the cross-section. We take this as strong evidence in favor of the process of economic development and diversification to be the result of a combinatorial, rather than aggregation, process by which locations accumulate characteristics that were until now not present in them.

Trends Prediction in Social Networks Using Influence Diffusion Models

Yaniv Altshuler
MIT Media Lab
77 Massachusetts Ave.
Cambridge, MA, 02139
yanival@media.mit.edu

Wei Pan
MIT Media Lab
77 Massachusetts Ave.
Cambridge, MA, 02139
panwei@media.mit.edu

Alex "Sandy" Pentland
MIT Media Lab
77 Massachusetts Ave.
Cambridge, MA, 02139
sandy@media.mit.edu

Overview: The importance of the ability to predict emergence and evolution of trends in human communities and social networks has been growing rapidly in the past few years with the increased dominance of social media in our everyday's life. Whereas many works focus on detection of anomalies in networks, there exist little theoretical work on the prediction of the likelihood of anomalous network patterns to globally spread and emerge into prominent "trends".

In this work we present an analytic model for the social diffusion dynamics of spreading patterns in (social) networks. Our proposed model generalizes a known local influence model, and is capable of predicting future trends based on the analysis of past social interactions between the community's members.

Contribution: Recently, a model for the local dynamics of social influence was developed (Pan, Aharony, and Pentland 2011), where the probability that a certain behavior would "migrate" from one user to her friends was shown to be well approximate by the $f(t) = 1 - e^{-(\xi_1+\xi_2 t)}$, where ξ_1, ξ_2 are derived from past activity of the social network. We have generalized this model, giving analytically provable lower and upper bounds for the probabilities that certain behaviors would evolve into global network trends.

Validation: We have validated our model using real-world date from two comprehensive social datasets. The first is the *Friends and Family* experiment (Aharony et al. 2011), held in MIT for over a year, where the complete activity of 140 users was analyzed, including data concerning their calls, SMS, MMS, GPS location and social media activities. The second dataset contains the complete financial transactions of over 2.5 million members of the *eToro* community – the world's largest "social trading" platform, allowing users to trade in currency, commodities and indices by selectively copying trading activities of prominent traders.

Model: One of the main difficulties of trends-prediction stems from the fact that the first spreading phase of "soon to be global trends" demonstrates significant similarity to other types of anomalous network patterns. In other words, given several observed anomalies in a social network, it is very hard to predict which of them would result in a wide-spread trend and which will quickly dissolve into oblivion.

We model the community, or social network, as a graph G, that is comprised of V (the community's members) and E (social links among them). We use $n = |V|$ to denote the size of the network, in which we are interested in predicting the future behavior of some observed anomalous pattern a. Notice that a can refer to a growing use of some new web service such as *Groupon*, or alternatively a behavior such as associating oneself with the "*99% movement*".

Note that "exposures" to trends are transitive. Namely, an "exposing" user generates "exposure agents" that are transmitted on the network's social links to "exposed users", which in turn if "convinced", can transmit them onwards. We hence model trends' exposure interactions as movements of random walking agents in a network, assuming that very user that was exposed to a trend a generates β such agents, on average (β can be estimated using past observations).

We assume that our network is (or can be approximated by) a scale free network $G(n, c, \gamma)$, namely, a network of n users where the probability that user v has d neighbors follows a power law :

$$P(d) \sim c \cdot d^{-\gamma}$$

We also define the following properties of the network (given for the complete definition of Theorem 1. The Proof of Theorem 1 is omitted due to space considerations) :

Definition 1. *Let $V_a(t)$ denote the group of network members that at time t advocate the behavior associated with the potential trend a.*

Definition 2. *Let us denote by $\beta > 0$ the average* "diffusion factor" *of a trend a. Namely, the average number of friends a user who have been exposed to the trend will be talking about the trend with (or exposing the trend in other ways).*

Definition 3. *Let P_Δ be defined as the probability that two arbitrary members of the network, u and v, have degrees ratio of Δ or higher :*

$$P_\Delta \triangleq Prob\left[deg(u) > \Delta \cdot deg(v)\right]$$

Definition 4. *We denote by σ_+ and σ_- the* "high temporal resistance" *and* "low temporal resistance" *of the network :*

$$\forall t, \forall \Delta_t \text{ , and for } \sigma(\Delta, t) \triangleq e^{\Delta \cdot \frac{\beta^{\Delta t} \cdot |V_a(t)}{n}} \quad :$$

$$\sigma_+ \triangleq \min \left\{ \begin{matrix} 1 \leq d \leq 2 \\ \Delta \geq 1 \end{matrix} \middle| 1 - \left(1 - \frac{1 - \frac{d-1}{d}^{\gamma - 1}}{c^{-1}(\gamma - 1)} \right) \frac{1 - P_\Delta}{\sigma(\Delta, t)^d} \right\}$$

$$\sigma_- \triangleq \max \left\{ \Delta \geq 1 \middle| 1 - \frac{1 - P_\Delta}{\sigma(\Delta, t)} \right\}$$

Definition 5. *Let $P_{Local-Adopt}(a, v, t, \Delta_t)$ denote the probability that at time $t + \Delta_t$ the user v had adopted trend a (for some values of t and Δ_t). This probability may be different for each user, and may depend on properties such as the network's topology, past interactions between members, etc.*

Definition 6. *Let P_{Local} denote the expected value of the local adoption probability throughout the network :*

$$P_{Local} = \underset{u \in V}{\mathrm{E}} \left[P_{Local-Adopt}(a, u, t, \Delta_t) \right]$$

Definition 7. *Let us denote by $P_{Trend}(\Delta_t, \frac{V_a(t)}{n}, \varepsilon)$ the probability that at time $t + \Delta_t$ the group of network members that advocate the trend a has at least $\varepsilon \cdot n$ members (namely, that $|V_a(t + \Delta_t)| \geq \varepsilon \cdot n$).*

Definition 8. *Let $\mathcal{N}_{v,a}(t)$ denote the number of friends of user v that at time t are exposing v to the trend a (namely, the number of friends of v that at time t have been exposed to the trend a and are conveying this information to v).*

Theorem 1. *For every Δ_t, $|V_a(t)|$, n, ε, the probability that at time $t + \Delta_t$ at least ϵ portion of the network's users advocate the trend a is :*

$$e^{-\varepsilon \cdot n \cdot \xi_G \cdot \xi_N^{\rho_{opt_-}}} \cdot \left(1 - \Phi \left(\sqrt{n} \cdot \frac{\varepsilon - \tilde{P_-}}{\sqrt{\tilde{P_-}(1 - \tilde{P_-})}} \right) \right)$$

$$\leq P_{Trend} \left(\Delta_t, \frac{|V_a(t)|}{n}, \varepsilon \right) \leq$$

$$e^{-\varepsilon \cdot n \cdot \xi_G \cdot \xi_N^{\rho_{opt_+}}} \cdot \left(1 - \Phi \left(\sqrt{n} \cdot \frac{\varepsilon - \tilde{P_+}}{\sqrt{\tilde{P_+}(1 - \tilde{P_+})}} \right) \right)$$

where :

$$\tilde{P_+} = e^{(\rho_{opt_+} - \Delta_t \cdot \sigma_+)} \cdot \left(\frac{\Delta_t \cdot \sigma_+}{\rho_{opt_+}} \right)^{\rho_{opt_+}}$$

$$\tilde{P_-} = e^{-\left(\frac{\Delta_t \cdot \sigma_-}{2} - \rho_{opt_-} + \frac{\rho_{opt_-}^2}{2\Delta_t \cdot \sigma_-} \right)}$$

and where :

$$\rho_{opt_-} \triangleq \underset{\rho}{\mathrm{argmin}} \left(P_{Local}^{\varepsilon \cdot n} \cdot P_{Trend} \left(\Delta_t, \frac{|V_a(t)|}{n}, \varepsilon \right) \right)$$

$$\rho_{opt_+} \triangleq \underset{\rho}{\mathrm{argmax}} \left(P_{Local}^{\varepsilon \cdot n} \cdot P_{Trend} \left(\Delta_t, \frac{|V_a(t)|}{n}, \varepsilon \right) \right)$$

and provided that $\rho_{opt_-} < \Delta_t \cdot \sigma_-$ and $\rho_{opt_+} > \Delta_t \cdot \sigma_+$, and where ξ_G denotes the network's adoption factor *and ξ_N denotes the network's* influence factor *:*

$$\xi_G = e^{-\frac{1}{n} \sum_{v \in V} s_v} \quad , \quad \xi_N = e^{-\frac{1}{n} \sum_{e(v,u) \in E} \left(\frac{w_{u,v}}{|\mathcal{N}_v|} + \frac{w_{v,u}}{|\mathcal{N}_u|} \right)}$$

Experimental results: The datasets were analyzed using the model given in (Pan, Aharony, and Pentland 2011), based on which we have experimentally calculated the values of β, ξ_G, ξ_N and σ_-.

Figures 1 and 2 demonstrate the probabilistic lower bound of Theorem 1 for trend emergence, as a function of the overall penetration of the trend at the end of the time period, under the assumption that the emerging trend was observed in 5% of the population. In other words, for any given "magnitude" of trends, what is the probability that network phenomena that are being advocated by 5% of the network, would spread to this magnitude. Notice that although longer spreading times slightly improve the penetration probability, the "maximal outreach" of trends (the maximal rate of global adoption, with sufficient probability) is dominated by the topology of the network.

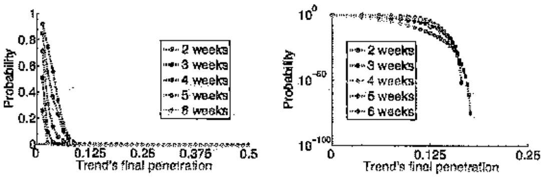


Figure 1: Trends spreading potential in the *eToro* network, for various penetration rates. Initial seed group is defined as 5% of the population. Each curve represents a different time period, from 2 weeks to 6 weeks.

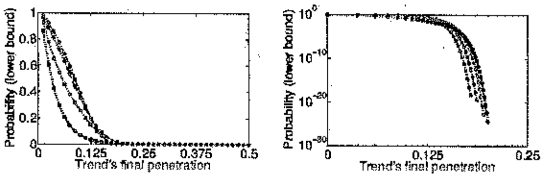


Figure 2: Trends spreading potential in the *Friends and Family* network, for various penetration rates. Initial seed group is defined as 5% of the population. Each curve represents a different time period, from 2 weeks to 5 weeks.

The upper bound of Theorem 1 can be used in two main ways. First, is provides analytic estimation regarding the success probability of various campaign goals, and can be used as an "impossibility result" for detecting unrealistic goals. This is illustrated in Figure 3, where the probability of obtaining various levels of trend penetration are given, as a function of the initial amount of resources (e.g. the size of the initial seeding group). As the desired goal become increasingly more ambitious (namely, convincing a growing number of users to adopt the trend), the success probability sharply decreases, while maintaining a monotonous connection between the size of the initial seeding group, and the success probability. Alternatively, multiplying the success probability by the size of the initial seeding group, we obtain the campaign's *"utilization function"*, analytically predicting the marginal utilization of any increase in the resources allocated to the campaign. This in turn can also be used for calculating the optimal size of the seeders group (per any penetration goal), as illustrated in in Figure 4, demonstrating how a change in the campaign goals can significantly change the optimal campaign strategy. Figure 4 is based on real-world network data of the *eToro* community. Note how the utilization function for a campaign intended to convince 30% of the community would be monotonous, while a campaign intended to convince 50% of the community would be a non-monotonous function, with a single global maximum.

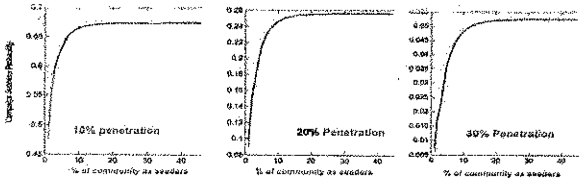


Figure 3: An illustration of Theorem 1 using the *eToro* network. Upper bound of the campaign success probability for three different goals (penetration of 10%, 20% and 30%) as a function of the size of the initial seeding group.

References

Aharony, N.; Pan, W.; Ip, C.; Khayal, I.; and Pentland, A. 2011. Social fmri: Investigating and shaping social mechanisms in the real world. *Pervasive and Mobile Computing*.

Pan, W.; Aharony, N.; and Pentland, A. 2011. Composite social network for predicting mobile apps installation. In *Proceedings of the 25th Conference on Artificial Intelligence (AAAI)*, 821 – 827.

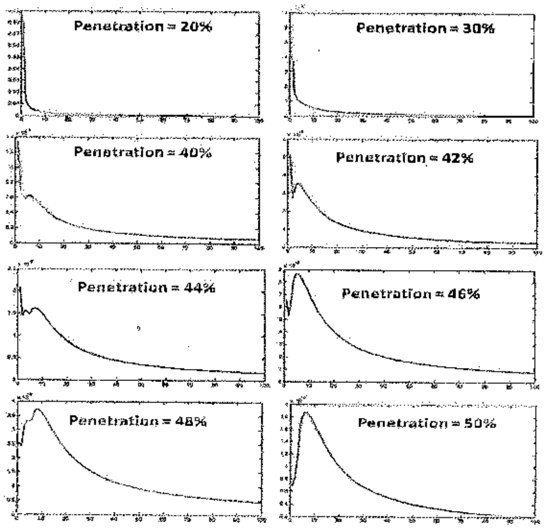


Figure 4: An illustration of Theorem 1 using the *eToro* network, analytically demonstrating the utilization function of various campaigns goals, calculated as the success probability multiplied by the size of the seeding group. Each graph corresponds to a different campaign goal (defined by the required penetration rate) and shows the marginal utilization of the resources (i.e. the seeders) as a function of the amount of resources invested (i.e. number of seeders). Notice how a monotonous utilization function for 20% suddenly gets distorted as we approach 40% penetration goal. Several local maxima appear around 42% - 48%, and at 50% a new global maximum is settled, with a non-monotonous function.

The Fallacy of Pigeonholing Global Users into Local Models

Ceren Budak Divyakant Agrawal Amr El Abbadi
Department of Computer Science UCSB
Santa Barbara, CA 93106-5110, USA
{cbudak, agrawal, amr}@cs.ucsb.edu

ABSTRACT

Recent studies on the diffusion of information in social networks have largely focused on models based on the influence of local friends. In this paper, we challenge this approach and revive earlier theories introduced by social scientists in the context of diffusion of innovations to model user behavior. To this end, we study various diffusion models in two different online social networks; Digg and Twitter. We first evaluate the applicability of two representative *local* influence models and show that the behavior of most social networks users are not captured by these *local* models. Next, driven by the theories introduced in the diffusion of innovations research, we introduce a novel diffusion model called *Gaussian Logit Curve Model (GLCM)* that models user behavior with respect to the behavior of the general population. Our analysis shows that *GLCM* captures user behavior significantly better than the *local model*, especially in the context of Digg. Aiming to capture both the *local* and *global* signals, we introduce various hybrid models and evaluate their significance through statistical methods. Our methodology models each user separately, automatically determining which users are more driven by their local relations and which users are better defined through adopter categories, therefore capturing the complexity of human behavior.

1. INTRODUCTION

The advent of online social networks has generated an ever increasing interest in understanding how and why people share information online. Such an understanding can help devise methods to control or maximize [?] the reach of an information item or even build social networks that provide the right tools for sharing information. Various models introduced in the context of online social networks, aim to explain the behavior of a given user by the behavior of his/her friends. This entails making the simplification of modeling users as being aware of *only* their immediate surrounding.

The goal of building models of diffusion is hardly new, however. Social scientists have long been building models to study the *diffusion of innovations*. Although the theory of diffusion of innovations brings up the importance of friendship relations as well, it reaches beyond that and identifies various other characteristics of human behavior that are vital to this process. For instance, it categorizes people into one of five categories, *innovators, early adopters, early majority, late majority* and the *laggards*, based on their innovativeness. This categorization is done on the basis of how early/late a person adopts an innovation *with respect to the rest of the population*. Therefore, adoption behavior is explained not solely based on *friend behavior* but based on the behavior of *the entire population*. One of the most interesting findings of this research, which was first proposed by Rogers [?] is that users belong to one of the five adopter categories *irrespective of the innovation*. This means that an innovator will lead and not follow an innovation no matter what the nature of the innovation is. Similarly a laggard is intrinsically reluctant to adopt an idea before a large population has done so.

With the advent of online social networks, researchers largely moved away from this point of view and moved on to local models where the behavior of users are explained by the events in their local neighborhood. In this work, one of our main goals is to ask the question "should the global models be tossed away so swiftly?" and make a case for re-integration of these characteristics for understanding user behavior. To this end, we first study two representative *local* models [?, ?] in two different online social networks; Digg and Twitter and show that such models are inadequate in capturing user behavior. In the context of Digg, our goal is to define the probability of voting on a story. In the case of Twitter, our goal is to capture the first time a user uses a specific hashtag. Driven by the lack of fit of the local models and inspired by research in the diffusion of innovations, we propose a novel diffusion model called *Gaussian Logit Curve Model (GLCM)* that models user behavior with respect to the entire population and captures the innovativeness of a given user based on its actions. We test the fit of *GLCM* in Twitter and Digg and show a better fit compared to the local models, especially in the case of Digg. To bridge the gap between these two perspectives, we introduce several hybrid models that incorporate the social effect as well as the global awareness. Through the use of various statistical methods and prediction, we show that the hybrid models perform significantly better than the local models while performing slightly better than *GLCM*. For instance in Twitter, prediction results show that the two local models have a value of at least 0.8 *AUC* (which corresponds to *good* prediction) for only 11% and 8% of users respectively while this number is 76% for *GLCM*, and around 84% for the hybrid models. With a per-user modeling technique where each user is captured by the best model that fits them, this number reaches 92%.

We envision modeling as capturing the characteristics of each user rather than aiming to find one parameter, or even one model, to capture all users. Through a per-user modeling approach, our methods are able to capture the behavior of every user with the model that best fits them rather than applying a one-model-fits-all technique. This goal of modeling each user as a distinct entity introduces certain challenges. Methods such as traditional logistic regression fail for various users due to small samples. Therefore, we use Firth logistic regression which provides a better fit. Overall, our approach, which uses a cross social networks analysis methodology, challenges the current single model approach and thus the generality of many commonly used social influence models. Through the testing of various models on different networks, we believe our work takes the first step for creating a general and practical framework to evaluate diffusion models.

2. LOCAL MODELS

We study the problem of progressive diffusion where the users who adopt an item become active and do not become inactive again and focus our efforts on *per-user* modeling.

Linear Friendship Model (LFM): defines the odds of adoption as a linear function of the number of *already-active* friends the agent has [?]. Therefore, the probability $p(x_{lfm})$ of activation for an agent with x_{lfm} already-active friends can be defined as [?]:

$$p(x_{lfm}) = \frac{e^{\alpha_{lfm} * x_{lfm} + \beta}}{1 + e^{\alpha_{lfm} * x_{lfm} + \beta}} \tag{1}$$

This model can be tested through logistic regression. The coefficient α_{lfm} measures *social correlation*: a large value of α_{lfm} indicates a large degree of correlation. The additive nature of this model is similar to the *independent cascade model (ICM)* and the *Linear Threshold Model (LTM)* [?].

Friend Saturation Model (FSM): The notion of additive influence of friends was challenged in a recent study [?] which claims that the effect of multiple recommendations by friends quickly saturates and is approximated as constant. The *Friend Saturation Model* [?] that addresses this notion can be formulated as another logistic regression with one categorical explanatory variable x_{fsm} ($x_{fsm} = 0$ for no active friends and $x_{fsm} = 1$ for at least one active friend).

Regression Results for *LFM* and *FSM* in Digg and Twitter: Bias reduced (Firth) logistic regression is applied to extract α_{lfm} and α_{fsm} values for 41348 users who have voted on at least 10 stories. *LFM* has statistical significance for 8257 users while this number is 7711 for *FSM*. Even though a large portion of users (5786 out of 8257) have positive social correlation, there is a group that has *negative* social correlation instead. The results for Twitter show a better fit for the local models. For the 19666 users that used at least 20 distinct hashtags, 12028 had a good fit for *LFM* (similar results for *FSM*), providing a better fit compared to the results obtained for Digg data set. The α_{lfm} values for the users that pass the likelihood ratio test are mostly greater than 1. The results overall indicate that the effect of the social is higher in Twitter compared to Digg. However, the behavior of 2/5 of the users still is not captured by the local models.

3. GAUSSIAN LOGIT CURVE MODEL

Our goal is to define the odds of adoption for a given user as a function of the size of the adoption, i.e. the number of people that have already adopted it through a gaussian logit. Such modeling for each user can help determine their optimum time to adopt an idea therefore capturing the adopter category they belong to in the community as defined in the context of diffusion of innovations. In order to approximate the fit of a gaussian distribution to the odds of adoption, we use a technique that has been used in ecology to model species response curves to presence-absence data [?]. Similar to how species are modeled to respond to an ecological variable, we model $p(x_{glcm})$, the probability of adopting an idea as a function of an environmental variable x_{glcm} that captures the user's response to the behavior of the general crowd. For this purpose, we introduce the *Gaussian Logit Curve Model (GLCM)*, in which the logit-transform of probability [?] is a quadratic function as follows:

$$log(\frac{p(x_{glcm})}{1 - p(x_{glcm})}) = b_0 + b_1 x_{glcm} + b_2 x_{glcm}^2 = a - \frac{1}{2}\frac{(x_{glcm} - \mu)^2}{\sigma^2} \tag{2}$$

where x_{glcm} is the natural logarithm of the total number of people who have adopted the idea so far, μ is the optimum point in the innovation curve for a given user to adopt the innovation. Tolerance of a user, i.e. how much the user varies from his/her mean behavior is captured by σ. A low sigma value means that the probability of adoption changes quickly as the current number of adopters in the network moves further from the mean (μ) of the given user. The parameter a is related to the maximum value of $p(x_{glcm})$. We choose x_{glcm} as the logarithm of the number of adopters since the number of adopters can be a noisy parameter as it varies largely for different stories. Using the quadratic model, the optimum mean and deviation values can be extracted from the b values as follows:

$$\mu = b_1/(2b_2) \tag{3}$$

$$\sigma = \frac{1}{\sqrt(-2b_2)} \tag{4}$$

$$p_{max} = p(\mu) = \frac{1}{1 + \exp(-b_0 - b_1\mu - b_2\mu^2)} \tag{5}$$

Regression Results for *GLCM* in Digg and Twitter: The results of the Firth logistic regression show that the behavior of a much larger population can be described by the global model (= 25050) compared to *LFM* (=8257) or *FSM* (=7711). This result suggests that the users in Digg have consistent adopter behaviors, i.e. the timing users choose to vote on a topic w.r.t. the votes from the general crowd is largely independent of the story and mostly a function of intrinsic characteristics of the user itself. Similarly, a large fraction the *Twitter* population (19484 out of 19666 users) has a good fit for *GLCM*.

4. HYBRID MODELS

Our analysis of the local and global models indicates that the actions of the local neighborhood and the entire community can have different levels of effect on the adoption behavior of a given user. However, exactly how these two notions interact is not clear. Therefore, we tested 5 different hybrid models that represent different interactions between the local and global signals. Due to space limitations, here we introduce only one of these hybrid models:

LFMorGLCM is an additive model where the probability of adoption depends on both the global and the local signals. For this model, the logit can be defines as: $logit(p) = b_0 + b_1 x_{glcm} + b_2 x_{glcm}^2 + b_3 x_{lfm}$ where x_{lfm} is the natural logarithm of the number of active friends and x_{glcm} is the logarithm of the total number of active users. Our analysis shows that 13990 Digg users pass the likelihood ratio test for this model while this number is 13672 for Twitter.

5. MODEL EVALUATION

In order to capture the predictiveness of the models, we perform 20 different experiments for each user where the data is randomly divided into test and training sets. Training data consists of 80% of the data points while the rest is marked as test data. Since a user adopting an item is a rare event, simple methods such as precision are not applicable as an evaluation metric. Therefore, the *receiver operating characteristic (ROC)* is used to evaluate the models. The result of *ROC* analysis can be summarized in one value; the *area under the ROC (AUC)*. *AUC* measures the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one. An *AUC* value between 0.5-0.6, 0.6-0.7, 0.7-0.8, 0.8-0.9 and 0.9-1 can be evaluated as *fail, poor, fair, good* and *excellent* respectively.

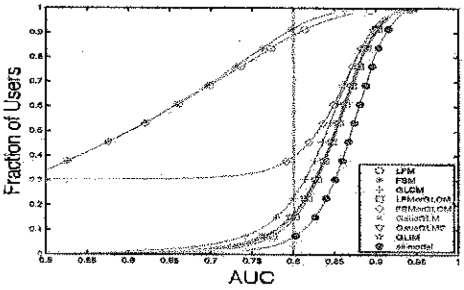


Figure 1: Area under ROC results for cross validation in Twitter

The results for Twitter are provided in Figure 1. Each curve in this figure corresponds to the CDF distribution of AUC values for a given model across all the users. For instance a point (x,y) on a given curve means that *y* fraction of users have *AUC* value of less than or equal to *x* for that given model. Therefore, a curve that is closer to the *x*-axis is preferred. It is obvious that the local models *LFM* and *FSM* consistently perform worse than the global and the hybrid models. Of the hybrid models, *FSMorGLCM* performs the worse while the others have comparable performance. For Digg, *GLCM* performs very similar to the hybrid models while in Twitter we see some improvement in performance by using one of the hybrid models (except *FSMorGLCM*). These results show that a hybrid model alone provides "some" added value but the difference is not drastic.

Capturing the "best" model through user-centric modeling: One of the advantages of performing user-centric modeling is that it is not necessary to find one model to fit all users. If for each user, the best model is chosen based on their *AUC* values, the fit can be improved as given by the curves named *all-model* in Figure 1. For instance, consider the performance of the models in the context of Twitter. The vertical line at $x = 0.8$ shows that the fraction of users for which the models *LFM, FSM, GLCM, LFMorGLCM, FSMorGLCM, GausGLM, GausGLM2* and *GLIM* perform at least good (x=0.8) on average are 0.1107, 0.0828, 0.7680, 0.8568, 0.6009, 0.8393, 0.8354 and 0.8309 respectively. However, if there is no restriction to use one model to fit all users and rather choose the best model per user, this value reaches 0.9294 as given in the curve named *all-model*. These results demonstrate the importance of capturing each user as a separate entity with distinct intrinsic characteristics.

6. CONCLUSION

In this paper we studied the diffusion of information in two different networks, Digg and Twitter and investigated the validity of commonly used local influence models. We studied two different local models to address this question. The results indicate that Digg users show little social correlation while Twitter has overall more social correlation. Due to the large number of users whose behavior is not captured by the local model, we introduced a novel diffusion model called *Gaussian Logit Curve Model* that is inspired by the research in the theory of diffusion of innovations. *GLCM* captures a user's global behavior, i.e. how fast or slow they vote on stories in Digg or they use hashtags in Twitter compared to the rest of the social network community. *GLCM* gives significant improvement over the local models for Digg while the results for Twitter, although still impressive, are not as strong as the results for Digg. In order to capture both the local and the global signals we introduce a number of hybrid models. Prediction results show that the hybrid models provide a great improvement over the local models while the improvement is not as pronounced for *GLCM*.

The results presented in this paper also raise the higher level question for our research community: "How generalizable are the social influence models that have been used in the area of social networks? Are we, as a community, too eager to look for social effects in any online tool that simply has a friend button?" Given the outcomes of this study, there are various future research directions. As future work, we aim to investigate the effectiveness of maximization of diffusion techniques introduced in the context of purely local models and to construct optimal methods for the *per-user* modeling technique. It is also important to investigate new and efficient techniques for predicting the future popularity of information items based on the new models. Finally, through a cross-network, cross-model evaluation framework our work takes the first step for a general framework to study models of diffusion. We believe it is an important task to enrich this framework by studying other social networks and diffusion models.

Apple vs. Android: a comparison of social effects in adoption

Johannes Bjelland, Geoffrey Canright, Kenth Engø-Monsen, Pål Roe Sundsøy
Telenor Research and Future Studies
Telenor ASA, Oslo, Norway
johannes.bjelland, geoffrey.canright, kenth.engo-monsen, pal-roe.sundsoy@telenor.com

Rich S. Ling
IT University / Telenor ASA
Copenhagen, Denmark / Oslo, Norway
rili@itu.

I. INTRODUCTION

It has long been known among marketers that our social network matters when we make purchasing decisions, and that having positive word of mouth about a product can be a key to success; see e.g. [1] for a review of studies on social networks within marketing. Traditionally, data on social networks have been difficult to collect, but in recent years researchers have gained access to massive social network data from e.g. online instant messaging services [8][5] and phone log data [2][4][3][6][9]. Such data has made it possible to study e.g. social churn [3], service uptake [2] among telecom customers, and product adoption on an Instant Messaging network [8].These studies confirm that consumer behavior is dependent on the communication network. We have in a recent study [6] shown how the structure of the *adopter network*—the social network of adopters—develops over time, and how social spreading can be measured by studying this network. In this paper, we do a comparative study of social spreading effects for two competing types of smartphones - the Apple iPhone, and smartphones based on Google's Android OS.

II. METHOD

Our social network is built by collecting anonymized call data records, aggregated over a 3-month period, and then using the communication links (voice and sms) as proxy for the social relationships. Other studies have shown that mobile phone activity is a good way to measure real social relationships [3]. To remove error sources due to 'non-personal' relationships, we have applied some filtering of the dataset, regarding extreme outlier nodes (based on combinations of extreme usage and degree) as machines, and removing them. Only traffic between Telenor customers is used; calls to other operators are excluded. For this study we have also included weak links –relations with limited SMS/voice traffic in the period. In total we end up with a network containing around 2.5 million nodes and 45 million edges.

We also use handset type data to associate a handset type with each node in the social network. With these data we can define the 'adoption network' – the social network among adopters [6]. This is simply the sub network consisting of adopters and their common links. We can then study the development of the adoption network for iPhones (viewed as a single 'product') [6]—and for Android phones, over time (again making no distinction among the various models of Android phones). These same data allow us to measure conditional adoption probabilities between neighbors on the network, which we use as an indicator of social effects. Finally, we use postcode information—very coarse-grained geographic information—on subscribers to map smartphone adoption to geographical areas in Norway.

III. RESULTS

In a previous paper [6], we looked at the growth of the iPhone adoption network over time, showing clearly the development of a 'social monster'—a giant connected component of the adoption network which shows the fastest growth. We equated the strength of this monster with the presence of iPhone adopters in the 'dense core' of highly central subscribers—a sign of success of the product in taking off.

In Figure 1, we compare the growth of the Apple adoption network with that of the Android adoption network, on a quarterly basis. In each case, we start with the quarter in which the 'product' was first launched. While we see no dramatic difference in the first–quarter picture (Fig 1(a)), it is clear that already, two quarters later (Fig 1(c)), the Apple 'monster' (Largest Connected Component - LCC) is growing much more rapidly than the Android monster. This holds not only for total number of adopters in the LCC, but also in terms of their percentage of all adopters: two quarters after launch, the Apple LCC has ca 38% of all adopters, while the Android LCC has around 28%.

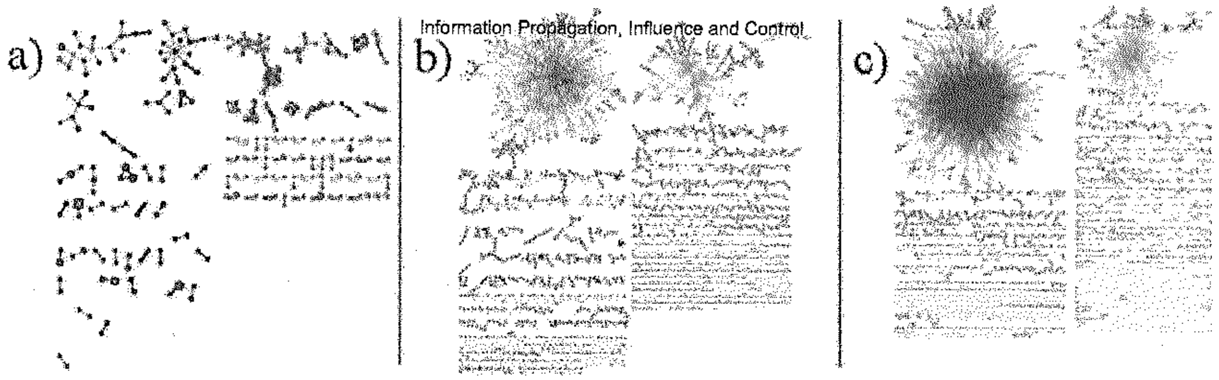


Figure 1
The figure shows the evolution of the iPhone (Red/Left) and Android (Blue/Right) adoption networks during the 3 first quarters after launch of the first respective brands. The nodes are customer with iPhone(red) and Android (blue). Links indicate communication between the nodes. Figure a) is the quarter when the handset first appears in the market, b) is the next quarter and c) is third quarter after product launch. Isolated nodes are not shown – i.e iPhone customers that do not know other iPhone buyers or Android Customers that do not call other Android customers will not appear in this visualization.

For another indicator of social adoption, we look at the number of inter-adopter links (adoption pairs) in each adoption network, over time. Figure 2 tracks the number of adoption pairs for each product, versus the total number of adopters. The black dotted curve in Figure 2 gives the number of adopter pairs expected, for the given total number of adopter pairs on the fixed call network, if adoption was purely random. We see that both products generate many times the number of adopter pairs expected from this random reference model. Thus, both products show significant social adoption—but, again, the effect is clearly weaker for Android. (The ratio between the empirical number of adopters, and that number found in the random reference model, was studied in Ref. [6] and termed 'kappa'.)

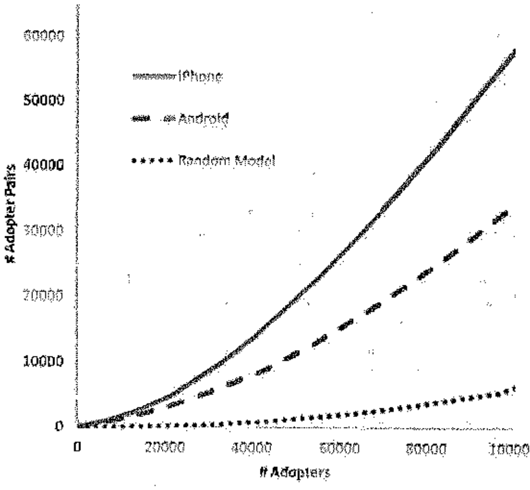


Figure 2
The plot shows the number of adoption pairs (Connected customers adopting same type of handset) vs. the total number of customers having the brand (x-axis). Red solid line is iPhone, blue stippled line is Android and black dotted line is the random simulation model.

In Figure 3 we plot yet another indicator of social adoption. Here we look at $p_X(k)$ —the conditional probability that, given that a node has k neighbors adopting product X, the node in question has also adopted product X.

Since random adoption gives a flat $p_X(k)$, the positive slope of the empirical results in Figure 3 are again taken as evidence for social effects (of some kind) in adoption—for both products. The difference between Apple and Android is seen here in that the Android curve has more weight at small k—flattening out at large k—while the Apple curve has less weight at small k, but grows more steeply, and more nearly linearly, with increasing *k*. These data were taken in Q3/2011. In this period, the Apple and Android penetration were approximately equal (around 18% each). Hence we see that p(k) is *under*represented at small k (compared to the random case, ie, a flat line at p(k) = 18%), and *over*represented at large k, for both products—but the skew is greater for Apple than for Android. Taking this skew as an indicator of social adoption, we find again that Apple is 'more social' than Android.

We next perform a numerical experiment, in which we merge the two adoption networks—forming a 'smartphone adoption network'—and then scramble the placement of the iPhones and Android phones randomly among the smartphone users in this network. The result is a simple network model of a 'brand-indifferent' smartphone user. What we find, remarkably enough, is that $p_X(k)$ is strongly affected by this re-scrambling for iPhone users, but *essentially unchanged* for Android users (compare the solid green curve if Fig 3 with the green triangles). That is, by this measure, Android users do not care (statistically) whether their contacts use Android or iPhone—while iPhone users certainly do have a preference.

We can also see this by simply taking the total number of friends using each type of smartphone, for each type of user. We find that the average Apple user had over <u>two times</u> as many Apple friends as statistically expected from no preference—while all other results (number of Apple friends of Android users, and number of Android friends of Apple and Android users) were statistically consistent with no

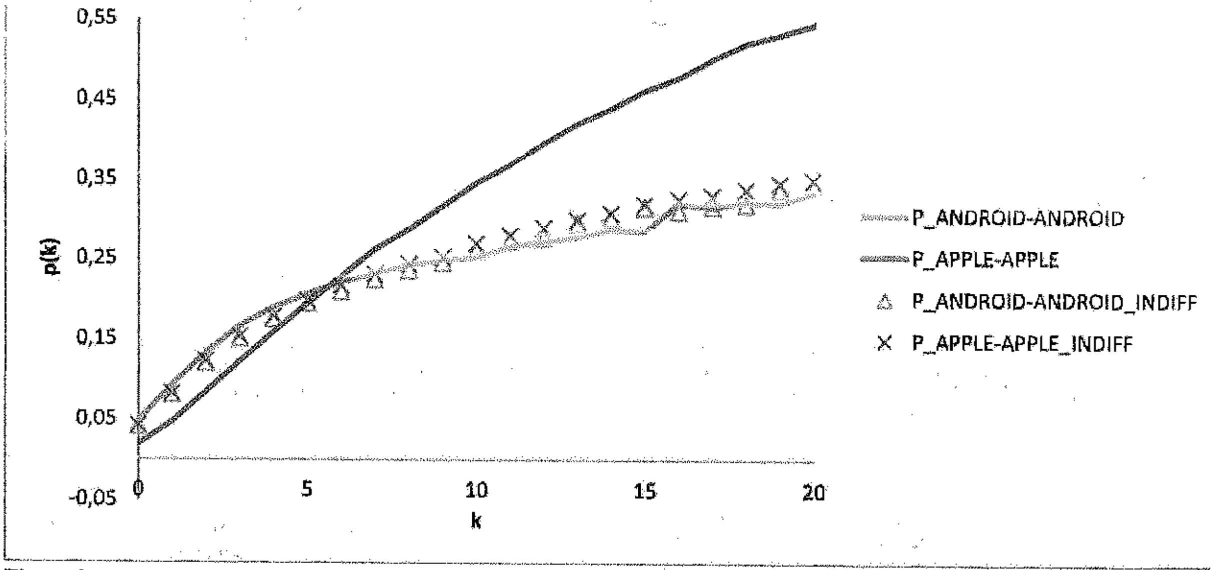


Figure 3

The plot shows ego's handset adoption probability, given k contacts with same handset. Red solid line is the empirical iPhone adoption probability, and green solid line is the same for Android. Results are also shown for numerical experiments in which the brands chosen are randomly scrambled—keeping the total number of each constant—among smartphone adopters (the 'brand-indifferent' model). Green triangles and purple x's show, respectively, the Android and Apple pairs found when users are randomly scrambled in this 'brand-indifferent' way over the smartphone network. We note that the true empirical distribution of Android-Android pairs is almost identical to that seen for the 'brand-indifferent' model—while the empirical Apple results show a strong bias towards Apple adoption with many Apple-adopting friends.

Apple/Android preference. In short: restricted to smartphone users, we again find that Apple users have more friends, and a stronger preference for their 'own kind'. Finally, when we drop the restriction to smartphone users, we still find that iPhone users have more contacts on the call graph (our proxy social network) than do Android users—averaging around 45 for the former, compared to around 33 for the latter.

Inspired by these results, we have examined the geographic distributions of these two products. Our method here was to first aggregate Apple and Android adoption totals over Norwegian postal codes, and then take the ratio of the two. What we find (not shown here) is that Apple is dominating in Norway's cities. Since these results are also from Q3/2011, there are roughly equal numbers of Apple and Android phones—so that Apple cannot win everywhere. Thus we see a rather stark urban/rural dichotomy, with Apple dominating the cities, and Android turning up dominant in scattered spots in the countryside.

We conjecture (but have not yet tested) that the high-centrality users (as measured by eigenvector centrality) are concentrated geographically in the cities (just as they are concentrated, by definition, in the dense core of the social network). In any case, all of the above results give a picture of Apple users as being more attracted to other Apple users than are Android users to other Android users—but also, more social in general.

[1] C. Van den Bulte and S. Wuyts,"Social Netwurks and Marketing", Marketing Science Institue 2007

[2] S. Hill, F.Provost and C. Volinsky, "Network-Based Marketing: Identifying Likely Adopters via Consumer Networks", Statistical Science. 2006, Vol. 21, No. 2, 256-276

[3] Dasgupta, K., Singh, R., Viswanathan, B., Chakraborty, D., Mukherjea, S., Nanavati, A. A., and Joshi, A. 2008. "Social ties and their relevance to churn in mobile telecom networks". In Proceedings of the 11th international Conference on Extending Database Technology: Advances in Database Technology (Nantes, France, March 25 - 29, 2008). EDBT '08, vol. 261. ACM, New York, NY, 668-677.

[4] J.-P. Onnela,J. Saramäki,J. Hyvönen, G. Szabó,D. Lazer,K. Kaski,J. Kertész and A.-L. Barabási,"Structure and tie strengths in mobile communication networks." Proc Natl Acad Sci U S A. 2007 May 1; 104(18): 7332–7336.

[5] Sinan Aral, Lev Muchnik, and Arun Sundararajan. "Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks." Proceedings of the National Academy of Sciences, 106(51):21544–21549, December 2009.

[6] P Sundsøy, J Bjelland, G Canright, K Engø-Monsen, R Ling, "Product adoption networks and their growth in a large mobile phone network", IEEE Advanced in Social Network Analysis and Mining (ASONAM 2010).

[7] N Eagle, A. Pentland, D Lazer, P Alex "Inferring friendshop network structure by using mobile phone data", National Academy of Sciences 106.36 (2009) 15274-15278, 2009.

[8] Rushi Bhatt, Vineet Chaoji, and Rajesh Parekh. 2010. Predicting product adoption in large-scale social networks. In *Proceedings of the 19th ACM international conference on Information and knowledge management* (CIKM '10). ACM, New York, NY, USA, 1039-1048.

[9] P. Sundsøy, J. Bjelland, K. Engø-Monsen, G. Canright, R. Ling, "Comparing and visualizing the social spreading of products on a large-scale social network", To appear in "The influence of Technology on Social Network Analysis and Mining, Tanzel Ozyer et.al (Springer 2012).

Connectivity & Collective Action in Social Networks

Navid Hassanpour, Ji Liu, Sekhar Tatikonda*.

August 20, 2012

Historical examples of mass political action such as social revolutions are often cascades of collective action resulting from local mobilization. In this study, we examine the possibility of universal action as a result of local learning and other-regarding decisions to act. We propose a dynamic variation of the threshold model (Granovetter 1978) in social networks. Similar to Gould (1993), Chwe (1999), Centola et al. (2007), and Watts (2002) we study the relation between network structure and the diffusion of collective action. In addition to structural conditions, such as diffusion asymptotics, we examine the difference between two classes of learning: observational and communicative (Acemoglu and Ozdaglar 2011). Hence we examine the possibility of cascades–or partial diffusion–based on two parameters: network connectivity and learning mechanism. In particular, we want to link levels of graph connectivity to the possibility of action cascades.

Intertwined Learning and Action Dynamics: The model we propose is a combination of action and threshold dynamics. At each time unit, actors update their thresholds based on a weighted average of their own and their neighbors' thresholds; this is known as the the DeGroot learning model (Golub and Jackson 2007). They also decide on taking action or abstaining based on the comparison between their personal thresholds and the proportion of their social network neighbors who acted in the previous round.

The existing strategic models of collective action, such as global games, condition decisions on noisy representations of the situation, e.g. power of the government (Dahleh et al. 2012). We instead take the thresholds to be the citizens' assessment of the government power: the higher they are, the less likely the person is to act against the status quo. This self-sufficient, citizen centric definition of thresholds relies on the assumption that regardless of material might, governing is not sustainable when citizens stop to believe in the legitimacy of power. Our model also decouples threshold dynamics from action dynamics, hence allowing for a separate examination of private beliefs–reflected in thresholds–and public actions, similar to the dissonance studied in Kuran (1989). We will show that the actions can fluctuate while beliefs are hardened and vice versa.

Asymptotics & Transient Behavior: In the first section of the project, we examine the dynamic equations for actions and thresholds in the case of canonical network configurations such as fully connected, star, ring, and infinite path topologies. For a number of these cases, we find conditions under which cascades of collective action are possible. When not, we find the size of islands of action defined as the "radius of diffusion."

Emphasis on Radius of Diffusion: We examine the minimum number of radicals needed to ignite a cascade or generate a stable island of action. We show that observational learning can not help diffusion much: communication is needed to ease the spread of action. When radicals, i.e. agents with low thresholds, are in minority, highly connected graphs are often detrimental

---

*The authors are affiliated with Yale University: `navid.hassanpour@yale.edu`, `ji.liu@yale.edu`, `sekhar.tatikonda@yale.edu`

to cascades, see figure (1). In such situations we demonstrate that adding links does not help collective action. This is in line with the conclusion in Hassanpour (2012) on the consequences of media shut-down during the Egyptian uprising of 2011.

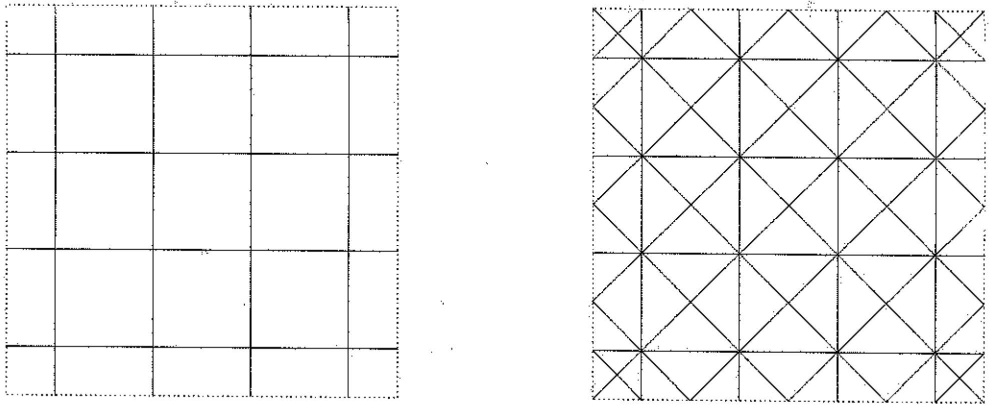


Figure 1: When learning is observational, islands of activity are possible on the simpler grid, while they are not sustainable on the more connected pattern on the right. Connectivity is not always helpful. In both networks, agents are located at the intersection of horizontal and vertical lines.

References

Acemoglu, D. and Ozdaglar, A. (2011), 'Opinion dynamics and learning in social networks', *Dynamic Games and Applications* 1(1), 3–49.

Centola, D., Eguiluz, V. M. and Macy, M. W. (2007), 'Cascade dynamics of complex propagation', *Physica A* **374**, 449–456.

Chwe, M. (1999), 'Structure and strategy in collective action', *American Journal of Sociology* **105**(1), 128–156.

Dahleh, M., Tahbaz-Salehi, A., Tsitsiklis, J. N. and Zoumpoulis, S. I. (2012), 'On global games in social networks of information exchange', *working paper* .

Golub, B. and Jackson, M. O. (2007), 'Naive learning in social networks and the wisdom of crowds', *American Economic Journal: Microeconomics* .

Gould, R. V. (1993), 'Collective action and network structure', *American Sociological Review* **58**(2), 182–196.

Granovetter, M. S. (1978), 'Threshold models of collective behavior', *The American Journal of Sociology* **83**(6), 1420–1443.

Hassanpour, N. (2012), 'Media disruption and revolutionary unrest: Evidence from mubaraks quasi-experiment', *Political Communications* .

Kuran, T. (1989), 'Sparks and prairie fires: A theory of unanticipated political revolution', *Public Choice* **61**(00), 41–74.

Watts, D. J. (2002), 'A simple model of global cascades on random networks', *Proceedings of the National Academy of Sciences* **99**(9), 5766–71.

An Analysis of One-Dimensional Schelling Segregation

Christina Brandt* Nicole Immorlica† Gautam Kamath‡ Robert Kleinberg§

Introduction

In 1969, economist Thomas Schelling introduced a landmark model of racial segregation. Elegantly simple and easy to simulate, it provided a persuasive explanation of an unintuitive result: that local behavior can cause global effects that are undesired by all [4]. We investigate his model in an attempt to better understand the dynamics which result in residential segregation.

In Schelling's model, individuals of two races, denoted x and o, are placed near one another. This represents a mixed-race city, where individuals of different races live in close proximity. Individuals are satisfied if at least a fraction τ of the other agents in a small local neighborhood around them are of the same type. Unhappy agents can move locations, either by inserting themselves into new positions or exchanging locations with other agents. In Schelling's experiments, he found that on average, an individual i with $\tau = \frac{1}{2}$ ended up in a significantly more segregated neighborhood with approximately 80% of i's neighbors having i's type.

However, it is surprisingly difficult to analytically prove or even rigorously define the segregation phenomenon observed qualitatively in simulations. Much of this difficulty lies in the fact that the dynamics may converge to a variety of states (complete segregation, complete integration, and various partially segregated states), and so the underlying Markov chain does not have a well defined unique stationary distribution. Prior work [7, 8, 9, 10] circumvents this difficulty by introducing perturbations in the dynamics, allowing individuals to perform detrimental actions with vanishingly small probability. The research then analyzes the degree of segregation in stochastically stable states, finding generally that as time approaches infinity, complete segregation is inevitable.

We instead analyze the one-dimensional segregation dynamics directly, providing the first rigorous analysis of an unperturbed Schelling model. Our model considers a society of n individuals arranged in a ring network. We start from a random initial configuration in which each individual is assigned type x or o independently and uniformly at random. We parameterize the neighborhood size in the Schelling dynamics by w. An individual is said to be *happy* if at least w of his $2w$ nearest

---

[*] Department of Computer Science, Stanford University. Supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1147470. Email: `cbrandt@stanford.edu`

[†] Department of Electrical Engineering and Computer Science, Northwestern University. Supported in part by NSF awards CCF-1055020 and SMA-1019169, a Microsoft New Faculty Fellowship, and an Alfred P. Sloan Foundation Fellowship. Email: `nickle@eecs.northwestern.edu`

[‡] Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology. Supported by an MIT Akamai Presidential Graduate Fellowship. Email: `g@csail.mit.edu`

[§] Department of Computer Science, Cornell University. Supported in part by NSF awards CCF-0643934 and AF-0910940, AFOSR grant FA9550-09-1-0100, a Microsoft Research New Faculty Fellowship, an Alfred P. Sloan Foundation Fellowship, and a Google Research Grant. Email: `rdk@cs.cornell.edu`

neighbors are of the same type as him and *unhappy* otherwise. We consider the model of dynamics under which random pairs of unhappy individuals of opposite types trade places in each time step.

Methods and Results

In stark contrast to previous analyses of approximate dynamics, we prove that once the dynamics converge, most individuals reside in nearly integrated neighborhoods; that is, the average run length in the final configuration is independent of n and only polynomial in w. Thus, contrary to the established intuition, the local dynamics of Schelling's model do not induce global segregation in proportion to the size of the society but rather induce only a small degree of local segregation.

We observe that a sequence of at least $w+1$ consecutive individuals of the same type is stable. Every individual in this configuration is happy, so they will never move away. We name this type of configuration a *firewall*. It can be shown that, with high probability, all happy individuals in the final configuration will be in a firewall. Our analysis centers around showing that there will be many firewalls, thus providing an upper bound on the length of any individual firewall. In particular, if we have a firewall of one type, but know that it is between two nearby firewalls of the opposite type, this limits how large the firewall may grow.

In order to track the production of firewalls, we analyze how configurations known as *firewall incubators* develop over time. A firewall incubator is a segment of length $O(w)$ which has a strong imbalance in the types of the contained individuals at the beginning of the process. Intuitively, we'd expect that these incubators are likely to become firewalls. By applying the central limit theorem, we show that a segment is a firewall incubator with constant probability.

We then examine how the dynamics affect a firewall incubator over time. We would like that an incubator experiences enough "good" swaps to become a firewall, and does not undergo enough "bad" swaps to lose its initial imbalance in types. Using a probabilistic lemma known as the Ballot Theorem [1, 2, 3], we show that a firewall incubator becomes a firewall with probability inversely proportional to w.

This reasoning depends on a key assumption: the proportion of unhappy individuals of each type is close to balanced for the majority of the process. We justify this through an application of Wormald's theorem [5, 6], which allows us, under suitable technical conditions, to approximate a discrete-time stochastic process with a continuous-time differential equation. This is technically non-trivial due to complications with infinite differential equations.

Together, these steps allow us to conclude our main result,

Theorem 1. *Consider the segregation process with window size w on a ring network of size n, starting from a uniformly random initial configuration. There exists a constant $c < 1$ and a function $n_0 : \mathbb{N} \to \mathbb{N}$ such that for all w and all $n \geq n_0(w)$, with probability $1 - o(1)$, the process reaches a configuration after finitely many steps in which no further swaps are possible. The average run length in this final configuration is $O(w^2)$. In fact, the distribution of runlengths in the final configuration is such that for all $\lambda > 0$, the probability of a randomly selected node belonging to a run of length greater than λw^2 is bounded above by c^λ.*

References

[1] D. André. Solution directe du problème résolu par M. Bertrand. *Comptes Rendus de l'Académie des Sciences, Paris*, 105:436–437, 1887.

[2] J. Bertrand. Solution d'un problème. *Comptes Rendus de l'Académie des Sciences, Paris*, 105:369, 1887.

[3] M. Renault. Four proofs of the ballot theorem. *Mathematics Magazine*, 80(5):345–352, 2007.

[4] T. Schelling. Models of segregation. *The American Economic Review*, pages 488–493, 1969.

[5] N. C. Wormald. Differential equations for random processes and random graphs. *Annals of Applied Probability*, 5:1217–1235, 1995.

[6] N. C. Wormald. The differential equation method for random graph processes and greedy algorithms. In M. Karoński and H. J. Prömel, editors, *Lectures on Approximation and Randomized Algorithms*, pages 73–155. PWN, 1999.

[7] H. P. Young. *Individual strategy and social structure: an evolutionary theory of institutions.* Princeton University Press, Jan. 2001.

[8] J. Zhang. A dynamic model of residential segregation. *Journal of Mathematical Sociology*, 28(3):147–170, 2004.

[9] J. Zhang. Residential segregation in an all-integrationist world. *Journal of Economic Behavior & Organization*, 54(4):533–550, 2004.

[10] J. Zhang. Tipping and residential segregation: A unified Schelling model. *Journal of Regional Science*, 51:167–193, Feb. 2011.

On Bitcoin and Red Balloons *

Moshe Babaioff[†] Shahar Dobzinski[‡] Sigal Oren [§] Aviv Zohar[¶]

Many large decentralized systems rely on information propagation to ensure their proper function. We examine a common scenario in which only participants that are aware of the information can compete for some reward, and thus informed participants have an incentive *not* to propagate information to others. One recent example in which such tension arises is the 2009 DARPA Network Challenge (finding red balloons). We focus on another prominent example: Bitcoin, a decentralized electronic currency system.

Bitcoin represents a radical new approach to monetary systems. It has been getting a large amount of public attention over the last year, both in policy discussions and in the popular press [3, 6]. Its cryptographic fundamentals have largely held up even as its usage has become increasingly widespread. We find, however, that it exhibits a fundamental problem of a different nature, based on how its incentives are structured. We propose a modification to the protocol that can eliminate this problem.

Bitcoin relies on a peer-to-peer network to track transactions that are performed with the currency. For this purpose, every transaction a node learns about should be transmitted to its neighbors in the network. As the protocol is currently defined and implemented, it does not provide an incentive for nodes to broadcast transactions they are aware of. In fact, it provides an incentive not to do so. Our solution is to augment the protocol with a scheme that rewards information propagation. Since clones are easy to create in the Bitcoin system, an important feature of our scheme is Sybil-proofness.

We show that our proposed scheme succeeds in setting the correct incentives, that it is Sybil-proof, and that it requires only a small payment overhead, all this is achieved with iterated elimination of dominated strategies. We complement this result by showing that there are no reward schemes in which information propagation and no self-cloning is a dominant strategy.

Introduction

In 2009, DARPA announced the DARPA Network Challenge, in which participants competed to find ten red weather balloons that were placed at various locations across the United States [2]. Faced with the daunting task of locating balloons spread across a wide geographical area, participating teams attempted to recruit individuals from across the country to help. The winning team from MIT, incentivized balloon hunters by offering them rewards of \$2000 per balloon they locate [5]. Recognizing that notifying individuals from all over the US about these rewards is itself a difficult undertaking, the MIT team cleverly offered additional rewards of \$1000 to a person who directly recruits a balloon finder, a reward of \$500 to his recruiter, and so on. These additional payments created the incentive for participants to spread the word about MIT's offer of rewards and were instrumental in the team's success. In fact, the additional rewards are necessary: each additional balloon hunter competes with the participants in his vicinity, and reduces their chances of getting the reward.

MIT's scheme still requires further improvement. As it is, a participant can create a fake identity, invite the fake identity to participate, and use that identity to recruit others. This Sybil attack increases the participant's reward by 50%. Reward schemes should be resistant to such attacks.

A related setting is a raffle, in which people purchase numbered tickets in hopes of winning some luxurious prize. Each ticket has the same probability of winning, and the prize is always allocated. As more tickets are sold, the winning probability of a specific ticket decreases. In this case again, there is a clear tension between the organizer of the raffle, who wants as many people to find out about the raffle, and the participants who have already purchased tickets and want to increase their individual chances of winning. The lesson here is simple, to make raffles more successful participants should be incentivized to spread the word. One example of a

---

[*] This is an extended abstract. The full version has appeared in EC'12 and is available from the authors' websites.

[†] Microsoft Research, Silicon Valley. moshe@microsoft.com

[‡] Department of Computer Science, Cornell University. shahar@cs.cornell.edu

[§] Department of Computer Science, Cornell University. sigal@cs.cornell.edu

[¶] Microsoft Research, Silicon Valley. avivz@microsoft.com raffle already implementing this is Expedia's "FriendTrips" in which the more friends you recruit the bigger your probability of winning.

Our goal is to design reward schemes that incentivize *information propagation* and counter the dis-incentive arising from the competition from other nodes, and are *Sybil proof* while having a *low overhead* (a total reward that is not too high). In particular, we identify the need for such incentives in the Bitcoin protocol, our main example for the rest of this letter. First, we introduce Bitcoin and explain where the incentive problem shows up.

Bitcoin. Bitcoin is a decentralized electronic currency system proposed by Satoshi Nakamoto[1] in 2008 as an alternative to current government-backed currencies [4]. Bitcoin has been actively running since 2009, and has been getting a large amount of public attention over the last year. It represents a radical new approach to monetary systems which has appeared in policy discussions and in the popular press. Its cryptographic fundamentals have largely held up even as its usage has become increasingly widespread.

Bitcoin's appeal lies mainly in the ability to quickly transfer money over the internet, and in its relatively low transaction fees.[2] As of September 2012, there are 10 milliou units of currency in circulation (called *Bitcoins*) which are traded at a value of approximately 11 USD per bitcoin.

Bitcoin relies on a peer-to-peer network to verify and authorize all transactions that are performed with the currency. Transactions are cryptographically signed by the owner of the bitcoins that wishes to transfer them, and are sent to nodes in the peer-to-peer network for authorization. Each node in the network is supposed to propagate the transaction to its neighbors. Upon receiving a transaction, each node verifies that it is properly signed by the bitcoins' owner, and then tries to "authorize" the transaction by attempting to solve a computationally hard problem (basically inverting a hash function). This authorization process is a key ingredient in maintaining Bitcoin's security (refer to [4] for details). Once a node successfully authorizes a transaction, it sends the "proof" (the inverted hash) to all of its neighbors. They in turn, send the "proof" to all of their neighbors and so on. Finally, all nodes in the network "agree" that the transaction has taken place and was authorized.

In compensation for their efforts, nodes are offered a payment in bitcoins for successful authorizations. The system is currently in its initial stages, in which nodes are paid a predetermined amount of bitcoins that are created "out of thin air". This also slowly builds up the bitcoins supply. But Bitcoin's protocol specifies an exponentially decreasing rate of money creation that effectively sets a cap on the total number of bitcoins that will be in circulation. As this payment to nodes is slowly phased out, bitcoin owners that want their transactions approved are supposed to pay fees to the authorizing nodes.

This is where the incentive problem manifests itself. A node in the network has an incentive to keep the knowledge of any transaction that offers a fee for itself, as any other node that becomes aware of the transaction will compete to authorize the transaction first and claim the associated fee.

The consequences of such behavior may be devastating: as only a single node in the network works to authorize each transaction, authorization is expected to take a very long time.

We stress that false identities are a prominent concern in Bitcoin. In fact, the Bitcoin protocol is built around the assumption that nodes can create false identities, thus, for a transaction to be approved, nodes that control a majority of the CPU power in the network should accept it, rather than just a majority of the nodes. The latter is vulnerable to Sybil attacks. Therefore any reward scheme for transaction distribution must discourage such attacks.

The Model

We present our model for information propagation in Bitcoin's authorization protocol. For a more detailed presentation, refer to [1].

We assume for simplicity that the network consists of a forest of complete d-ary directed trees, each of them of height H. We model the authorization process of a single transaction in two phases: a distribution phase and a computation phase.

In the beginning of the *distribution phase* the buyer sends the details of the transaction to the t roots of the trees (which we term *seeds*). Each node v that is aware of the transaction can send the information to any of its children, before sending to any child it can add any number of fake identities. All of v's fake identities are connected to the same set of children. A node can condition its behavior only on the *length* of the referral chain above it, which can possibly include false identities that were produced by its ancestors.

---

[1] The name Satoshi Nakamoto appears to be an alias. The real identity of Bitcoin's creator remains a mystery.

[2] There are additional properties that some consider as benefits: Bitcoins are not controlled by any government, and its supply will eventually be fixed. Additionally, it offers some degree of anonymity.

In the *computation phase* each node that is aware of the transaction tries to authorize it. If there are k such nodes, each of them has the same probability of $\frac{1}{k}$ to authorize it first. We assume that there is a minimal payment for authorization, normalized to 1, which is necessary to motivate the nodes to work on authorizing the transaction.

When a node succeeds in authorizing a transaction we can reward nodes on the chain (starting at some seed) to that node. This chain may contain false identities as well, but cryptographic tools ensure that no node can remove its ancestors from the chain.

Reward Schemes

We suggest a rewarding scheme family called the $(\beta, \mathcal{H})$-almost-uniform family, we then combine schemes from this family to create a hybrid scheme that possesses better qualities.

$(\beta, \mathcal{H})$-Almost-Uniform Schemes. The rewards of schemes in this family are defined as follows: Suppose that a node v has authorized the transaction, and has a chain of l nodes through which it has received the transaction. If $l > \mathcal{H}$ no node is rewarded (so nodes "far" from the seed do not attempt to authorize the transaction). Otherwise, each node in the chain except v gets a reward of β, and v gets a reward of $1 + (\mathcal{H} - l + 1)\beta$. Hence, the total rewards that the scheme allocates is $\beta \cdot \mathcal{H} + 1$.

Given that there are $\Omega(\beta^{-1})$ seeds, the $(\beta, \mathcal{H})$-almost-uniform scheme creates the incentives for each node to propagate information to all its children without duplicating itself. Specifically, we show:

Theorem. *If there are $\Omega(\beta^{-1})$ seeds, the $(\beta, \mathcal{H})$-almost-uniform scheme guarantees that only strategy profiles that exhibit information propagation and no duplication up till a height of $\mathcal{H}$ survive every order of iterated removal of dominated strategies.*

This gives us two interesting schemes, for two different values of β that offer tradeoffs between the total payment and the number of seeds that need to be initially notified. The first scheme the $(1, \mathcal{H})$-almost-uniform scheme which requires only a constant number of seeds and its total payment is always $O(\mathcal{H})$. The second scheme is the $(\frac{1}{\mathcal{H}}, \mathcal{H})$-almost-uniform scheme. This scheme works if the number of seeds is $\Omega(\mathcal{H})$. Its total payment is 2.

The Hybrid Scheme. We combine the $(\frac{1}{H}, H)$- and $(1, 1 + \log_d H)$-almost-uniform schemes to create a hybrid scheme that requires only a constant number of seeds and pays only a constant amount in expectation. We obtain the following result:

Theorem. *In the hybrid rewarding scheme, if the number of seeds is at least* 14, *the only strategies that always survive iterated elimination of dominated strategies exhibit information propagation and no duplication. In addition, there exists an elimination order in which the only strategies that survive exhibit information propagation and no duplication. Furthermore, the expected sum of payments is at most* 3.

References

[1] Moshe Babaioff, Shahar Dobzinski, Sigal Oren, and Aviv Zohar. On bitcoin and red balloons. EC '12.

[2] DARPA. The DARPA network challenge. Available online at http://archive.darpa.mil/networkchallenge/, 2009.

[3] Joshua Davis. The crypto-currency: Bitcoin and its mysterious inventor. The New Yorker, October 10, 2011.

[4] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. Available online at http://bitcoin.org/bitcoin.pdf, 2008.

[5] Galen Pickard, Wei Pan, Iyad Rahwan, Manuel Cebrián, Riley Crane, and Anmol Madan. Time critical social mobilization. *Science*, 334:509–512, 2011.

[6] James Surowiecki. Cryptocurrency. Technology Review, MIT. September/October, 2011. Available online: http://www.technologyreview.com/computing/38392/.

On Threshold Models over Finite Networks

Elie M. Adam, Munther A. Dahleh and Asuman Ozdaglar*

Networks intertwine with every aspect of our modern lives, be it through sharing ideas, communicating information, shaping opinions, performing transactions or delivering utilities. Explicitly, we may cite social networks, financial networks, economic networks, communication networks and power networks. Interactions over those many different types of networks require agents to coordinate with their neighbors. In economic networks, technologies that conform to the standards used by other related firms are more productive; in social networks, conformity to the behavior of friends is valuable for a variety of reasons. The desire for such coordination can lead to cascading behavior: the adoption decision of some agents can spread to their neighbors and from there to the rest of the network. One of the most commonly used models of such cascading behavior is the linear threshold model introduced by Granovetter in 1978. This model is used to explain a variety of aggregate level behaviors including diffusion of innovation, voting, propagation of rumors and diseases, spread of riots and strikes, and dynamics of opinions.

Most analyses of this model in the literature assume that one of the behaviors adopted by the agents (represented by the nodes of a graph) is irreversible, meaning that agents can only make a single switch into this behavior and can never switch out from it. However, incurring this progressive property in behavior dilutes several perspectives of the dynamics: whereas some situations are best captured by such a variant, many others cannot be captured but by allowing players to revert back to previous actions. A main motivation for example would be opinion dynamics in social network: in most situations a player changes opinions back and forth. This said, the literature lacks a satisfactory characterization of the limiting properties of such a model.

In this paper, we consider a model of cascade effects based on binary linear threshold dynamics over finite graphs. We start from an explicit coordination game set over a finite undirected network. The payoff of each agent is the sum of the payoff in a *two player and two action coordination game* the agent plays pairwise with each of the neighbors (the action is fixed across all interactions). We then study the behavior induced by best response dynamics, whereby each agent changes the played action to that which yields highest payoff given the actions of the neighbors. We first show that best response dynamics are identical to the dynamics traced by the linear threshold model with heterogeneous thresholds for the agents. However, crucially, actions can change multiple times. Thus, the dynamics of interest for the set of problems posed here cannot be studied using existing results and in fact have a different mathematical structure. The main contribution of this paper is to fully characterize these dynamics. We provide three sets of results.

We establish that agent behavior cycles among different actions in the limit, we term such limiting behavior as *convergence cycles.* We characterize the length of those convergence cycles.

---

[*] All authors are with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139. (emails: eadam@mit.edu, dahleh@mit.edu, asuman@mit.edu)

Ultimately, we show that for any graph structure on the players, any threshold distribution over the players and any initial action configuration played by the players, the limiting behavior of the dynamics get absorbed into action configuration cycles of length at most two. In other words, at the limit, every agent either plays one action, never deviating, or keeps on switching actions at every time step. We then characterize the time required to reach convergence cycles, termed as *convergence time.* Building on the framework set up, we show that for some positive integer m, given any graph structure on the players, any threshold distribution over the players and any initial action configuration, the dynamics reach a non-degenerate cycle or a fixed-point in at most mn^2 time steps where n is the number of players. We mention that similar results on convergence cycles and quadratic convergence time for linear threshold models (termed differently) have appeared in the literature on Cellular Automata. They consider the same dynamics, and prove the same bound on the length for convergence cycles and a quadratic bound on the convergence time. We approach the problem from a different perspective. We build a combinatorial framework for the analysis and begin by imposing restrictions on the model to obtain the essence of it, the simplest case. We then extend the results removing one restriction at a time. We take care to show how crucial properties carry on and morph as we increase the complexity of the model. Our aim is to give the reader enough familiarity and insight from those transformations to be able to manipulate the model as needed. We furthermore improve the convergence time bound from quadratic to be uniformly not more than the size of the network whenever the graph in concern is either a cycle graph or a tree.

We then study the complexity of counting and decision problems that arise in this model. We are interested in characterizing the number of limiting states the system could get absorbed in. We begin by arguing that no 'insightful' uniform upper-bound or lower-bound can be established. Considering only the case of a cycle graph, the number of fixed-points may vary at least from 2 to $2^{n/3}$ depending on the threshold distribution. Instead, we turn to study how tractable it is to count the convergence cycles. We proceed to show that given a graph structure on the player and a threshold distribution over the players as input, the problem of counting the number of limiting configuration classes (i.e. either fixed-points or non-degenerate cycles), the problem of counting the number of fixed-points and the problem of counting the number of non-degenerate cycles are all #P-Complete. We further consider the problem of deciding whether an action configuration over the network is reachable along the dynamics. To this end, we show that given a graph structure on the player, a threshold distribution over the players and some action configuration played by the players as input, the problem of deciding whether that action configuration is reachable is NP-Complete. We also show that given a graph structure on the player, a threshold distribution over the players and some *reachable* action configuration played by the players, the problem of counting the number of action configuration preceding that reachable action configuration is #P-Complete.

Of central importance in the study of cascades over networks is the resilience of networks to invasion by certain types of behavior (e.g., cascades of failures or spread of epidemics). For the new dynamics defined by our problem, we define a measure of resilience of a network to such invasion that captures the minimal 'cost of recovery' needed when the model is confronted with a perturbation in the agents' action profile. We prove achievable uniform lower-bounds and upper-bounds on the resilience measure, we compute the resilience measure of some network structures and provide basic insight on how different network structures affect this measure.

A Nonparametric Method for Early Detection of Trending Topics

Stanislav Nikolov*,† and Devavrat Shah†
{snikolov,devavrat}@mit.edu
†Department of EECS, Massachusetts Institute of Technology
*Twitter, Inc.

Abstract

Online social networks can be used as networks of human sensors to detect important events [3] — from a global breaking news story to an incident down the street. It is important to be able to detect such events as early as possible. To do so, we propose a nonparametric method that predicts *trending topics* on Twitter by comparing a recent activity signal for a topic to a large collection of historical activity signals for trending and non-trending topics. We posit that the signals observed for each class of topics were generated by an unknown set of *latent source* signals for that class according to a stochastic model depending on the *distance* between the observation and its latent source, and propose a class estimator based on this model. Using our method, we are able to detect trending topics in advance of Twitter 79% of the time, with a mean early advantage of 1 hour and 26 minutes, while maintaining a true positive rate of 95% and a false positive rate of 4%. In addition, our method allows for tradeoffs between error types and relative detection time, scales to large amounts of data, and provides a broadly applicable framework for nonparametric classification.

Empirical Observations

On Twitter, users can post short, public messages known as *Tweets*. There are over 400 million Tweets written every day, many of which can be considered *about* one or more topics. For example, this tweet by one of the authors (Twitter handle @snikolov) *"Stuyvesant High School Taps 'Stuy Mafia' at Google, Foursquare to Enhance Computer Science Program via @Betabeat http://betabeat.com..."* is about "Stuyvesant High School", "Computer Science", and so on. Because of the public nature of Twitter, topics can spread and gain popularity. Topics that gain sudden widespread popularity start *trending* i.e. they are featured on a list of top ten *trending topics* on Twitter.

Trending topics can typically be detected by a sudden high-magnitude spike in activity over some baseline of activity [2][1]. However, this sudden spike is often preceded by lower magnitude activity that is indicative of the topic's imminent popularity. This suggests that we can detect trending topics earlier by observing this early activity. Thus, we propose to predict whether a topic will become *trending* by comparing recent time series of activity for the topic to historical time series of activity leading up to other topics becoming trending, and historical time series of activity for topics that did not become trending.

We define the *activity signal* for a topic in terms of the rate $\rho[n]$ of Tweets about that topic over time, at time bins $n = 1, \ldots, N_{obs}$. We observe that activity is typically characterized by spikes above a baseline rate, so we further transform the rate to normalize away the baseline ($\rho_b[n] = (\rho[n]/b)^\beta$, $b = \sum_n \rho[n]/N_{obs}$) and emphasize spikes ($\rho_{b,s}[n] = |\rho_b[n] - \rho_b[n-1]|^\alpha$), according to parameters $\alpha \geq 1, \beta \geq 1$ (we used $\alpha = 1.2$, $\beta = 1$). In addition, we convolve the result with a smoothing window to eliminate noise and effectively measure the volume of Tweets in a sliding window ($\rho_{b,s,c}[n] = \sum_{m=n-N_{smooth}+1}^{n} \rho_{b,s}[m]$). Finally, because the spread of topics can reasonably be thought of as a branching process, and branching processes exhibit exponential growth, we measure the volume $\rho_{b,s,c}$ at a logarithmic scale ($\rho_{b,s,c,l}[n] = \log \rho_{b,s,c}[n]$).

Data Model

Activity signals, even within a single class, are incredibly diverse. Rather than training a model to distinguish between the activity signals of trending and non-trending topics, we assume no model structure at all and instead propose the following nonparametric model relating observed activity signals (*observations*) to their class labels. Suppose there are two classes: + (topics that were trending at some point during the period of interest) and − (topics that were never trending during the period of interest). We posit that there are a number of distinct *latent source* signals in each class that account for all observations in that class. Let us call them $\mathbf{t}_1, \ldots, \mathbf{t}_n$ for + and $\mathbf{q}_1, \ldots, \mathbf{q}_\ell$ for −. Each observation labeled + is assumed to be a noisy version of one of the latent sources $\mathbf{t}_1, \ldots, \mathbf{t}_n$. Similarly, each observation labeled − is assumed to be a noisy version of one of the latent sources $\mathbf{q}_1, \ldots \mathbf{q}_\ell$. We do not know what the latent source signals are or even how many there are. We only know a stochastic model that relates an observation to its latent source.

Let the observation $\mathbf{s}$ be the most recent N_{obs} samples of an infinite stream $\mathbf{s}_\infty$ of activity. An observation $\mathbf{s}$ is *generated* by a latent source signal $\mathbf{q}$ according to the stochastic model

$$\mathbb{P}(\mathbf{s} \text{ generated by } \mathbf{q}) \propto \exp(-\gamma d(\mathbf{s}, \mathbf{q})) \tag{1}$$

where d is a symmetric, positive definite, and convex distance function (we used the Euclidean norm) and γ is a scale parameter. To determine whether the observation belongs to + or −, we make use of a set of example, or *reference* activity signals $\mathcal{R}_+$ from + and $\mathcal{R}_-$ from −. Under our model the observation must belong to + if it has the same latent source as one of the reference signals in

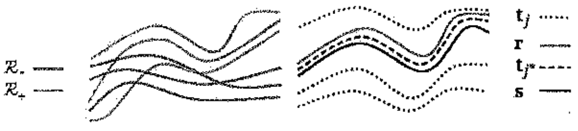


Figure 1: **Left:** Reference signals from each class. **Right:** Finding the latent source signal $\mathbf{t}_{j^*}$ that minimizes $d(\mathbf{s}, \mathbf{t}_j) + d(\mathbf{r}, \mathbf{t}_j)$, i.e. the latent source signal most likely to have generated both $\mathbf{s}$ and $\mathbf{r}$.

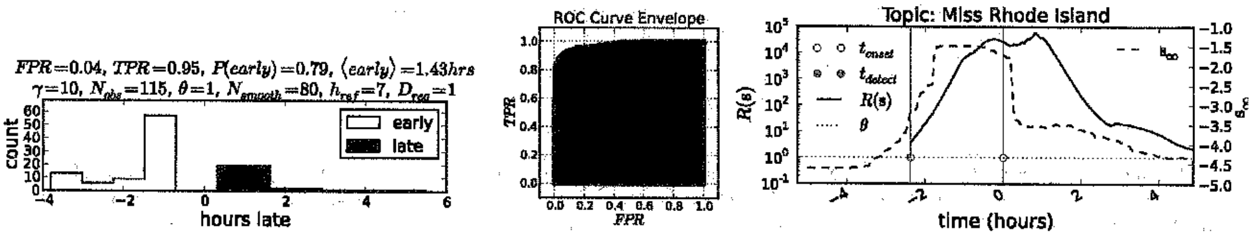


Figure 2: Our method is capable of early detection of trending topics while maintaining a low rate of error and provides the flexibility for tradeoffs between error types.

$\mathcal{R}_+$. Similarly, the observation must belong to $-$ if it has the same latent source as one of the reference signals in $\mathcal{R}_-$. Hence, the probability that the observation belongs to $+$ is

$$\mathbb{P}(+ \mid \mathbf{s}) = \sum_{\mathbf{r} \in \mathcal{R}_+} \mathbb{P}(\mathbf{s} \text{ belongs to } +, \mathbf{s} \text{ shares a latent source with } \mathbf{r}) = \sum_{\mathbf{r} \in \mathcal{R}_+} \sum_{j=1}^{n} \mathbb{P}(\mathbf{s} \text{ generated by } \mathbf{t}_j, \mathbf{r} \text{ generated by } \mathbf{t}_j)$$

$$\propto \sum_{\mathbf{r} \in \mathcal{R}_+} \sum_{j=1}^{n} \exp\left(-\gamma\left(d(\mathbf{s}, \mathbf{t}_j) + d(\mathbf{r}, \mathbf{t}_j)\right)\right) \approx \sum_{\mathbf{r} \in \mathcal{R}_+} \exp\left(-\gamma \min_j \left(d(\mathbf{s}, \mathbf{t}_j) + d(\mathbf{r}, \mathbf{t}_j)\right)\right) \approx \sum_{\mathbf{r} \in \mathcal{R}_+} \exp\left(-C\gamma d(\mathbf{s}, \mathbf{r})\right). \qquad (2)$$

The second to last approximation relies on the fact that for large enough γ, the term with the smallest exponent will dominate the sum over the latent sources. For the last approximation, we observe that the global minimum of $d(\mathbf{s}, \mathbf{t}) + d(\mathbf{r}, \mathbf{t})$ over all signals $\mathbf{t}$ is $Cd(\mathbf{r}, \mathbf{s})$ for some $C > 0$ and is achieved at $\mathbf{t}^* = (\mathbf{s} + \mathbf{r})/2$. The approximation is valid when the minimizing latent source $\mathbf{t}_{j^*}$ is sufficiently close to the global minimizer $\mathbf{t}^*$. We approximate $\mathbb{P}(- \mid \mathbf{s})$ in a similar fashion. In essence, because the latent sources are unknown, we cannot directly compare the observation with them. Instead we compare the observation to the reference signals, using them as a proxy for the latent sources. Figure 1 illustrates this. We classify the observation according to $f(\mathbf{s}) = \text{sgn}(R(\mathbf{s}) - \theta)$, where $R(\mathbf{s}) = \mathbb{P}(+|\mathbf{s})/\mathbb{P}(-|\mathbf{s})$ and θ is a threshold. We can optionally require D_{req} consecutive "detections" to declare something a trending topic. This approach has an appealing interpretation: to classify an observation, one simply computes the distance from the observation to all examples from each class. This can be done in parallel on enormous data sets.

Results and Conclusion

We collected 500 topics that were trending at some point during June 2012, and 500 that were not. We then collected 10% of the Tweets from June 2012 containing those topics. We used a 50/50 split between reference signals and observations and performed detection in a window of size $2h_{ref}$ hours, centered around the true onset of the trending topic if the topic was trending or chosen randomly otherwise. We performed detection for a range of parameters $\gamma, N_{obs}, h_{ref}, N_{smooth}, \theta$ and D_{req} to evaluate tradeoffs between detection errors and relative detection time. Using this small sample of Tweets, our method is capable of detecting trends in advance of Twitter 79% of the time, with a mean early advantage of 1 hour and 26 minutes, while maintaining a 95% true positive rate and a 4% false positive rate (Figure 2 (left)). Figure 2 (right) shows an example early detection in which our method detected the trending topic "Miss Rhode Island" over 2 hours in advance of Twitter. In Figure 2 (center) the envelope of ROC curves for all combinations of parameters shows the ability of our method to perform well under a variety of tradeoffs between types of error. It should be noted that Twitter's trending topic detection method may need to be more conservative to avoid low-quality trending topics, and that on Twitter, trending topics compete for the top ten spots, whereas our method is based on a score threshold alone. Nevertheless, our results demonstrate the effectiveness of our method for trending topic detection, as well as its potential as a broadly applicable framework for scalable nonparametric classification in the presence of large amounts of data.

References

[1] Mathioudakis, M., and Koudas, N. Twittermonitor: Trend detection over the twitter stream. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data* (New York, NY, USA, 2010), SIGMOD '10, ACM, pp. 1155–1158.

[2] @twitter. To trend or not to trend. `http://blog.twitter.com/2010/12/to-trend-or-not-to-trend.html` (accessed 9 September, 2012).

[3] Zhao, S., Zhong, L., Wickramasuriya, J., and Vasudevan, V. Human as real-time sensors of social and physical events: A case study of twitter and sports games. *CoRR abs/1106.4300* (2011).

Active Influence in Dynamical Models of Structural Balance in Social Networks

Tyler H. Summers[†] and Iman Shames[‡]*

[†]ETH Zürich [‡]University of Melbourne

August 19, 2012

Abstract

In mathematical models for social network analysis, links in the network often have positive connotations, such as friendship, collaboration, information sharing, etc. However, negative interactions in social networks, such as antagonism or distrust, also play a key role in both structure and dynamics of social networks and are receiving increased attention in the literature [4, 8, 7].

The concept of *structural balance* is an old idea in sociology, tracing back to social psychology research in the 1940s by Heider [5]. The theory begins with notions of tension and balance in three-agent networks. Imagine that a person has two good friends who hate each other. There is a tension in this situation that is resolved when either the person takes one side and ends the friendship with the other or when the feuding friends reconcile their differences. Similarly, there is a tension amongst three people unfriendly with one another that is resolved when two of them to form an alliance against the other. The theory was generalized to networks of n agents in the 1950s by Cartwright and Harary [3], who model the network as a complete signed graph in which n vertices represent agents, a complete edge set represent relationships amongst all agents in the network, with positive signs associated with positive relationships and negative with negative relationships. They showed that n-agent structurally balanced networks are those that can be partitioned into two factions, such that within each faction all relationships are friendly and between factions all relationships are hostile. The theory has found various applications, e.g. in models of international relations [2], but has remained mainly static, focusing only on network structure.

Dynamic models for structural balance are quite recent and provide a new and interesting perspective. They are particularly interesting in light of new online networks that provide real-world dynamical data for social networks with both positive and negative relationships. For example, users on the product review website Epinions can display both trust and

---

*tsummers@control.ee.ethz.ch, iman.shames@unimelb.edu.au

distrust of other users; on the technology news website Slashdot, users can designate other users as either friend or foe; and on Wikipedia, users can vote for or against another person becoming an administrator. The structure of such networks has been studied in [7, 8].

Suppose that a network is in an initial state that is not structurally balanced. How might the state of the network evolve toward a structurally balanced state? Discrete dynamical models, in which a relationship is either positive or negative, have been proposed in [1] and [10]. In these models, the system evolves by flipping the sign on certain edges to increase the number structurally balanced triangles in the network. However, these models suffer from the existence of so-called "jammed states", in which the system becomes stuck in structurally unbalanced local minima. More recently, Kulakowski et al [6] and Marvel et al [9] have proposed and analyzed continuous dynamical models, in which a real-valued "friendliness level" is associated with each relationship (positive values indicate friendliness and negative values indicate hostility). Collecting the friendliness levels in the network into a symmetric matrix, the dynamic model is given by the matrix differential equation

$$\dot{X} = X^2. \tag{1}$$

In this model, Marvel et al [9] show that for generic initial conditions, the system converges to a structurally balanced state in finite time. Further, the final balanced state is determined by the outer product of the eigenvector associated with the largest eigenvalue of the initial state matrix.

In this paper, we study external influence in dynamic models of structural balance. In particular, we suppose again that a network is in an initial state that is not structurally balanced and that a single agent can influence the state of the network (in an international relations context, this might be due to certain foreign policy actions). We consider the following problem:

Problem 1. *Given a symmetric matrix X_0, find, if it exists, a symmetric matrix ΔX_0 with entries equal to zero except for possibly the entries of the ℓ-th row and the ℓ-th column such that*

$$\widetilde{v}_i v_i^\star > 0, \quad \forall i$$

where $\widetilde{v}$ is the eigenvector associated with the largest eigenvalue of $\widetilde{X}_0 = X_0 + \Delta X_0$ and $v^\star$ is a vector with desired sign pattern and entries of either 1 or -1.

We show that it is possible for a single agent to choose such a perturbation and achieve any desired structurally balanced state given any initial state. We also present a method to compute and optimize the influence that is required to achieve the desired state. The results are applied to structural balance in an international relations network, using data from United Nations General Assembly voting records dating from 1946 to 2008 to estimate friendliness levels amongst various countries. Our results give interesting historical interpretations and provide guidance for how a country might optimally achieve certain foreign policy goals.

References

[1] T. Antal, PL Krapivsky, and S. Redner. Dynamics of social balance on networks. *Physical Review E*, 72(3):036121, 2005.

[2] R.M. Axelrod and D.S. Bennett. *A landscape theory of aggregation.* Cambridge Univ Press, 1991.

[3] D. Cartwright and F. Harary. Structural balance: a generalization of heider's theory. *Psychological review*, 63(5):277, 1956.

[4] D. Easley and J. Kleinberg. *Networks, crowds, and markets.* Cambridge Univ Press, 2010.

[5] F. Heider. Attitudes and cognitive organization. *The Journal of Psychology*, 21(1):107–112, 1946.

[6] K. Kulakowski, P. Gawronski, and P. Gronek. The heider balance-a continuous approach. *Arxiv preprint physics/0501073*, 2005.

[7] J. Leskovec, D. Huttenlocher, and J. Kleinberg. Predicting positive and negative links in online social networks. In *Proceedings of the 19th international conference on World wide web*, pages 641–650. ACM, 2010.

[8] J. Leskovec, D. Huttenlocher, and J. Kleinberg. Signed networks in social media. In *Proceedings of the 28th international conference on Human factors in computing systems*, pages 1361–1370. ACM, 2010.

[9] S.A. Marvel, J. Kleinberg, R.D. Kleinberg, and S.H. Strogatz. Continuous-time model of structural balance. *Proceedings of the National Academy of Sciences*, 108(5):1771–1776, 2011.

[10] S.A. Marvel, S.H. Strogatz, and J.M. Kleinberg. Energy landscape of social balance. *Physical review letters*, 103(19):198701, 2009.

SODEXO: A System Framework for Deployment and Exploitation of Deceptive Honeybots in Social Networks

Quanyan Zhu, Coordinated Science Laboratory and Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA. Email: zhu31@illinois.edu

Andrew Clark, Department of Electrical Engineering, University of Washington, Seattle, WA 98195 USA. Email: awclark@uw.edu

Radha Poovendran, Department of Electrical Engineering, University of Washington, Seattle, WA 98195 USA. Email: rp3@uw.edu

Tamer Başar, Coordinated Science Laboratory and Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA. Email: basar1@illinois.edu

Abstract:

As social networking sites such as Facebook and Twitter are becoming increasingly popular, a growing number of malicious attacks, such as phishing and malware, are exploiting them. Among these attacks, social botnets have sophisticated infrastructure that leverages compromised users accounts, known as *bots* to automate the creation of new social networking accounts for spamming and malware propagation. Traditional defense mechanisms are often passive and reactive to non-zero-day attacks. In this paper, we adopt a proactive approach for enhancing security in social networks by infiltrating botnets with honeybots. We propose an integrated system named SODEXO which can be interfaced with social networking sites for creating deceptive honeybots and leveraging them for gaining information from botnets. We establish a Stackelberg game framework to capture strategic interactions between honeybots and botnets, and use quantitative methods to understand the tradeoffs of honeybots for their deployment and exploitation in social networks. We design a protection and alert system that integrates both microscopic and macroscopic models of honeybots and optimally determines the security strategies for honeybots. We corroborate the proposed mechanism with extensive simulations and comparisons with passive defenses.

References:

Q. Zhu, A. Clark, R. Poovendran, and T. Başar, "SODEXO: A system framework for deployment and exploitation of deceptive honeybots in social networks," in Arxiv preprint 1207.5844, 2012. URL: http://arxiv.org/abs/1207.5844

SODEXO: A System Framework for Deployment and Exploitation of Deceptive Honeybots in Social Networks

Quanyan Zhu, Andrew Clark, Radha Poovendran and Tamer Başar

***Abstract*— As social networking sites such as Facebook and Twitter are becoming increasingly popular, a growing number of malicious attacks, such as phishing and malware, are exploiting them. Among these attacks, social botnets have sophisticated infrastructure that leverages compromised users accounts, known as *bots*, to automate the creation of new social networking accounts for spamming and malware propagation. Traditional defense mechanisms are often passive and reactive to non-zero-day attacks. In this paper, we adopt a proactive approach for enhancing security in social networks by infiltrating botnets with honeybots. We propose an integrated system named SODEXO which can be interfaced with social networking sites for creating deceptive honeybots and leveraging them for gaining information from botnets. We establish a Stackelberg game framework to capture strategic interactions between honeybots and botnets, and use quantitative methods to understand the tradeoffs of honeybots for their deployment and exploitation in social networks. We design a protection and alert system that integrates both microscopic and macroscopic models of honeybots and optimally determines the security strategies for honeybots. We corroborate the proposed mechanism with extensive simulations and comparisons with passive defenses.**

Keywords: social networks; cyber security; game theory; botnet; malware propagation; Stackelberg games

I. INTRODUCTION

Online social networks such as Facebook and Twitter are employed daily by hundreds of millions of users to communicate with acquaintances, follow news events, and exchange information. The growing popularity of OSNs has led to a corresponding increase in spam, phishing, and malware on social networking sites. The fact that a user is likely to click on a web link that appears in a friend's Facebook message or Twitter feed can be leveraged by attackers who compromise or impersonate that individual.

An important class of malware attacks on social networks is social botnets [1], [2]. In a social botnet, an infected user's device and social networking account are both compromised by installed malware. The compromised account is then used to send spam messages to the user's contacts, containing links to websites with the malware executable. As a result, compromising a single well-connected user could lead to hundreds or thousands of additional users being targeted for spam, many of whom will also become members of the botnet and further propagate the malware. The most prominent example of a social botnet to date is Koobface, which at its peak had infected 600,000 hosts [1].

Current methods for mitigating malware, including social botnets, in social networks are primarily based on URL blacklisting. In this defense mechanism, links that are suspected to contain spam or malware are added to a centralized blacklist controlled by the owner of the social network. After a link has been blacklisted, the social networking site will no longer communicate with the IP address indicated by the link, even if a user clicks the link [3].

While blacklisting can slow the propagation of malware, there remain several drawbacks to this approach. First, automated methods for blacklisting links often fail to detect spam and malware; one survey suggests that 73% of malicious links go undetected and are not added to the blacklist [4]. Second, automated blacklisting creates the risk of valid accounts and messages being classified as spam, degrading the user experience. Third, even for links that are correctly identified as pointing to malware, there is typically a large delay between when links are detected and blacklisted. One study estimates this delay as 25 days on average, while at the same time most clicks on malware links occur within the first 48 hours of posting [5].

A promising approach to defending against social botnets is through deception mechanisms. In a deceptive defense, the defender generates fake social network profiles that appear similar to real profiles and waits to receive a link to malware. The defender then follows the link to the malware site, downloads the malware executable, and runs it in a quarantined, sandbox environment. By posing as an infected node and interacting with the owner of the botnet, the defender gathers links that are reported to the blacklist either before or shortly after they are posted, reducing the detection time and increasing the success rate. Currently, however, there is no systematic approach to modeling social botnets and the effectiveness of deception, as well as designing an effective strategy for infiltrating the botnet and gathering information.

In this paper, we introduce an analytical framework for SOcial network DEception and eXploitation through hOneybots (SODEXO). Our framework has two components, *deployment* and *exploitation*. The deployment component models how decoy accounts are introduced into the online social network and gain access to the botnet. The exploitation component characterizes the behavior of the decoys and the botnet owner after infiltration has occurred, enabling us to

The research was partially supported by the AFOSR MURI Grant FA9550-10-1-0573, and also by an NSA Grant through the Information Trust Institute at the University of Illinois.

Q. Zhu and T. Başar are with the Coordinated Science Laboratory and Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA. Email: `{zhu31, basar1}@illinois.edu`

Andrew Clark and Radha Poovendran are with the Department of Electrical Engineering, University of Washington, Seattle, WA 98195 USA. Email: `{awclark, rp3}@u.washington.edu`

model the effect of the decoy on the botnet operation.

For the deployment component, we first develop a dynamical model describing the population of a social botnet over time. We derive the steady-state equilibria of our model and prove the stability of the equilibria. We then formulate the problem of selecting the optimal number of honeybots in order to maximize the information gathered from the botnet as a convex optimization problem. Our results are extended to include networks with heterogeneous node degree.

We model the exploitation of the botnet by the honeybots as a Stackelberg game between the botmaster and the honeybots. In the game, the botmaster allocates tasks, such as spam message delivery, among multiple bots based on their trustworthiness and capabilities. The honeybots face a trade-off between obtaining more information by following the commands of the botmaster, and the impact of those commands on other network users. We derive closed forms for the optimal strategies of both the botmaster and honeybots using *Stackelberg equilibrium* as a solution concept. We then incorporate the utility of the honeybot owner under the Stackelberg equilibrium in order to select an optimal deployment strategy.

The paper is organized as follows. The related work is reviewed in Section II. In Section III, we describe the architecture of our proposed framework for deceptive defense. In Section IV, we model the exploitation phase of the botnet, in which the honeybot gathers the maximum possible information while avoiding detection by the botmaster. In Section V, we model the deployment and population dynamics of the infected nodes and honeybots. Section VI describes the Protection and Alert System (PAS), which provides a unifying framework for controlling deployment and exploitation. Section VII presents our simulation results. Section VIII concludes the paper.

II. RELATED WORK

Social botnets are becoming a serious threat for network users and managers, as they possess sophisticated infrastructure that leverages compromised users accounts, known as *bots*, to automate the creation of new social networking accounts for spamming and malware propagation [2].In [6], a honeypot-based approach is used to uncover social spammers in online social systems. It has been shown that social honeypots can be used to identify social spammers with low false positive rates, and that the harvested spam data contain signals that are strongly correlated with observable profile features, such as friend information and posting patterns. The goal of [6], however, is not to infiltrate the botnet, but to use honeypots to differentiate between real and spam online profiles.

In [4], a zombie emulator is used to infiltrate the Koobface botnet to discover the identities of fraudulent and compromised social network accounts. The authors arrived at the conclusion that "to stem the threat of Koobface and the rise of social malware, social networks must advance their defenses beyond blacklists and actively search for Koobface content, potentially using infiltration as a means of early detection." This insight coincides with our proactive approach for defending social networks using deceptive social honeybots.

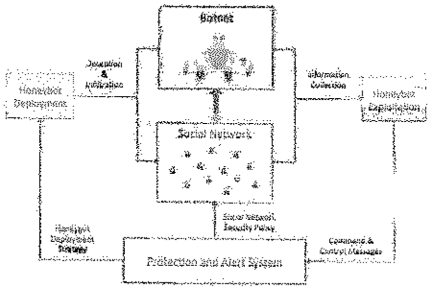


Fig. 1. System architecture of honeybot deceptive mechanism in social networks

Deception provides an effective approach for building proactively secure systems [7], [8]. Considerable amount of work can be found using deception for enhancing cyber security. In recent literature on intrusion detection systems, honeypots have been used to monitor suspicious intrusions [9], [10], and provide signatures of zero-day attacks [11]. In [12], to enhance the security of control systems in critical infrastructure, deception has been proposed to make the system more difficult for attackers to plan and execute successful attacks. At present, however, there has been no analysis on the impact of deception on malware propagation in social networks.

In order to establish a formal method to evaluate the performance of deceptive social honeybots against botnets, we employ a game- and system- theoretic approach to model the strategic behaviors of botnets and the deployment and exploitations of honeybots. Such approaches have become pivotal for designing security mechanisms in a quantitative way [13]. In [14], an optimal control approach to modeling the maximum impact of a malware attack on a communication network is presented. In [15], the authors have proposed an architecture for a collaborative intrusion detection network and have adopted a game-theoretic approach for designing a reciprocal incentive compatible resource allocation component of the system against free-rider and insider attacks.

III. SYSTEM ARCHITECTURE

In this section, we introduce our honeybot-based defense system named SODEXO for protecting social networks against malicious attacks. Fig. 1 illustrates the architecture of SODEXO. Our framework consists of two components, namely, honeybot deployment (HD) and honeybot exploitation (HE). HD deals with the distribution of honeybots within social networks and the deception mechanisms to infiltrate the botnet to learn and monitor the activities in botnets. HE aims to use the successfully infiltrated honeybots to collect as much information as possible from the botnet. The behaviors of the two blocks are coordinated by a Protection and Alert System (PAS), which uses the gathered information to generate real-time signatures and alerts for the social network (Fig. 2).

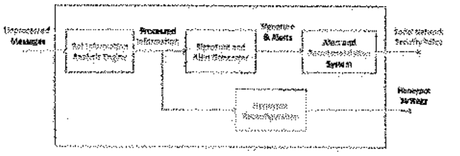


Fig. 2. Architecture of the protection and alert system

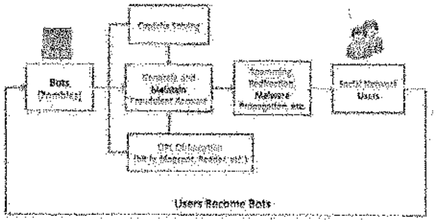


Fig. 3. Illustration of the interactions between social networks and botnets

The introduction of honeybots into a social network allows a proactive defense and monitoring of the social network against botnets. The SODEXO architecture bears its resemblance to feedback control systems. The HE component behaves as a security sensor of the social network; PAS can be seen as a controller which takes the "measurements" from HE and yields a honeypot deployment strategy; and HD acts as an actuator that updates the honeypot policy designed by PAS. In the following subsections, we discuss in detail each component of SODEXO.

A. Honeybot Deployment (HD)

A honeypot is deployed by first creating an account on a social networking site. The account profile is designed to imitate a real user, as in [6]. Once deployed, the honeypot sends a set of friend requests to a set of randomly chosen other users. The honeypot continues sending friend requests to random users until the desired number of neighbors, denoted d, has been reached. The honeypot monitors the message traffic of its neighbors, which may include personal messages, wall posts, or Twitter feeds, and follows any posted link. If the link points to malware and has not been blacklisted, then the honeypot becomes a member of the social botnet and proceeds to the exploitation stage.

B. Honeybot Exploitation (HE)

The HE component of SODEXO takes advantage of the successfully infiltrated honeybots to gain as much information as possible from the botnet. The information is obtained in the form of command and control messages. The honeybots need to gain an appropriate level of trust from the bots and respond to the C&C messages while minimizing harm to the legitimate social network users and avoiding legal liability. Honeybots work collaboratively to achieve this goal. In the case where honeybots are commanded to send spam or malware to network users, they can send them to each other to remain active in the botnet. Depending on the sophistication of the botnet, honeybots can sometimes be detected using mechanisms described in [16], [17]. In this case, a higher growth rate of honeybot population will be needed to replace the detected honeybots. Hence, the performance of HE heavily depends on the effectiveness of HD, and in turn, HD should change its policy based on the sophistication of botnets and the amount of information learned in HE.

C. Protection and Alert System (PAS)

The major role of PAS is to provide security policies for HD based on the information learned from HE. Fig. 2 illustrates two major functions of PAS. The first step of PAS is to process the messages and logs gained from honeybots. Using data mining and machine learning techniques, it is possible that the structure of botnets can be inferred from network traffic information [18] and botnet C&C channels in a local area network can be identified [19]. These information can be used by the network administrator to detect the location of botmasters and remove them from the network.

The second important task of HD is to generate signatures for detecting malware and spam, which are then used to update the libraries of intrusion detection systems, blacklists of spam filters, and user alerts or recommendations. The process of reconfiguration of IDSs and spam filters can be done either offline or real-time as in [20] and [21].

D. Botnet Propagation Model

Fig. 3 illustrates a mechanism used by botnets to infect social network users, which has been found in the Koobface botnet [2], [4]. The botnet maintains a fixed domain that bots or zombies regularly contact to report uptime statistics and request links for spamming activity. The bots aim to obtain fresh user accounts and send malicious messages. The bot messages contain a malicious URL obfuscated by shortening services such as bit.ly or wrapped by an innocuous website including Google Reader and Blogger. Clicking on the URL of these messages eventually redirects to a spoofed Youtube or Facebook page that attempts to trick the victim into installing malware masquerading as a Flash update. Unsuspecting users become infected by clicking on these messages. Infected users are recruited to spam their own social network friends, leading to a wide propagation of malware within social network users.

Once a user has been compromised, it makes frequent attempts to connect with one or more command control (C&C) bots to retrieve commands from the botnet for further actions. These commands are usually issued from another compromised computer for the purpose of concealing the botmaster's real identity [16], leading to a hierarchical botnet architecture. Fig. 4 illustrates the structure of a typical botnet, where a single botmaster sends messages to two C&C bots and then they send to bots.

IV. SYSTEM MODEL FOR HONEYBOT EXPLOITATION

In this section, we introduce a system model for hierarchical botnets and employ a Stackelberg game framework to model the interactions between the botnet and infiltrating honeybots.

A. Theoretical Framework

Consider a botmaster B that sends requests to a set of C&C bots $\mathcal{M} = \{1, 2, \cdots, m\}$ with $m = |\mathcal{M}|$. Each C&C bot $i \in \mathcal{M}$ sends commands to a set of compromised bot nodes $\mathcal{N}_i$ with $n_i = |\mathcal{N}_i|$. We assume that the botnet is a three-level tree architecture and, without loss of generality, we can assume $\cap_{i\in\mathcal{M}} \mathcal{N}_i = \emptyset$ since a single bot controlled by multiple C&C bots can be modeled using multiple duplicate bots. Let H be a honeybot that communicates with node $i \in \mathcal{M}$, i.e., $H \in \mathcal{N}_i$. We assume that all honeybots work together as a team, and hence one honeybot node H under one C& C subtree can conceptually represent a group of collaborative honeypots who have succeeded in infiltrating the same botnet.

We let $p_{ij} \in \mathbb{R}_+$ be the number of messages or commands (in bytes) per second sent from C&C bot i to bot node $j \in \mathcal{N}_i$. Likewise, p_{ji} denotes the number of response messages per second to C&C node $i \in \mathcal{M}$ from node $j \in \mathcal{N}_i$.

Each C&C node i maintains a trust value $T_{ij} \in [0,1]$ associated with a bot or honeybot node $j \in \mathcal{N}_i$. The trust values indicate the quality of response and performance of bot nodes. The trust values also inherently model the detection mechanisms in botnets, which have been discussed in [16], [22]. For botnets with such mechanisms, low trust values indicate the inefficiency of a bot or a high likelihood of being a honeybot. For those without such mechanisms, we can take $T_{ij} = 1$, for all $j \in \mathcal{N}_i$, i.e., equivalently seeing all bots are all equally trusted.

One C&C bot needs to send commands to a large population of bot nodes. Hence, the goal of C&C bot $i \in \mathcal{M}$ is to allocate its communication resources $\mathbf{p}_i := [p_{ij}, j \in \mathcal{N}_i]$ to maximize the utility of its subtree network $U_i : \mathbb{R}_+^{n_i} \to \mathbb{R}$, which is the sum of utilities obtained from each bot j, i.e.,

$$U_i(\mathbf{p}_i) = \sum_{j\in\mathcal{N}_i} U_{ij}, \tag{1}$$

where $U_{ij} : \mathbb{R}_+ \to \mathbb{R}$ is the individual utility of C&C bot i from bot $j \in \mathcal{N}_i$, which is chosen to be

$$U_{ij}(p_{ij}) := T_{ij} p_{ji} \ln(\alpha_i p_{ij} + 1). \tag{2}$$

The choice of logarithmic function in (2) indicates that the marginal utility of the C&C bot diminishes as the number of messages increases. It captures the fact that the bots have limited resources to respond to commands, and a larger volume of commands can overwhelm the bots, which leads to diminishing marginal utility of node i. $\alpha_i \in \mathbb{R}_{++}$ is a positive system parameter that determines marginal utility.

The utility of C&C bot i is also proportional to the number of messages or responses per second from bot j, indicated by $p_{ji} \in \mathbb{R}_+$. The number of response messages from bot j indicates the level of activity of a bot. We can see that when $p_{ji} = 0$ or $T_{ij} = 0$ in (2), then bot i is believed to be either inactive or fake, and it is equivalently removed from the subtree of C&C node j in terms of the total utility (1). Note that T_{ij} in (2) evaluates the quality of the responses while p_{ji} evaluates the quantity. The product of T_{ij} and p_{ji} captures the fact that the botnet values highly active and trusted bots.

We consider the following C&C bot optimization problem (BOP) of every node $i \in \mathcal{M}$:

$$\begin{aligned} \text{(BOP)} \max_{\mathbf{p}_i\in\mathbb{R}_+^{n_i}} \quad & U_i := \textstyle\sum_{j\in\mathcal{N}_i} T_{ij} p_{ji} \ln(\alpha_i p_{ij} + 1) \\ \text{s.t.} \quad & \textstyle\sum_{j\in\mathcal{N}_i} c_{ij} p_{ij} \le C_i. \end{aligned} \tag{3}$$

The constraint (3) in (BOP) is a capacity constraint on the communications using C&C channel, where C_i is the total capacity of the channel. The cost $c_{ij} \in \mathbf{R}_{++}$ is the cost on sending commands to bots. The cost is also dependent on the size of messages from C&C bot i to its controlled bots. It has been found in [4] that Twitter has larger volume of spam messages than Facebook. This is due to the fact that Twitter messages are often shorter than facebook messages, and hence the cost for commanding bots spamming with Twitter messages is relatively less than the one for Facebook.

Let $\mathcal{F}_i := \{\mathbf{p}_i \in \mathbb{R}_+^{n_i} : \sum_{j\in\mathcal{N}_i} c_{ij} p_{ij} \le C_i\}$ be the feasible set of (BOP). We let $\mathcal{L}_i : \mathbb{R}_+^{n_i} \times \mathbb{R} \to \mathbb{R}$ be the associated Lagrangian defined as follows:

$$\mathcal{L}_i(\mathbf{p}_i, \lambda_i) = \sum_{j\in\mathcal{N}_i} T_{ij} p_{ji} \ln(\alpha_i p_{ij} + 1) + \lambda_i \left(\sum_{j\in\mathcal{N}_i} c_{ij} p_{ij} - C_i \right) \tag{4}$$

Since the feasible set is nonempty and convex, and the objective function is convex in $\mathbf{p}_i$, it is clear that (BOP) is a convex program, and hence we can use the first-order optimality condition to characterize the optimal solution to (BOP):

$$\frac{\partial \mathcal{L}_i}{\partial p_{ij}} = \sum_{j\in\mathcal{N}_i} \frac{\alpha_i T_{ij} p_{ji}}{\alpha_i p_{ij} + 1} + \lambda_i c_{ij} = 0, \tag{5}$$

which leads to

$$p_{ij} = -\frac{T_{ij}}{\lambda_i c_{ij}} - \frac{1}{\alpha_i}. \tag{6}$$

Due to the monotonicity of logarithmic functions in (2), the optimal solution is found on the Pareto boundary of feasible set. Hence by letting $\sum_{j\in\mathcal{N}_i} c_{ij} p_{ij} = C_i$, we obtain Lagrangian multiplier λ_i from (6) as follows.

$$\lambda_i = -\frac{\sum_{j\in\mathcal{N}_i} T_{ij} p_{ji}}{C_i + \frac{1}{\alpha_i} \sum_{j\in\mathcal{N}_i} c_{ij}}. \tag{7}$$

We make following assumptions before stating Theorem 1.

(A1) The product $T_{ij} p_{ji} \ne 0$ for all $j \in \mathcal{N}_i, i \in \mathcal{M}$.

Assumption (A1) states that all bots controlled by C&C bot i are both active and trusted. This assumption is valid because for a controlled bot j that is either inactive ($p_{ij} = 0$) or untrusted ($T_{ij} = 0$) can be viewed as the one excluded from the set $\mathcal{N}_i$. Hence Assumption (A0) is equivalent to the statement that $\mathcal{N}_i$ contains all active and trusted bots.

Theorem 1: Under Assumption (A1), (BOP) admits a unique solution when α_i is sufficiently large.

$$p_{ij} = \left(\frac{T_{ij} p_{ji}}{\sum_{j\in\mathcal{N}_i} T_{ij} p_{ji}} \right) \left(\frac{C_i + \frac{1}{\alpha_i} \sum_{j\in\mathcal{N}_i} c_{ij}}{c_{ij}} \right) - \frac{1}{\alpha_i}. \tag{8}$$

Proof: Assumption (A1) ensures that (BOP) is strictly convex in p_{ij} for all $j \in \mathcal{N}_i$. Hence the result follows directly from (6) and (7). Since α_i is a system parameter, we can choose α_i sufficiently large so that the solution obtained in

(8) is nonnegative. ■

B. Stackelberg Game

In this section, we formulate a two-stage Stackelberg between honeybots and C&C nodes. Honeybots behave as leaders who can learn the behaviors of the C&C bots once they succeed in infiltrating the botnet and choose the optimal strategies to respond to the commands from C&C bots.

The goal of honeypots is to collect as much information as possible from the botmaster. We consider the following game between honeypots and a C&C bot. The honeypot node H firsts chooses a response rate p_{Hi} to the commands from C&C bot i, and then C&C bot i observes the response and chooses an optimal rate to send information to honeybot H according to (BOP). We make the following assumption on the real bots in the network.

(A2) The real bots are not strategically interacting with the C&C bot i, i.e., they send messages to bot i at a constant rate $p_{ij}, j \neq H, j \in \mathscr{N}_i$.

The above assumption holds because bots are non-human driven, pre-programmed to perform the same routine logic and communications as coordinated by the same botmaster [19]. Under Assumption (A2), the strategic interactions exist only between honeybots and C&C nodes.

The honeypot node H has a certain cost when it responds to the botnet. This can be either because of the potential harm that it can cause on the system or due to the cost of implementing commands from the botmaster. We consider the following honeypot optimization problem (HOP), where node H aims to maximize its utility function $U_H : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}_+$ as follows:

$$\text{(HOP)} \max_{p_{Hi} \in \mathscr{F}_H} \quad U_H(p_{iH}, p_{Hi}) := \ln(p_{iH} + \xi_H) - \beta_i^H p_{Hi},$$

where $\xi_H \in \mathbb{R}_{++}$ is a positive system parameter; β_i^H is the cost of honeybot H responding to the bot node i; p_{Hi} is the message sending rate from honeybot node H to C&C bot i and p_{Hi} is the rate of C&C bot i sending commands to H.

$\mathscr{F}_H$ denotes the feasible set of the honeypot node H. We let $\mathscr{F}_H = \{p_{Hi}, 0 \leq p_{Hi} \leq p_{Hi,\max}\}$, where $p_{Hi,\max} \in \mathbb{R}_{++}$ is a positive parameter that can be chosen to be sufficiently large. The logarithmic part of the utility function (9) is used to model the property of diminishing returns of an information source. The value of receiving an additional piece of information from the C&C bot decreases as the total number of messages received by the honeypot increases.

The interactions between honeypot H and C&C node i can be captured by the Stackelberg game model $\Xi_S := \langle (i,H), (U_i, U_H), (\mathscr{F}_i, \mathscr{F}_H) \rangle$, and Stackelberg equilibrium can be used as a solution concept to characterize the outcome of the game.

Definition 1 (Stackelberg Equilibrium): Let $\pi_{iH}^*(\cdot) : \mathbb{R}_+^{n_i} \to \mathbb{R}_+$ be the unique best response of the C&C bots to the response rate p_{Hi} of the honeypots. An action profile $(\mathbf{p}_i^*, \ p_{Hi}^*) \in \mathscr{F}_i \times \mathscr{F}_H$ is a *Stackelberg equilibrium* if $\mathbf{p}_i^* = \pi_{iH}(p_{Hi}^*)$, and the following inequality holds $U_H(\pi_{iH}(p_{Hi}^*), p_{Hi}^*) \geq U_H(\pi_{iH}(p_{Hi}), p_{Hi}), \ \forall p_{Hi} \in \mathscr{F}_H$.

Theorem 2: Under Assumption (A1), the nonzero-sum continuous-kernel Stackelberg game Ξ_S admits a Stackelberg equilibrium.

Proof: The utility function of C&C bot i is strictly convex for all $p_{Hi} \neq 0$ under Assumption (A1). Since $\mathscr{F}_H$ and $\mathscr{F}_B$ are compact sets, by Corollary 4.4 of [23], the game admits a Stackelberg equilibrium solution. ■

Under Assumption (A1), the unique best response $\pi_{iH}(\cdot)$ can be obtained from (8) for sufficiently large α_i as follows:

$$p_{iH} = \pi_{iH}(p_{Hi}) = C_H \left(\frac{T_{iH} p_{Hi}}{T_{iH} p_{Hi} + I_{-H}} \right) - \frac{1}{\alpha_i}, \qquad (9)$$

where $I_{-H} = \sum_{j \neq H, j \in \mathscr{N}_i} T_{ij} p_{ji}$ is the number of responses from real bots weighted by their trust values and $C_H := \frac{C_i + \frac{1}{\alpha_i} \sum_{j \in \mathscr{N}_i} c_{ij}}{c_{iH}}$.

Letting $\bar{\xi}_H = 1/\alpha_i + \xi_H$ and substituting (9) in (HOP), we arrive at the following optimization problem faced by the honeybot node H:

$$\max_{p_{Hi} \in \mathscr{F}_H} U_H(\pi_{iH}(p_{Hi}), p_{Hi}) := \ln \left(C_H \left(\frac{T_{iH} p_{Hi}}{T_{iH} p_{Hi} + I_{-H}} \right) + \bar{\xi}_H \right) - \beta_i^H p_{Hi}. \qquad (10)$$

Theorem 3: Under Assumptions (A1) and (A2), the Stackelberg equilibrium solution $(\mathbf{p}_i^*, \ p_{Hi}^*)$ of the game Ξ_S is unique and can be found as follows:

$$\begin{aligned} p_{Hi}^* \ = \ & \frac{C_H I_{-H}}{2 T_{iH} (C_H + \xi_H)} \left(\sqrt{1 + 4 \frac{T_{iH}(C_H + \bar{\xi}_H)}{I_{-H} C_H \beta_i^H}} - 1 \right) \\ & + \frac{I_{-H} \xi_H}{T_{iH} (C_H + \xi_H)}, \end{aligned} \qquad (11)$$

and $p_{iH}^* = \pi_{iH}(p_{Hi}^*)$ and $p_{ij}^* = \pi_{ij}(p_{ij})$ for $j \neq H, j \in \mathscr{N}_i$.

Proof: The problem described in (13) is a convex program with the utility function U_H convex in p_{Hi} and convex set $\mathscr{F}$. Hence the first-order optimality condition yields

$$C_H I_{-H} T_{iH} = \beta_i^H (I_{-H} + p_{Hi} T_{iH})(C_H p_{Hi} T_{iH} + (I_{-H} + p_{Hi} T_{iH}) \bar{\xi}_H), \qquad (12)$$

which is a quadratic equation to be solved for p_{Hi} and its nonnegative solution of (12) is given in (11). Since $p_{Hi,\max}$ is chosen sufficiently large and (11) is non-negative, p_{iH}^* is a feasible solution. The equilibrium solution for bot i hence follows from (9). ■

In order to provide insights into the solution obtained in (11), we make the following assumptions based on common structures of the botnets.

(A3) The real bots controlled by C&C bot i have identical features, i.e., $c_{ij} = \bar{c}_i$, $p_{ij} = \bar{p}_i$ and $T_{ij} = \bar{T}_i$ for all $j \neq H, j \in \mathscr{N}_i$.

(A4) The size of the real bots controlled by C&C bot i is much larger than the size of honeybots.

(A5) We let $\bar{\xi}_H = 0$.

Assumption (A5) is valid due to the freedom of choosing parameter ξ_H in (HOP). Without loss of generality, we can let $\xi_H = \frac{1}{\alpha_i}$ and hence $\bar{\xi}_H = 0$. Assumption (A3) holds if real bots controlled by C&C bot i are of the same type, for example, Windows non-expert Facebook users. These type of users are commonly the target of botnets. Under (A3), we can simplify the expressions in (11) and obtain $C_H = \frac{C_i}{c_i} + \frac{n_i}{\alpha_i}$, $I_{-H} = n_i^B \bar{T}_i \bar{p}_i$.

Assumption (A4) is built upon the fact that one C&C node in botnets often controls thousands of bots and the size of honeybots are often comparably small due to their implementation costs [24]. Under (A2), we have $I_{-H} \gg p_{Hi} T_{iH}$, then (13) can be rewritten as

$$\tilde{U}_H(\pi_{iH}(p_{Hi}), p_{Hi}) = \ln\left(C_H\left(\frac{T_{iH} p_{Hi}}{I_{-H}}\right) + \bar{\xi}_H\right) - \beta_i^H p_{Hi}. \tag{13}$$

Corollary 1: Under Assumptions (A1), (A2) and (A4), the Stackelberg equilibrium solution $(\mathbf{p}_i^*,\ p_{Hi}^*)$ of the game Ξ_S is given by

$$p_{Hi}^* = \left(\frac{1}{\beta_i^H} - \frac{I_{-H}\xi_H}{C_H T_{iH} + T_{iH}\xi_H}\right)^+, \tag{14}$$

where $(\cdot)^+ = \max\{0, \cdot\}$; $p_{iH}^* = \pi_{iH}(p_{Hi}^*)$ and $p_{ij}^* = \pi_{ij}(p_{ij})$ for $j \neq H, j \in \mathcal{N}_i$.

Proof: From (A4), we can rewrite (12) by replacing $I_{-H} + p_{Hi} T_{iH}$ with I_{-H}. Since all the terms in (14) is bounded, we can let $p_{Hi,\max}$ be sufficiently large and arrive at (14). The result then follows from Theorem 3. ■

Corollary 2: Let the size of real bots under C&C be n_i^B and the size of the honeybots represented by super node H n_i^H. Note that $n_i = n_i^B + n_i^H$. Under Assumptions (A1) - (A5), the Stackelberg equilibrium of the game Ξ_S is given by

$$p_{Hi}^* = \frac{1}{\beta_i^H}, \quad p_{ij}^* = \pi_{ij}(p_{ij}), \tag{15}$$

for $j \neq H, j \in \mathcal{N}_i$, and the equilibrium solution of C&C node i is composed of two terms given by $p_{iH}^* = p_{iH,S}^* + p_{iH,N}^*$, with the first term independent of n_i^H,

$$p_{iH,S}^* = \frac{T_{iH}}{T_{iH} + \beta_i^H n_i^B \bar{T}_i \bar{p}_i}\left(\frac{C_i}{c_i} + \frac{n_i^B}{\alpha_i}\right) - \frac{1}{\alpha_i}, \tag{16}$$

and the second term dependent on n_i^H,

$$p_{iH,S}^* = \frac{n_i^H T_{iH}}{T_{iH} + \beta_i^H n_i^B \bar{T}_i \bar{p}_i}. \tag{17}$$

Proof: The result immediately follows from Corollary 1 using (A3) and (A5). ■

Remark 1: From Corollary 2, we can see that under Assumption (A1), the equilibrium response strategy is inversely proportional to the unit cost β_i^H. We can see that the number of command and control messages harvested from the botnet is affine in the number of successfully infiltrated honeybots. The growth rate of the number of messages is given by

$$r_{iH}^* := \frac{\partial p_{iH}^*}{\partial n_i^H} = \frac{T_{iH}}{\beta_i^H n_i^B \bar{T}_i \bar{p}_i + T_{iH}}. \tag{18}$$

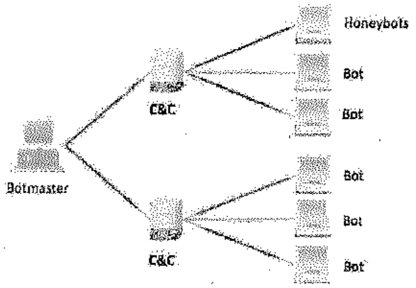


Fig. 4. Illustration of a hierarchical social botnet

The growth rate is dependent of the trust value T_{iH}. Honeybots can harvest more information from the botnet if they are more trusted. The growth rate is also dependent on the number of the real bots controlled by C&C bot i. As $n_i^B \to \infty$, the growth rate $r_{iH}^* \to 0$, i.e., size of honeybots will not affect the number of messages received by the network.

Trust values can change over time and can either modeled by a random process or by some assessment rules adopted by the attacker. We can separate this into different subsections of discussion. We can also consider a dynamic optimization problem as well by having belief/trust as the state. This can be done through using beta or Dirichlet distributions.

V. MODEL OF HONEYBOT DEPLOYMENT AND BOTNET GROWTH

In what follows, a macroscopic model of the dynamics of the number of bots at time t, denoted $x_1(t)$, and the number of honeybots, denoted $x_2(t)$, is presented. We then formulate an optimization problem for determining the number of honeypot nodes to introduce into the network.

A. *Botnet and honeypot growth models*

The bots are assumed to send spam messages, containing links to malware, with rate r. Each message is sent to each of the d neighbors of the bot, where d is the average node degree. Hence in each time interval dt, $rd\ dt$ spam messages are sent. Since the number of valid nodes is $N - x_1(t)$, the number of messages reaching valid nodes is equal to $rd\frac{N - x_1(t)}{N}\ dt$.

The number of nodes that become bots depends on the behavior of the valid users and the number of links that have been blacklisted. Each user clicks on a spam link with probability q. If the link has been blacklisted, then the user will be blocked from visiting the infected site; otherwise, the user's account is compromised and the device becomes part of the botnet.

To determine the probability that a link has been blacklisted, we assume that each bot is independently given a set of k malicious links, out of M links total. The probability that a given link has been given to a specific honeybot is therefore $\frac{k}{M}$. Hence the probability that a link has not been blacklisted is the probability that that link has not been given to any honeybot, which is equal to $\left(1 - \frac{k}{M}\right)^{x_2}$. We assume that:

(A6) The number of links given to each honeybot, k, satisfies $k \ll M$.

Under (A6), $(1-\frac{k}{M})^{x_2}$ can be approximated by $\left(1-\frac{kx_2}{M}\right)$.

Finally, we assume that the infected devices are discovered and cleaned with rate μ_1. This leads to dynamics

$$\dot{x}_1(t) = rdqx_1\left(1-\frac{kx_2}{M}\right)\frac{N-x_1}{N} - \mu_1 x_1. \tag{19}$$

Honeybot nodes are inducted into the botnet in a similar fashion. We make the following assumptions regarding the honeybot population:

(A7) The number of honeybots that are not part of the botnet, denoted z, is constant.

(A8) The number of honeybots is small compared to the total number of users, so that $\frac{z}{z+N} \approx \frac{z}{N}$.

Assumption (A7) can be guaranteed by creating new, uninfected honeybots when existing honeybots infiltrate the botnet. As with real users, honeybot nodes cannot follow links that have been blacklisted; however, unlike real users, honeybot nodes will attempt to follow any link with probability 1. The botmaster detects and removes honeybots with rate μ_2. The honeybot population is therefore defined by

$$\dot{x}_2(t) = rdx_1\left(1-\frac{kx_2}{M}\right)\frac{z}{N} - \mu_2 x_2. \tag{20}$$

Proposition 1: The dynamics defined by (19) and (20) have two equilibria, given by $(x_1,x_2) = (0,0)$ and

$$x_1^* = \frac{N\mu_2 M(rdq-\mu_1)}{rdq\mu_2 M + rdkz\mu_1}, \quad x_2^* = \frac{\left(rd-\frac{\mu_1}{q}\right)z}{\frac{rdkz}{M}+\mu_2}. \tag{21}$$

Proof: Equation (19) reaches equilibrium if $x_1 = 0$ or if $rdq\left(1-\frac{kx_2}{M}\right)\frac{N-x_1}{N} - \mu_1 = 0$. If $x_1 = 0$, then (20) reaches equilibrium when $x_2 = 0$.

If $rdq\left(1-\frac{kx_2}{M}\right)\frac{N-x_1}{N} - \mu_1 = 0$, solving for x_1 yields $x_1 = N\left(1-\frac{\mu_1}{rdq\left(1-\frac{kx_2}{M}\right)}\right)$. Substituting into (20) gives the equilibrium of (21). ■

The quantity rdq corresponds to the rate at which new nodes are inducted into the botnet, while μ_1 is the rate at which nodes are cleaned and exit the botnet. Thus if $rdq < \mu_1$, then the number of bots converges to zero, while $rdq > \mu_1$ implies that the number of bots converges to a nonzero steady-state value.

Since network security policies are typically updated intermittently, while the dynamics of (19) and (20) converge rapidly, we base our subsequent analysis on the steady-state values of x_1 and x_2, and derive the optimal number of honeybots to introduce into the system in steady-state. In order to prove that this problem is well-defined, we first examine the stability properties of each equilibrium in the following theorem.

Theorem 4: If $\mu_1 > rdq$, then $(x_1,x_2) = (0,0)$ is asymptotically stable. If $\mu_1 < rdq$, then (x_1^*,x_2^*) is asymptotically stable in the limit as $M \rightarrow \infty$, $N \rightarrow \infty$.

Proof: An equilibrium point of a nonlinear dynamical system is asymptotically stable if its linearization is asymptotically stable at that point [25, Theorem 3.7]. The linearization of (19) and (20) around (0,0) is given by

$$A_{00} = \begin{pmatrix} rdq-\mu_1 & 0 \\ \frac{rdz}{N} & -\mu_2 \end{pmatrix}.$$

If $\mu_1 > rdq$, then $-A_{00}$ is diagonally dominant, and hence has eigenvalues with positive real part [ref]. The eigenvalues of A_{00} therefore have negative real part, implying that the linearization around $(0,0)$ is asymptotically stable. The linearization $A_{x_1^*x_2^*}$ around (x_1^*,x_2^*) is given by

$$A_{x_1^*x_2^*} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \text{ where}$$

$$a_{11} = \mu_1 - \frac{rdq}{N}\left(\frac{\mu_2 M + \frac{\mu_1 kz}{q}}{\mu_2 M + rdkz}\right) \cdot \left(-N + \frac{2\mu_1 rdkz + \mu_2 M}{rdq\mu_2 M + rdkz\mu_1}\right),$$

$$a_{12} = -\frac{rdqkN}{M}\left(1-\frac{\mu_1 rdkz + \mu_2 M}{rdq\mu_2 M + rdkz\mu_1}\right) \cdot \left(\frac{\mu_1 rdkz + \mu_2 M}{rdq\mu_2 M + rdkz\mu_1}\right),$$

$$a_{21} = \frac{rdz}{N}\left(\frac{\mu_2 M + \mu_1 kz/q}{\mu_2 M + rdkz}\right),$$

$$a_{22} = -\frac{rdkz}{M}\left(1-\frac{\mu_1 rdkz + \mu_2 M}{rdq\mu_2 M + rdkz\mu_1}\right) - \mu_2.$$

To prove that $A_{x_1^*x_2^*}$ has eigenvalues with negative real part, we examine $-A_{x_1^*x_2^*}$. The second row is clearly diagonally dominant, since the diagonal element is positive and the off-diagonal element is negative. The first row is diagonally dominant if

$$-\frac{rdqkN}{M}\left(1-\frac{\mu_1 rdkz + \mu_2 M}{rdq\mu_2 M + rdkz\mu_1}\right)\left(\frac{\mu_1 rdkz + \mu_2 M}{rdq\mu_2 M + rdkz\mu_1}\right) < \mu_1 + rdq\left(\frac{\mu_2 M + \frac{\mu_1 kz}{q}}{\mu_2 M + rdkz} - \frac{2\mu_1 rdkz + \mu_2 M}{rdq\mu_2 M + rdkz\mu_1}\right). \tag{22}$$

In the limit as $M \rightarrow \infty$, the left-hand side of (22) converges to zero while the right-hand side reduces to $\mu_1 + rdq\left(1-\frac{1}{rdqN}\right)$, which is positive for N sufficiently large. Hence $-A_{x_1^*x_2^*}$ is diagonally dominant, and therefore has eigenvalues with positive real part, implying that (x_1^*,x_2^*) is a stable equilibrium point. ■

B. Computation of system parameters

The parameter μ_2 determines the rate at which honeypot nodes are discovered and removed by the botmaster, and hence can be calculated by observing the lifetime of deployed honeypots (see Section VI). Similarly, the number of received messages p and the cost τ can be estimated by averaging over the set of deployed honeypots over time. The fraction of malicious links $\frac{k}{M}$ that are given to a single bot or honeypot is estimated by using the assumption that links are distributed independently and uniformly at random by the botmaster, so that the probability that a given link has been received by a honeypot is $\left(1-\frac{k}{M}\right)^{x_2}$. This probability can be estimated by analyzing the set of malicious links received by new honeypots, which combined with knowledge of x_2 enables computation of $\frac{k}{M}$. The rate at which spam messages are sent by bots, denoted r, is estimated by the number of instruction messages received by the honeypots.

The parameters μ_1 and q, equal to the rate at which bots are removed from the botnet, and the fraction of malicious links that are followed by users, depend on user behavior. These parameters can be estimated using existing data sets of user behavior [26]. Furthermore, to obtain an upper bound on the effectiveness of the botnet, the parameter q can be set equal to 1, implying that a valid user always clicks any link to the malware executable (the worst case). The average node degree, d, is estimated based on existing analyses of the degree distribution of social networks [27].

C. *Extension to heterogeneous networks*

Typical social networks follow a non-uniform degree distribution. We present a model for the bot and honeypot population dynamics as follows. Let N_d denote the total number of users with degree d, and let $x_1^d(t)$ and $x_2^d(t)$ denote the number of bots and honeypots with degree d at time t. The average degree of the network is equal to $\bar{d}$. We make the assumption that:

(A9) The average degree of the infected users is equal to the average degree of the social network as a whole.

The total number of spam messages sent by bots in time interval dt is equal to $r\bar{d}x_1\, dt$, each of which has not been blacklisted with probability $\left(1-\frac{kx_2}{M}\right)$. The probability that the recipient of a message has degree d and has not been infected is equal to

$$Pr(\text{degree } d, \text{ not infected}) = \frac{N_d - x_1^d}{N_d}\frac{dN_d}{\bar{d}N} = \frac{(N_d - x_1^d)d}{\bar{d}N}.$$

This implies that the dynamics of $x_1^d(t)$ are given by

$$\dot{x}_1^d(t) = \frac{rdq}{N}x_1\left(1-\frac{kx_2}{M}\right)(N_d - x_1^d) - \mu_1 x_1^d, \qquad (23)$$

where N_d is the number of user accounts of degree d.

Similarly, the probability that the recipient of a spam message is a honeypot node of degree d that has not been infected yet is $\frac{z_d^*}{N}$, where z_d is the number of honeybots of degree d that have not joined the botnet, leading to dynamics of $x_2^d(t)$ described by

$$\dot{x}_2^d(t) = \frac{rd}{N}x_1\left(1-\frac{kx_2}{M}\right)z_d^* - \mu_2 x_2^d. \qquad (24)$$

Proposition 2: The dynamics (23) and (24) have equilibrium points at $x_1^d = x_2^d = 0$ for all d and

$$x_1^{d*} = \left[rdq\left(1-\frac{kr\bar{d}_z\left(1-\frac{\mu_1}{r\bar{d}q}\right)z}{rkzd_z+\mu_2 M} - \frac{\mu_1}{r\bar{d}q}\right)+\mu_1\right]^{-1} \qquad (25)$$
$$\cdot\left(rdq\left(1-\frac{kr\bar{d}_z\left(1-\frac{\mu_1}{r\bar{d}q}\right)z}{rkzd_z+\mu_2 M} - \frac{\mu_1}{r\bar{d}q}\right)\right),$$

$$x_2^{d*} = \frac{rd}{\mu_2}\left(1-\frac{kr\bar{d}_Z\left(1-\frac{\mu_1}{r\bar{d}q}\right)z}{rkzd_Z+\mu_2 M} - \frac{\mu_1}{r\bar{d}q}\right)z_d. \qquad (26)$$

Proof: Summing (23) over d yields

$$\dot{x}_1(t) = \frac{x_1\bar{d}rq}{N}\left(1-\frac{kx_2}{M}\right)(N-x_1) - \mu_1 x_1,$$

which implies that in steady-state we have

$$x_1^* = N\left(1-\frac{\mu_1}{r\bar{d}q\left(1-\frac{kx_2}{M}\right)}\right). \qquad (27)$$

Similarly, summing $x_2^d(t)$ over d results in

$$\dot{x}_2(t) = \frac{rx_1}{N}\left(1-\frac{kx_2}{M}\right)z\bar{d}_Z - \mu_2 x_2,$$

which combined with (27) gives $x_2^* = \frac{r\bar{d}_Z\left(1-\frac{\mu_1}{r\bar{d}q}\right)}{\frac{rkzd_Z}{M}+\mu_2}$. The steady-state value (26) can be obtained from (24). Similarly, x_1^{d*} can be obtained from (23). ■

VI. MODELING OF PROTECTION AND ALERT SYSTEM (PAS)

PAS is a coordination system that strategically deploys honeybots and designs security policies for social networks. In this section, we focus on optimal reconfiguration of honeybots as illustrated in Fig. 2. We introduce a mathematical framework for finding honeybot deployment strategies based on system models described in Sections IV and V.

A. *Relations between HD and HE*

We have adopted a divide-and-conquer approach in Sections IV and V, and have modeled the behavior of each system independently. However, the interdependencies between HD and HE are essential for PAS to make optimal security policies for the social network. The HE model in Section IV describes strategic operations of honeybots at a microscopic level while the HD model in Section V provides a macroscopic description of the population dynamics of bots and honeybots. These two models are interrelated through their parameters together with the feedback control from PAS.

The interactions between bots and honeybots in the HE model occur on a time scale of seconds. The analysis of Stackelberg equilibrium in Section IV captures the steady-state equilibrium after a repeated or learning process of the game. Hence the equilibrium can be reached on a time scale of minutes. On the other hand, the population dynamics in HD model evolve on a larger time scale (for example, days). Hence, we can assume that the Stackelberg game has reached its equilibrium when the populations evolve at a macroscopic level. Decisions made at PAS are on a longer time scale (for example, weeks) because the processing of collected information, learning of bots and honeybots in social networks, and high-level decision on security policy in reality demand considerable amount of human resources for coordination and supervision.

1) Trust Values and Detection Rate: The trust values T_{ij} used in HE model are related to the macroscopic detection and removal rate μ_2 in HD model. As we have pointed out earlier, zero trust values are equivalent to the removal of honeybots from the botnet. Hence we can adopt a simple dynamic model to describe the change of T_{ij} over a longer time period (say, days). We let T_{ij}^0 be the initial condition of the trust value. The evolution of T_{ij} over the macroscopic time scale can be modeled using the following ODE:

$$T_{ij}(t) = -\mu_2 T_{ij}(t), \quad T_{ij}(t_{ij}^0) = T_{ij}^0. \tag{28}$$

Note that honeybots have different initial time t_{ij}^0. Hence from (28), we obtain

$$T_{ij}(t) = e^{-\mu_2(t-t_{ij}^0)}, \quad t \geq t_{ij}^0, \tag{29}$$

i.e., the trust values exponentially decay with respect to the removal rate. a threshold can be set on. From (29), we can obtain the mean life time of a honeybot is $1/\mu_2$. Macroscopic parameter μ_2 can be estimated by the rate of change of working honeybots in the botnet, which can is known to the system, while T_{ij} is a microscopic parameter and is often unknown directly to honeybots. With the ODE model in (28), we can use μ_2 to estimate T_{ij}.

2) Honeybot and Bot Populations: In Section V, the populations of bots and infiltrating honeybots are denoted by x_1 and x_2, respectively, whereas in Section IV, the bot size under C&C bot is n_i^B. Under a hierarchical structure of botnet, illustrated in Fig. 4, the total bot and honeybot populations x_1, x_2 are given by

$$x_1 = \sum_{i=1}^{m} n_i^B, \quad x_2 = \sum_{i=1}^{m} n_i^H. \tag{30}$$

If all C&C bots are assumed to be identical, i.e., $n_i^B = \bar{n}^B, i \in \mathscr{M}, n_i^H = \bar{n}^H, i \in \mathscr{M}$, then $x_i = m\bar{n}^B$, and $x_i = m\bar{n}^H$

3) Activity Level of Bots: The rate $\bar{p}_i$ in (18) indicates the activity level of bots when they respond to or poll information from C&C node i. This level of activity is often correlated with parameter r, the rate of sending out spamming messages to the social network. Assume that all C&C bots are assumed to be identical, i.e., $\bar{p}_i = \bar{p}, i \in \mathscr{M}$, then we can let $\bar{p} = \alpha r$, where $\bar{p}$ is in messages/sec, r is in messages/sec and $\alpha \in \mathbb{R}_{++}$ is a unitless positive parameter.

B. Cross-Layer Optimal Honeybot Deployment

In what follows, we first derive the optimal honeybot deployment when the benefit from each honeypot is measurable. We then combine the analysis of Sections IV and V to determine the optimal honeybot deployment, taking into account the behavior of the deployed nodes during the exploitation phase.

The goal of the honeypot operator is to maximize the number of blacklisted links that are reported to the social network. Based on the analysis of Corollary 2, we assume that the number of blacklisted links is proportional to the number of honeypot nodes in the botnet in steady-state, $\mathbf{x}_2^*$. The variable is the number of honeypot nodes that have not yet been inducted into the botnet, z. This leads to a utility function given by $V_H(z) = px_2^*(z) - \tau(x_2^* + z)$, where p and τ represent the benefit (information gathered) and cost of maintaining a single honeypot node. Substituting (21) yields

$$\begin{aligned} V_H(z) &= p\frac{\left(rd - \frac{\mu_1}{q}\right)z}{\frac{rdkz}{M} + \mu_2} - \tau\left(\frac{\left(rd - \frac{\mu_1}{q}\right)z}{\frac{rdkz}{M} + \mu_2} + z\right) \\ &= \frac{(p-\tau)\left(rd - \frac{\mu_1}{q}\right)z}{\frac{rdkz}{M} + \mu_2} - \tau z. \end{aligned} \tag{31}$$

The value of z that maximizes (31) is given by the following proposition.

Proposition 3: The optimum value of z is given by

$$z^* = M\left(\frac{-\mu_2 + \sqrt{(p-\tau)(rd - \frac{\mu_1}{q})\mu_2/\tau}}{rdk}\right). \tag{32}$$

Proof: Differentiating $V_H(z)$ with respect to z yields

$$\frac{\partial V_H}{\partial z} = \frac{(p-\tau)\left(rd - \frac{\mu_1}{q}\right)\mu_2}{\left(\frac{rdk}{M}z + \mu_2\right)^2} - \tau.$$

By inspection, $\frac{\partial V_H}{\partial z}$ is a strictly decreasing function of z, so that $V_H(z)$ is strictly concave. Setting this expression equal to zero implies (32). ■

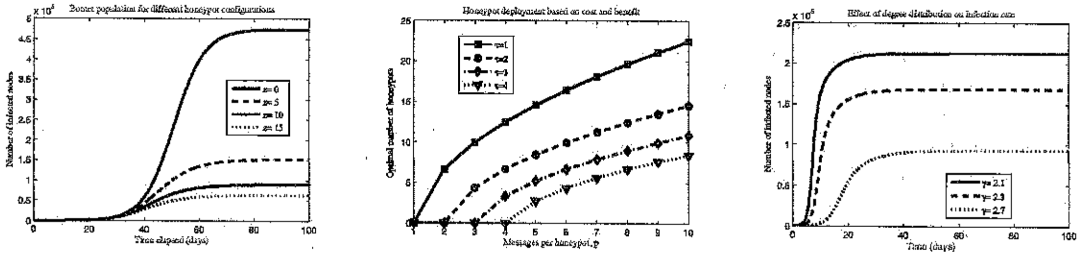
Remark 2: Eq. (32) has several implications for the design of honeybot systems. First, for malware that propagates rapidly (corresponding to a large rd value), fewer honeybots are needed, since the malware will quickly spread to the deployed honeybot. Second, if μ_2 is large, then honeybots are rapidly detected and removed by the botmaster, and hence the cost of deploying honeybots outweighs the benefits.

The utility function (31) can be augmented by incorporating the impact on the exploitation phase. In particular, (18) implies that $p = \frac{1}{1+\frac{\beta_i^H x_1^* T_{i'r}}{T_{iH}}}$, which we write as $p = \frac{1}{1+\zeta x_1^*} \approx \frac{1}{\zeta x_1^*}$ when the number of bots is sufficiently large. The utility function V_H can then be written as

$$V_H = \left(\frac{1}{\zeta x_1^*} - \tau\right)x_2^* - \tau z \tag{33}$$

An efficient algorithm for maximizing (33) can be derived using the following theorem.

Theorem 5: The problem of selecting z to maximize V_H



(a) (b) (c)

Fig. 5. Simulation of our framework for a network of $N=10^6$ users, where each user has probability $q=0.01$ of following a malicious link, message are sent at a rate of 0.4 messages per bot per day, and infected nodes are cleaned after 5 days on average. (a) Effect of increasing the number of honeypot nodes on the botnet population. Deployment of a small number of honeypots can greatly reduce the number of bots present. Note that the population converges quickly to its equilibrium value. (b) The optimum number of bots based on (32) for different costs τ and benefits p. The total number of honeypots remains small for each case. (c) Effect of degree distribution on the botnet population for number of honeypots $z=5$. Each network is scale-free, with exponent γ varying between networks. A higher connectivity results in a larger number of bots.

in (33) is equivalent to

$$\begin{aligned}
\max_{\theta,\phi,x_2^*,z} \ & \frac{1}{\zeta}\left(-\frac{rdq\theta^2}{Nrdq\phi-\mu_1}+\frac{M/4k}{N\left(rdq\left(1-\frac{kx_2^*}{M}\right)-\mu_1\right)}\right) \\
& -\tau x_2^* - \tau z, && (34)\\
\text{s.t.} \quad & \theta = x_2^* - \frac{M}{2k}, \\
& \phi = 1-\frac{kx_2^*}{M}, && (35)\\
& x_2^* \leq \frac{\left(rd-\frac{\mu_1}{q}\right)z}{\frac{rdkz}{M}+\mu_2}, && (36)\\
& \frac{1}{\zeta}\frac{rdq\mu_2M+rdkz\mu_1}{(rdq-\mu_1)\mu_2M} \geq \tau, \\
& z \geq 0, \quad 0 \leq x_2^* \leq N,
\end{aligned}$$

which is a convex program.

Proof: The optimization problem of selecting z to maximize V_H can be written as

$$\begin{array}{ll}\text{maximize} & \left(\frac{1}{\zeta x_1^*(z)}-\tau\right)x_2^*(z)-\tau z \\ z\in\mathbb{R}_+ & \end{array} \qquad (37)$$

If $\frac{1}{\zeta x_1^*(z)}-\tau<0$, then the objective function is monotone decreasing in z, leading to an optimal value $z=0$. To avoid this, we require $\frac{1}{\zeta x_1^*(z)}\geq\tau$, leading to constraint (37). Using (21), we have

$$\frac{x_2^*}{\zeta x_1^*}=\frac{\left(x_2^*-\frac{kx_2^{*2}}{M}\right)rdq}{N\left(rdq\left(1-\frac{kx_2^*}{M}\right)-\mu_1\right)}=\frac{rdq\left(-\frac{k}{M}\left(x_2^*-\frac{M}{2k}\right)^2+\frac{M}{4k}\right)}{N\left(rdq\left(1-\frac{kx_2^*}{M}\right)-\mu_1\right)}. \qquad (38)$$

Substituting $\theta=\left(x_2^*-\frac{M}{2k}\right)$ and $\phi=\left(1-\frac{kx_2^*}{M}\right)$ leads to the objective function

$$V_H(x_2^*,z,\theta,\phi)=\frac{1}{\zeta}\left(-\frac{rdq\theta^2}{Nrdq\phi-\mu_1}+\frac{M/4k}{N\left(rdq\left(1-\frac{kx_2^*}{M}\right)-\mu_1\right)}\right)$$
$$-\tau x_2^*-\tau z \qquad (39)$$

Since quadratic over linear and inverse functions are convex, the first two terms of (39) are concave, and hence V_H is concave.

Finally, the fact that the objective function is increasing as a function of x_2^* and decreasing as a function of z implies that the constraint (36) holds with equality at the optimum, so that the relationship between x_2^* and z in (21) is satisfied. This constraint is convex due to Proposition 3. ■

The convex optimization approach presented in Theorem 5 is used to select a honeybot deployment strategy in order to maximize the level of infiltration into the botnet and the amount of data gathered during the exploitation phase. Once inducted into the botnet, the honeybots follow the Stackelberg equilibrium strategy of Section IV and use the collected data to generate malware signatures and create URL blacklists. The parameters of (33) are updated in response to changes in botnet behavior observed during the exploitation phase.

VII. SIMULATION STUDY

We evaluated our proposed method using Matlab simulation study, described as follows. A network consisting of $N=10^6$ nodes was generated, with degree $d=100$ (consistent with observations of the average degree of social networks [27]). The rate at which malware messages are sent is given by $r=0.4$ messages per bot day, and the rate at which nodes are disinfected and removed from the botnet is $\mu_1=0.2$, an average lifetime for each bot of 5 days. These statistics are based on the empirical observations of [4]. Based on [5], we estimate that the probability of a user clicking on a spam link is given by $q=0.01$. It is assumed that the fraction of malware links given to each bot is equal to $k/M=0.01$. The rate at which honeybots are detected and removed is equal to $\mu_2=0.5$. In each case, we assume that there are 50 infected nodes and 0 honeybots present in the network initially.

The population dynamics of the bots, described by (19) and (20), are shown in Fig. 5(a). Each curve represents the number of infected users over time for a different level of honeybot activity, as described by the parameter z. In each case, the number of bots converges to its equilibrium value. The top curve (solid line) assumes $z=0$, i.e. no deception takes place and malicious links are detected through blacklists only. Employing deception through honeybots significantly reduces the botnet population, even when the number of honeybots is small relative to the population size. As additional honeybots are added, the botnet population continues to decline. However, the marginal benefit of adding a honeybot decreases as the number of honeybots grows large.

The optimum number of honeybots depends on the cost of introducing and maintaining honeybots, denoted τ, as well as the benefit p from each honeybot, as described in (32). The optimum nnmber of honeybots is given in Fig. 5(b). As the cost of introducing new honeybots is reduced, the optimal number of honeybots increases. In each case, the optimum number of honeybots remains small, at around 25 nodes, relative to the total network population of 10^6 nodes.

The effect of a heterogeneous degree distribution is shown in Fig. 5(c). The degree distribution was chosen to be scale-free, so that the probability that a node has degree d was proportional to $d^{-\gamma}$. Hence a higher value of γ corresponds to a less-connected network. The parameter γ had a significant impact on the rate of propagation of the botnet, even through for the chosen values of γ the average degrees of the three networks were similar.

VIII. CONCLUSION

In this paper, we studied the problem of defending against social botnet attacks through deception. We considered a defense mechanism in which fake honeybot accounts are deployed and infiltrate the botnet, impersonating infected users. The infiltrating honeybots gather information from command and control messages, which are used to form malware signatures or add spam links to URL blacklists. We introduced a framework for SOcial network Deception and EXploitation through hOneybots (SODEXO), which provides an analytical approach to modeling and designing social honeybot defenses. We decomposed SODEXO into deployment and exploitation components.

In the deployment component, we model the population dynamics of the infected users and honeybots, and show how the infected population is affected by the number of honeybots introduced. We derive the steady-state populations of infected users and honeybots and prove the stability of the equilibrium point. In the exploitation component, we formulate a Stackelberg game between the botmaster and the honeybots and determine the amount of information gathered by the honeybot in equilibrium. The two components are combined in the Protection and Alert System (PAS), which chooses an optimal deployment strategy based on the observed behavior of the botnet and the information gathered by the honeybots. Our results are supported by simulation studies, which show that a small number of honeybots significantly decrease the infected population of a large social network.

REFERENCES

[1] J. Baltazar, J. Costoya, and R. Flores, "The real face of koobface: The largest web 2.0 botnet explained," *Trend Micro Research*, vol. 5, no. 9, p. 10, 2009.

[2] G. Keizer, "Worm spreads on facebook, hijacks users' clicks," *Computerworld*, December 2008.

[3] "Twitter FAQ for developers." `https://dev.twitter.com/docs/faq`.

[4] K. Thomas and D. Nicol, "The Koobface botnet and the rise of social malware," in *5th International Conference on Malicious and Unwanted Software (MALWARE)*, 2010.

[5] C. Grier, K. Thomas, V. Paxson, and M. Zhang, "@spam: the underground on 140 characters or less," in *Proc. of the 17th ACM conference on Computer and communications security*, pp. 27–37, 2010.

[6] K. Lee, J. Caverlee, and S. Webb, "Uncovering social spammers: social honeypots + machine learning," in *Proc. of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, pp. 435–442, 2010.

[7] J. W. Caddell, *Deception 101 - Primer on Deception.* Strategic Studies Institute, 2004.

[8] "The Department of Defense Dictionary of Military and Associated Terms,"

[9] J. Yuill, D. Denning, and F. Feer, "Using deception to hide things from hackers: Processes, principles, and techniques," *Journal of Information Warfare*, vol. 5, no. 3, pp. 26–40, 2006.

[10] J. John, A. Moshchuk, S. Gribble, and A. Krishnamurthy, "Studying spamming botnets using botlab," in *Proc. of the 6th USENIX symposium on Networked systems design and implementation*, pp. 291–306, 2009.

[11] G. Portokalidis, A. Slowinska, and H. Bos, "Argos: an emulator for fingerprinting zero-day attacks for advertised honeypots with automatic signature generation," *SIGOPS Oper. Syst. Rev.*, vol. 40, pp. 15–27, Apr. 2006.

[12] M. McQueen and W. Boyer, "Deception used for cyber defense of control systems," in *Human System Interactions, 2009. HSI '09. 2nd Conference on*, pp. 624 –631, may 2009.

[13] M. Manshaei, Q. Zhu, T. Alpcan, T. Başar, and J.-P. Hubaux, "Game theory meets network security and privacy," *ACM Computing Survey*, vol. 45, September 2013 (to appear).

[14] M. H. R.Khouzani, S. Sarkar, and E. Altman, "Maximum damage malware attack in mobile wireless networks," in *Proc. of the 29th conference on Information communications (INFOCOM'10), San Diego, California, USA*, 2010.

[15] Q. Zhu, C. Fung, R. Boutaba, and T. Başar, "A game-theoretic approach to knowledge sharing in distributed collaborative intrusion detection networks: Fairness, incentives and security," in *Proc. of 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pp. 243–250, 2011.

[16] P. Wang, L. Wu, R. Cunningham, and C. C. Zou, "Honeypot detection in advanced botnet attacks," *International Journal of Information and Computer Security*, vol. 4, no. 1, pp. 30–51, 2010.

[17] C. C. Zou and R. Cunningham, "Honeypot-aware advanced botnet construction and maintenance," in *Proc. of the International Conference on Dependable Systems and Networks*, DSN '06, (Washington, DC, USA), pp. 199–208, IEEE Computer Society, 2006.

[18] G. Gu, J. Zhang, and W. Lee, "BotSniffer: Detecting botnet command and control channels in network traffic," in *Proc. of the 15th Annual Network and Distributed System Security Symposium (NDSS'08)*, February 2008.

[19] G. Gu, R. Perdisci, J. Zhang, and W. Lee, "BotMiner: Clustering analysis of network traffic for protocol- and structure-independent botnet detection," in *Proc. of the 17th USENIX Security Symposium (Security'08)*, 2008.

[20] Q. Zhu and T. Başar, "Dynamic policy-based ids configuration," in *Proc. of the 48th IEEE Conference on Decision and Control (CDC)*, pp. 8600 –8605, 2009.

[21] Q. Zhu and T. Başar, "Indices of power in optimal ids default configuration: theory and examples," in *Proc. of the Second international conference on Decision and Game Theory for Security*, GameSec'11, (Berlin, Heidelberg), pp. 7–21, Springer-Verlag, 2011.

[22] W. Ping, B. Aslam, and C. C. Zou, "Peer-to-peer botnets: The next generation of botnet attacks," in *Stavroulakis, P. and Stamp, M. (Eds), Handbook of Information and Communication Security*, (Washington, DC, USA), pp. 335–350, Springer Press, 2010.

[23] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory.* SIAM Series in Classics in Applied Mathematics, 1999.

[24] N. Provos and T. Holz, *Virtual honeypots: from botnet tracking to intrusion detection.* Addison-Wesley Professional, first ed., 2007.

[25] H. Khalil, *Nonlinear Systems.* Macmillan Publishing Company New York, 1992.

[26] C. Kanich, C. Kreibich, K. Levchenko, B. Enright, G. Voelker, V. Paxson, and S. Savage, "Spamalytics: An empirical analysis of spam marketing conversion," in *Proc. of the 15th ACM conference on Computer and communications security*, pp. 3–14, ACM, 2008.

[27] M. Gjoka, M. Kurant, C. Butts, and A. Markopoulou, "Practical recommendations on crawling online social networks," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 29, pp. 1872–1892, October 2011.

Game in the Newsroom: Greedy Bloggers for Picky Audience

Avner May*
avnermay@cs.columbia.edu,

Augustin Chaintreau*
augustin@cs.columbia.edu,

Nitish Korula
nitish@google.com

Silvio Lattanzi
silviol@google.com

1. INTRODUCTION

We rely increasingly on social networks, especially microblogging sites like Twitter and Weibo, to diffuse news information and inform our collective decisions. Every now and then social media *make the scoop* – breaking news is often tweeted before its widespread diffusion – and everyday they *make news viral* – so importantly that media outlets like the New York Times allow for unlimited access to articles from social media. Weblogs, originally a simple way to share recently viewed items on the web, now exhibit the intriguing properties of a global echo chamber that benefits everyone: Passive users embrace such collaborative curation, as recent reports showed that users landing on an article through social endorsements generate more advertising revenue through better engagement. A minority of contributors rise to prominent level, some of them generating their own advertising revenue as information brokers.

A major algorithmic problem is whether such networks make the most of a critical scarce resource: users' reading time. Indeed, intuitively the process of collaborative curation is offering them a diversity of offers. Instead of reading the full length of news published on a given topic, they can enjoy a selection, typically chosen by someone else to be the most noteworthy. Such curation may encourage more users to follow online news, and provide a potentially larger audience to each news item depending on its perceived quality. This process potentially increases online news audience and revenue. Indeed, the bloggers who participate in this curation implicitly receive an incentive to improve the system; generating a larger audience for their tweet feeds and blogs, even if they only serve as intermediaries, puts them in a strategic position and allows them to monetize their traffic. Here we address the following challenge: can the various interests of media outlets, blogosphere and readership be aligned for efficient information filtering. To initiate our study, we focus on a simple case where all news items are related to the same "topic", but covers it with various relative degrees of depth. Our main assumption is that users have different interest thresholds that determine the set of news that they should ideally receive. We present three main contributions:

- We analyze empirically a million tweets collected during one of the most active recent events of online discussion: the death of Osama Bin Laden. We exhibit and explain an apparent paradox: that on a given theme, the most active or prominent users tend to post articles that are *less* popular than occasional tweeters. (Section 2)
- Motivated by the rationale of information filtering, we propose a game representing strategic bloggers and readers with varying degrees of engagement. We study how competition can lead to stable and efficient outcomes, proving that "greedy" readers lead to instability but a bit of laziness on the reader side can be sufficient to ensure stability and efficiency. (Section 3)
- Motivated by these preliminary results, we discuss future steps to analyze and validate the processs of collaborative news curation. (Section 4)

Our work follows the classic hypothesis of a two-step flow of information [6], where opinion leaders that are also news savvy play a key role of intermediaries. This study was revived using empirical evidence of a similar effect occuring on Twitter [8], and more recently identifying mass media and intermediaries as critical to information spread in general [3]. Much remains to be explored to analyze the intrinsic effectiveness of such information spread: a recent study showed that social network allows users to be exposed to more diverse or equilibrate viewpoints [1]. We wish to go a step further, by proving that intermediaries emerge naturally from a collaborative curation ecosystem and we aim at informing design of efficient incentives for such information sharing systems. We are only aware of a few works analyzing news dissemination with a game theoretic background. In [4], a different strategic news posting is analyzed. It assumes a fixed network of followers or friends and the same interest perceived among users, proving that local incentive are sufficient for higher quality news to spread globally, but that this spread is limited when

---

*supported by NSF award ICES-1139915

users attempt to avoid spam. In [5], a multi-topic model is analyzed to model news aggregators choosing a subset of topics in a sequential game. We focus on information filtering and how this affects the formation of the follower relationship. Previous works have analyzed how to optimize attention for news, typically as a scheduling problem in the face of ephemerality [2, 7], but they typically consider the problem from the point of view of a single media outlet.

2. INFORMATION FILTERING EVIDENCE

Data was gathered using `DiscoverText.com` to harvest tweets from the Twitter public API. This set contains 1,977,716 tweets containing the words "Osama" or "Bin Laden" posted on the night and following day[1] after the raid of the U.S. Navy. Based on self-reported numbers from twitter's traffic at this time[2], we estimate that this represents a sample of at least 5% of the tweets generated during that period. These contain a total of 950,950 unique URLs posted by 787,727 users. We observe varying *URL's popularity* (half of the URLs were posted only once; a few of them are present in a thousand tweets; 835 were cited at least 100 times) as well as *Users' activity* (half of the users post only once; two of them, both legitimate users, posted more than a thousand URLs; 855 users posted a hundred URLs).

The main merit of this dataset is that it deals with one particular topic while providing a dynamic range of depth, from blockbuster URLs (*e.g.*, President Obama's speech video) probably of interest to everyone, to obscure niche developments that may only appeal to users highly interested in following this event. One intuition posits that prominent or institutional bloggers (who generally post more) are likely to select "good" content and see their URLs posted more often. We see the exact opposite: Figure 1 distinguishes URLs posted by users with different activity levels and plots their distribution popularity. We observe a smooth transition, in which URLs posted by users who post less often are usually more popular. Two possible effects may explain this apparent paradox: (1) It might be that users behave the same but those who are more active simply *run out* of popular content. A comparison with a null-hypothesis model – where a user's set of post is picked with the same popularity distribution, but with different sizes – proved that this depletion effect is too weak to explain the observed trend. (2) Alternatively, it means that users post URLs *selectively*, and that the ones who post less typically favor URLs that are of broad interest.

This result motivates us to study how filtering emerges and improves information sharing: In choosing who to follow, a reader may actually decide the level of details that she will receive on this topic.

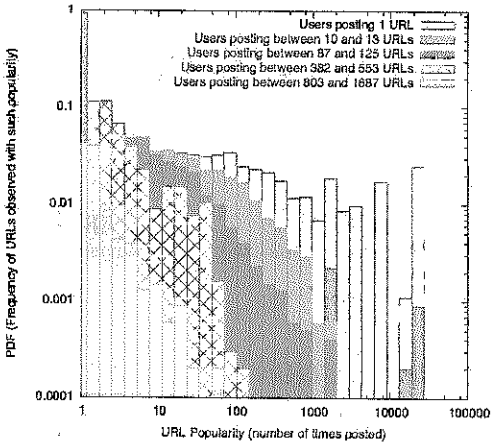


Figure 1: Histogram of tweets popularity and various posting activity levels. For one given theme, here the death of Osama Bin Laden, the less a user tweet, the more it is likely to share popular URLs disproportionally.

3. MODEL

Recall that we focus on content on a single topic. We assume that each item of content c has an inherent quality $q_c \in [0, 1]$. Items of quality 1 essentially appeal to the entire population of consumers or users, while items of quality $\varepsilon \approx 0$ appeal to few users, those high-information readers who are interested in almost all information on a topic. More formally, each user u has a quality *threshold* $\overline{q}_u$, and is interested in exactly the set of content items $C_u = \{c : q_c \geq \overline{q}_u\}$. Thus, a high-information user has a low quality threshold, while a disengaged user (who is only interested in the most important items) has a very high threshold. A user u gets positive utility from reading an item in her content set C_u, and negative utility from reading a low-quality item not in C_u. For now, we assume that any item in C_u contributes $+g$ to the utility of u, while an item not in C_u contributes $-h$ to the utility.[3]

How do users find the content that they are interested in? We posit that *intermediaries* (or *bloggers*) play a significant role by filtering: A user interested in a subset of the items may choose to follow a blogger who only posts about the more interesting items, instead of directly reading a primary source such as a newspaper (which probably produces much more content than she would like). A blogger b performs this filtering role by

---

[1] from 5/2/2011 at 3:30am EST to 5/3/2011 at 1:30pm EST.

[2] twitter.com/twittercomms/status/65125115272249344

[3] One can consider several variants of this model: For instance, the (positive or negative) utility might depend on the quality of the item, not just on whether it was above or below the threshold $\overline{q}_u$. Such models have similar features to the one discussed here.

picking a threshold q_b; he posts all articles with quality at least q_b. Thus, the set of items posted by b is $F_b = \{c : q_c > q_b\}$. Now, if user u "follows" blogger b, the utility of $u = |F_b \cap C_u| \cdot g - |F_b \setminus C_u| \cdot h$, where the former term denotes the positive utility from those articles posted by b that u is interested in, and the latter term denotes the negative utility from those items posted by b which are below u's quality threshold. The strategic decision faced by bloggers is what threshold to pick: If users are more likely to follow bloggers that give them high utility, what posting threshold is likely to maximize the number of followers of a blog?[4]

To understand how bloggers pick a posting threshold, we study the game where each blogger b decides on his quality threshold q_b arbitrarily to maximize his audience.[5] In this game, which we refer to as the *blog-positioning* game, the utility of a blogger is given by the number of users who choose to follow him (and not other bloggers).

First, suppose that each reader picks the best blogger (the one that gives her maximum utility among all bloggers) deterministically (in the case of ties, the reader picks one of the best bloggers uniformly at random). In this setting we can prove the following result:

THEOREM 1. *If each user picks the blogger that maximizes her utility, the blog-positioning game has no pure-strategy Nash Equilibrium.*

We note that it is not really reasonable for bloggers to implement mixed strategies (on one day, posting a lot of content, and on the next day, being very selective) in order to attract readers. Thus, the lack of pure-strategy equilibrium points to a flaw in the model's depiction of reality. In the real world, it would be very expensive / time-consuming for a reader to find the absolute best blogger; further, it is also unrealistic to assume that a reader would be able to accurately estimate her quality threshold. For those reasons we study a second model where readers *satisfice*: each reader picks a desired utility level, and among all the bloggers that give her at least this utility, she picks one uniformly at random. Equivalently, one can think of the decision of reader u as follows: Given her quality threshold $\overline{q}_u$ and a desired utility level, there exists an interval (containing $\overline{q}_u$) such that if a blogger picks a posting threshold in this interval, he gives u at least the desired utility. Thus, each reader defines an interval around her threshold, and she picks a blogger uniformly at random from among those who post in this interval. In this setting we can prove that:

THEOREM 2. *If users pick randomly among blogs that give them sufficient utility, the blog-positioning game has a pure-strategy Nash Equilibrium.*

This theorem follows from the following observation: we can (roughly speaking) model blogger dynamics as a congestion game, where congestion represents competition between blogs for the same readership.

We also show that the existence of intermediaries serves a valuable social function; in their absence, readers have no choice but to follow primary sources, which may produce far more content than they are interested in. Not only do bloggers filter content, by selfishly maximizing their follower counts, they ensure that (up to a factor of two), almost as many readers as possible will find a blog that gives them sufficiently high utility.

THEOREM 3. *If users pick randomly among blogs that give them sufficient utility, the blog-positioning game has Price of Anarchy equal to 2.*

4. DISCUSSION

We believe our work motivates further investigation.
Validation: We work on strengthening our information filtering evidence, especially according to other metrics of activity and popularity.
Readership bias: Lenient and equanimous users help stabilizing the system, but how would a limited bias in blog selection affect this property?
Bundling, Dynamicity: Blogs typically bundle topics together to attract users with diverse interest. We believe our analysis of system efficiency extends to multi-topics models and to account for information ageing.

5. REFERENCES

[1] J. An, M. Cha, K. Gummadi, and J. Crowcroft. Media landscape in Twitter: A world of new conventions and political diversity. *Proceedings of ICSWM*, 2011.

[2] L. Backstrom, J. M. Kleinberg, and R. Kumar. Optimizing web traffic via the media scheduling problem. In *KDD '09: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2009.

[3] M. Cha, F. Benevenuto, H. Haddadi, and K. Gummadi. The World of Connections and Information Flow in Twitter. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 2012.

[4] M. Gupte, M. Hajiaghayi, L. Han, L. Iftode, P. Shankar, and R. Ursu. News posting by strategic users in a social network. *Proceedings of WINE*, 2009.

[5] P. R. Jordan, U. Nadav, K. Punera, A. Skrzypacz, and G. Varghese. Lattice games and the economics of aggregators. *WWW '12: Proceedings of the 21st international conference on World Wide Web*. 2012.

[6] E. Katz. The Two-Step Flow of Communication: An Up-To-Date Report on an Hypothesis. *Public Opinion Quarterly*, 1957.

[7] F. Wu and B. A. Huberman. Popularity, novelty and attention. *EC '08: Proceedings of the 9th ACM conference on Electronic commerce*, 2008.

[8] S. Wu, J. M. Hofman, W. A. Mason, and D. J. Watts. Who says what to whom on twitter. *WWW '11: Proceedings of the 20th international conference on World wide web*. 2011.

---

[4] Bloggers may wish to maximize other objectives, such as the total number of pageviews on their blog, from which they derive advertising revenue. Similar issues arise, but we do not describe them in this abstract.

[5] Note that we allow the bloggers to change their quality threshold in response to the decisions of other bloggers.

A Comparative Study of Geographic Routing in Social Network Based on Mobile Phone Data

C. Herrera,[1,*] Thomas Couronne,[2] Zbigniew Smoreda,[2] C. M. Schneider,[1] R. M. Benito,[1,†] and M. C. González[1]

[1]*Department of Civil and Environmental Engineering,*
Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America
[2]*Orange Labs, Paris, France*

Understanding the principles driving the organization of complex systems is crucial for a broad range of fields such as, information and social sciences, economics and biology. In networks representing complex systems where their nodes are defined on geographic positions, spatial constraints may have a strong effect on the connectivity patterns. Links may either be spatially embedded, such as in roads or railway lines in transportation networks, or abstract entities, such as collaboration or friendship relations in social networks. In either case, space plays a crucial role in the network connectivity and its structure. In this work we focus on the study of social networks extracted from mobile phone data. The data has been filtered in such a way that the nodes are mobile phone users and the links represent a social interaction that indicates an acquaintance or a friendship as opposed to just an occasional phone call. This is possible because our current data spans over a period of at least six months in each of the countries we study (displayed in in Fig. 1).

The idea behind this work is based on the famous small-world experiment performed by the social psychologist Stanley Milgram [1] in the 1960s. Milgram's experiment really led to two striking discoveries, of which the existence of short paths between two randomly selected people in a country, was only the first. The second was that people in society, with knowledge of only their own personal acquaintances (local information), were collectively able to forward the letter to a distant target surprisingly fast. That such a decentralized routing scheme is effective suggest short paths in the underlying social network.

Many interesting questions remain open in relation to Milgram's, experiment, such as why should a social network contain such short paths and how people are able to select among hundreds of acquaintances the correct person to form the next link in the chain. To answer these questions several works have been carried out, both empirically and mathematically. Dodds et al. [2] repeated Milgram's experiment at large scale with e-mails, providing confirmation that geographical position of the nodes has a crucial role in the possibility of the network to be searchable. On the other hand, Lieben-Nowell et al. [3] proved a simple geo-greedy algorithm is able to efficiently route a message between different cities using the social network, with data from an online blog community where users declare the city where they live.

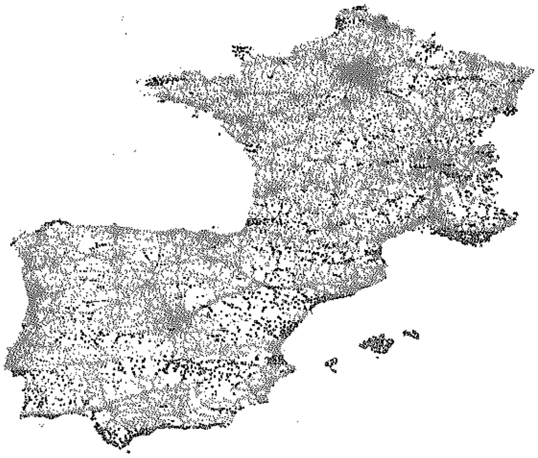


Figure 1. The spatial distribution of several thousand mobile phone towers over three countries, which are used for the Miligram experiment on top of the social network created with billions of phone calls.

Watts and Strogatz [4, 5] proposed a hierarchical network model that present the small world effect and high clustering. Kleinberg [6] has proposed several analytical treatable network models capable of finding short paths with high probability. He studied search algorithms in a graph where nodes are placed on a two-dimensional (2D) lattice with a fixed number of links whose placement is correlated with lattice distance to the other nodes. Under a specific form of the correlation, an algorithm with knowledge of the target's location can find the target in polylogarithmic time.

However, in the most general distributed search context, one may have very little information about the global structure of the network, and in most of real systems it cannot be considered as a simple 2D grid.

The goal of this research is to develop an efficient algorithm to be able to deliver a message in a small number of hops from a source user to a target user whose location

---

[*] Email: carloshy@mit.edu
Permanent address: Universidad Politécnica de Madrid, 28040 Madrid, Spain

[†] Permanent address: Grupo de Sistemas Complejos, and Departamento de Física y Mecánica, Escuela Técnica Superior de Ingenieros Agrónomos, Universidad Politécnica de Madrid, 28040 Madrid, Spain

Country	Number of Nodes	Number of Edges	$\langle k \rangle$	$\langle c \rangle$	$\langle c_r \rangle = \langle k \rangle / N$	$\langle l \rangle$	$\langle l_r \rangle = \ln(N)/\ln(\langle k \rangle)$
France	18 816 483	81 614 506	8.67	0.16	$9 \cdot 10^{-7}$	8.52	7.75
Portugal	1 218 377	4 003 452	6.57	0.26	$5 \cdot 10^{-7}$	8.35	7.44
Spain	5 929 366	16 137 920	5.44	0.21	$48 \cdot 10^{-7}$	10.36	9.20

Table I. Some properties of the social networks in the studied countries like the average degree $\langle k \rangle$, average clustering coefficient $\langle c \rangle$, and average shortest path length $\langle l \rangle$ as well as the corresponding values for random networks with the same size $\langle c_r \rangle$ and $\langle l_r \rangle$. The observed values are typical for small-world networks.

is given in terms of his/her coordinates in a city of the country under study. We want to discover the most efficient strategies for the case of limited information (without global knowledge of the social network), just using local information such as the identities and connectivity of a node's neighbors, and its neighbors' neighbors based on the structure of the empirical social networks.

We study three different countries, both at the inter-city and at the intra-city levels. We demonstrate that our search algorithm works well on real friendship networks at the inter-city level.

We use three different large data sets corresponding to six months of anonymized mobile phone calls of three European countries (France, Portugal, and Spain). The data sets include information of the most used tower coordinates (France and Portugal), or billing zip codes (Spain) for each of the users present in the call data.

From these data sets we build friendship networks based on reciprocal communications taking into account the user spatial distribution. We carry out first, an empirical study of the topological characteristics of the corresponding networks. In Table I, we present the general properties of these networks and their corresponding random networks for each country. It worth noticing that the network corresponding to France is much larger than the ones of Spain and Portugal, and it presents a higher average connectivity degree and the lowest average clustering coefficient. This behavior is a reflection of a higher market share.

In Fig. 1 we present the geographical maps of the three countries under study with the location of the corresponding communication towers. It can be noticed that although the towers are well spread over space, the distributions is not homogeneous presenting a higher concentration in the capitals and main cities.

Next, we analyze the individual's friendship interactions as a function of the geographic distance, and explore the relationship between network topology and geography and their joint importance in understanding how friendship spread through the populations of each country. To this end we present in Fig. 2, the friendship probability (measure as the number of links) as a function of the geographical distance of the towers associated to two connected nodes. It can be clearly observed that the probability of friendship decreases with the distance following almost a power law until approximately 700 km, for France and Spain, and 500 km in Portugal. After these distances it falls very fast due to approaching the

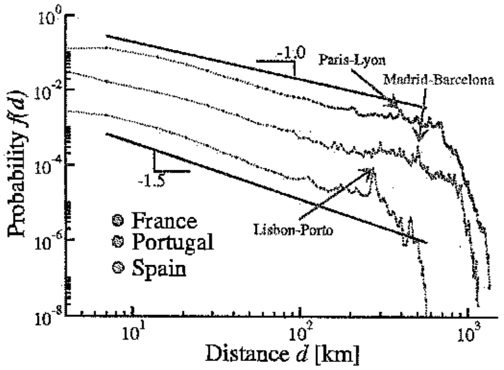


Figure 2. The distance dependent probability distribution follows a power-law until the distance reaches the country borders. The peak in the distributions related to the distance of the two largest cities is highlighted. Note that the probability distributions are vertically shifted for better visibility.

physical boundaries of the countries. It should be also remarked that there are peaks in the distributiom corresponding to the geographic distance between the main cities in each country. These peaks are usually explained with a radiation model [8]. In the context of friendships, however, this model should be corrected, since it is not symmetric. Therefore, we are working on an extension of the model for social networks with spatial structure both at the inter-city and at the intra-city level.

The length of the shortest path distributions of each network are shown in Fig.3. It can be clearly noticed that they present the small world property, being the average values of the shortest path 7.75, 7.44, and 9.2 for France, Portugal and Spain networks, respectively. These values are very close to the one expected for random networks with the same average connectivities of the empirical networks whose values are given in Table I. The fact that the network corresponding to Spain has a slightly larger value of the average length than the other two networks, is associated to the lower value in the connectivity degree, that can be explained by the smaller market coverage of the Spainish data set.

We have also developed an efficient algorithm to deliver a message in a small number of hops from a source user to a target user whose location is given in terms of his/her

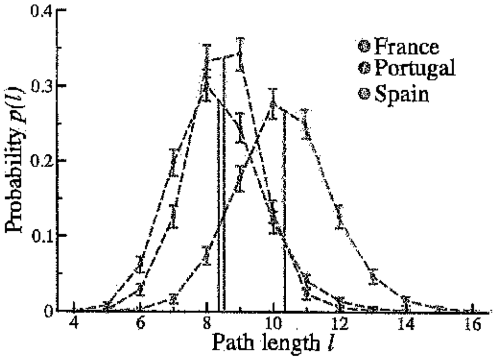


Figure 3. The shortest path length distribution between two users is centered around the mean values with small derivations. The actual average shortest path length show the characteristics of small-world networks.

coordinates in a city of the country under study using only information of the source's neighbors. This algorithm sends first the message to the neighbor whose location is geographically closer to the target. In case of more than one neighbor at the same distance, the algorithm select the one with the highest degree. We expect to demonstrate that our search algorithm works well on real friendship networks, the needed number of hops are similar to the optimal shortest path routing, and may be useful to help to reduce the network search traffic in many applications.

Due to the advantage that our data have enough resolution to face the challenge of the intra-city study, we are also studying the underlying social structure at this level, which was lacking in the previous studies [7]. We expect that the routing behavior within cities will be completely different that the one at the inter-city level. The distance alone does not have a crucial role predicting a connection in the social network. As an alternative, we will explore the community structure of the local graphs and we expect to get information of communities, related to school, work, sports, or other specific attribute of the users. This information will be critical to develop a routing algorithm at the intra-city level.

---

[1] Milgram, S., The small world problem. Psychology Today **1**, 62-67 (1967).
[2] Dodds, P.S., Muhamad, R., Watts, D.J., An Experimental Study of Search in Global Social Networks. Science **301**, 827-829 (2003).
[3] Lieben-Nowell, D., Novak, J., Kumar, R., Raghavan, P., Tomkins, A., Geographic routing in social networks. PNAS **102**, 11623 (2005).
[4] Watts, D.J., Strogatz S.H., Collective Dynamics of Small World Networks. Nature **393**, 440 (1998)
[5] Watts, D.J., Doods, P.S., Newman, M.E.J., Identity and Search in Social Networks. Science **296**, 1302-1305 (2002).
[6] Kleinberg, J., Complex networks and decentralized search algorithms. Proceedings of the International Congress of Mathematicians: Madrid, August 22-30, 2006: invited lectures, 1019-1044 (2006).
[7] Onnela J.-P., Arbesman S., González M.C., Barabási A.-L., Christakis N.A., The small world problem. Geographic Constraints on Social Network Groups. PLoS ONE **6**, e16939 (2011).
[8] Simini, F., González, M., Maritan, A., Barabási, A.-L., A universal model for mobility and migration patterns, Nature **484**, 96-100 (2012).

Homophily does not inflate contagion estimates in a social dilemma laboratory experiment

Jillian J. Jordan[a], David G. Rand[a,b], Samuel Arbesman[c,e], James H. Fowler[h,i] and Nicholas A. Christakis[c,d,f,g,1]
[a]Department of Psychology,
[b]Program for Evolutionary Dynamics,
[c]Institute for Quantitative Social Science, and
[d]Department of Sociology, Harvard University, Cambridge, MA 02138;
[e]Ewing Marion Kauffman Foundation, Kansas City, Missouri, United States of America;
[f]Department of Health Care Policy, and
[g]Department of Medicine, Harvard Medical School, Boston, MA 02115;
[h]Political Science Department, and
[i]Medical Genetics Department, University of California, San Diego, La Jolla, CA 92103;

Author email addresses:

jillianjordan@college.harvard.edu, drand@fas.harvard.edu, sarbesman@kauffman.org, james-fowler@ucsd.edu, and christakis@hcp.med.harvard.edu

Humans live in highly organized societies. Rather than interacting with each other at random, our relationships are governed by network processes. A key aspect of network interactions is social contagion: people are influenced by their social contacts, and a wide range of evidence suggests that, as a result, emotions, ideas, germs, and behaviors tend to spread across networks. Of course, social contagion is not the only factor that can lead to clustering of traits in social networks. Such clustering can also result from homophily, or the tendency of individuals to form connections with individuals that are similar to them. Thus, teasing apart the role of contagion versus homophily using statistical or experimental methods is a central issue in network science. Controlled laboratory experiments provide a powerful tool for addressing this issue. Previous work has taken advantage of laboratory experiments to demonstrate the presence of social contagion in the absence of homophily. A recent study from our group examined the spread of cooperative behavior by reanalyzing data from a classic experiment using the public goods game. By giving subjects no control over their interaction partners, the possibility of homophily was excluded. Nonetheless, subjects who were matched with more cooperative partners cooperated more themselves in subsequent rounds. This result provides clear evidence that this social behavior can be contagious, even among anonymous strangers in a laboratory setting with no homophily.

The question remains, however, whether homophily artificially inflates contagion estimates when it *is* possible. Here, we explore this issue using data from another controlled laboratory experiment examining cooperative behavior. We find that in this dataset, the possibility of homophily either has no effect on or increases clustering estimates. In contrast, however, homophily either has no effect on or decreases contagion estimates (i.e. estimates of the probability of changing state in the future based on the current behavior of one's neighbors). Thus, in the context of cooperative behavior, homophily does not appear to inflate estimates of social contagion, when approached longitudinally.

To be clear, any particular application does not prove or disprove the existence of homophily in general, nor the extent to which it might alter estimates of contagion in general. Virtually all social systems, and most social processes, will involve *both* homophily and contagion, and this has been appreciated since the 19th Century. The question is whether analytic techniques are able to adduce evidence for both processes, distinctly. The results here, and also analyses conducted by others, emphasize the fact that homophily can, depending on the circumstances, inflate, deflate, or not affect assessments of interpersonal influence. And these results emphasize the value of studying social contagion by looking longitudinally at the processes at hand, and considering changes in states.

More specifically, we examine clustering and contagion in settings where homophily does or does not occur by taking advantage of a dataset from a previously published cooperation experiment where homophily was excluded in one condition but possible in another. In this experiment, subjects recruited from Amazon Mechanical Turk engaged in a stochastically repeated social dilemma on a social network in groups averaging 20 people in size. Subjects were randomly assigned to positions on social networks, and each round, subjects chose between cooperating (paying a cost for each neighbor to receive a benefit) or defecting (paying no costs and generating no benefits). After making their decisions, subjects were informed of the decisions of each of their neighbors.

In the *fixed* condition, links were not updated; in all rounds, subjects played with the same neighbors, precluding the possibility of homophily. In the *fluid* condition, after each round, the network connections between 30% of subject pairs were randomly selected to be updated. One member of each pair was informed of the other's behavior in the preceding cooperation round and was offered the opportunity to alter their connection with that individual. If a link did not already exist between the pair, the subject could form a new connection, and if a link already existed, the subject could break it. Thus, in the fluid condition, subjects could preferentially form and maintain relationships with others that behaved similarly to them; that is, homophily was possible. These two experimental conditions thus allow us to compare estimates of clustering and contagion when homophily is and is not possible.

Before discussing the results, it is important to note that behavior in this experiment constitutes play in repeated cooperation games. Therefore, while changes in behavior over time fit the formal definition for social contagion, such shifts may also reflect conditional strategies employed in repeated games. Additionally, unlike traditional homophily, the dynamic network structure in the fluid condition creates asymmetrical assortment; cooperators preferentially assort with other cooperators, but defectors do not preferentially assort with other defectors. It is important to keep this in mind when considering our results.

First, we compare clustering of cooperation and defection in the fixed and fluid conditions. We measure clustering, or the extent to which neighbors are similar to each other, by examining the probability than an individual cooperates as a function of the percentage of her neighbors that cooperative in the current round. We use logistic regression with robust standard errors clustered on subject and experimental session, accounting for the non-independence of observations from the same subject or from subjects within the same session, and include controls for round number and the number of neighbors an individual has. We find a significant positive association between the probability of cooperating and the percentage of cooperating neighbors in the fluid condition but not in the fixed condition (although a regression including both conditions finds no significant interaction between condition and percentage of cooperative neighbors). This potential asymmetry probably arises from the fact that cooperators in the fluid condition were more likely to form and maintain connections with other cooperators, resulting in more similarity between neighbors in the fluid condition relative to the fixed condition. Thus, if considered alone, one would infer a significant clustering effect in the fluid but not fixed condition. This suggests that the possibility of homophily may inflate estimates of social contagion, if behavioral clustering is interpreted as evidence of social contagion, and highlights the often- discussed problems with this method.

Next, we examine social contagion by measuring changes in behavior across time. In these analyses, we examine the probability that an individual changed behavior between the previous round and the current round, as a function of her neighbors' behavior in the previous round. We examine transitions both from cooperation to defection, and from defection to cooperation, and again use logistic regression with robust standard errors and controls to relate these transition probabilities to the fraction of cooperating neighbors in the previous round.

For the contagion of defection, in both the fixed and fluid conditions, we observe a significant negative association between the probability of switching to defection and the percentage of cooperating neighbors. Thus, having a greater fraction of defecting neighbors makes one more likely to switch to defection oneself, both when homophily is or is not possible. In contrast, for the contagion of cooperation, the relationship between the probability of switching to cooperation and the percentage of cooperating neighbors is statistically significant in the fixed condition but not the fluid condition (although, as in the clustering analysis, there is no significant interaction between condition and percentage of defecting neighbors when considering all data together). Thus, in contrast to the clustering analysis, we find no evidence that homophily inflates estimates of the contagion of defection; if anything, homophily *deflates* contagion estimates in this experimental context: if considered alone, one would infer a significant contagion of cooperation effect in the fixed but not fluid condition.

This may be because in the fluid condition, individuals have the possibility of forming new connections with new neighbors, creating an incentive to switch to cooperation to attract new cooperative neighbors. Thus, individuals in the fluid condition may switch to cooperation even when they are paired with uncooperative neighbors, resulting in a deflated relationship between the previous cooperative behavior of one's neighbors and the probability of switching to cooperation. This suggests that the spread of selfish behavior may be a more robust property of human social networks than the spread of cooperative behavior, not because cooperative behavior itself is less stable (indeed, subjects cooperate more in fluid than fixed networks), but rather because cooperative behavior can result from a larger set of motivations.

Together, these results highlight the value of studying social contagion by exploiting a longitudinal approach that considers changes in state and also contribute to our understanding of the contagion of cooperative behavior.

Topic-Specific Communication Patterns in Email Data

Peter Krafft[1], Juston Moore[1], Hanna Wallach[1], Bruce Desmarais[2], James ben-Aaron[2]
{pkrafft, jmoore, wallach}@cs.umass.edu, desmarais@polsci.umass.edu, jbenaaro@acad.umass.edu
[1]Department of Computer Science, [2]Department of Political Science, University of Massachusetts Amherst

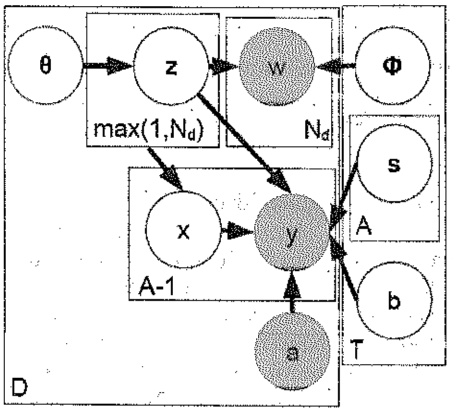
The structures of communication networks are critical to collaborative problem solving [1]. Email data sets can provide a unique view into the communication patterns of organizations and help researchers directly observe these communication networks. In this project, we are particularly interested in topic-specific communication patterns. Such information may be valuable to organizations that want to know, for example, whether there are any breaks in communication about particular topics. Organizations may wish to remedy these breaks in order to prevent lack of communication between relevant communities. We introduce a novel Bayesian admixture model for the text and network attributes of email data that can be used to jointly infer topics of communication, a partition of the email network into topic-specific subnetworks, and two-dimensional visualizations of those subnetworks. After validating our model with a link prediction task and two data fit experiments, we use a new email dataset to show how different inferred topic-specific subnetworks exhibit different behavior, and how our visualizations could be used to help assess the appropriateness of an organization's current structure.

There has been a great deal of work on probabilistic generative models for jointly modeling text and network data. Although there has been some work on modeling email data (e.g., the author--recipient topic model [2]), much of the recent work in this area has focused on networks of documents, such as web pages and the links between them [3] or academic papers and their citations [4]. In contrast to these types of data, email networks are networks of individuals; documents are associated with connections between nodes rather than with nodes themselves. Additionally, much recent work on jointly modeling text and network data has focused on link prediction (e.g., the relational topic model [3]) and does not offer the ability to extract qualitative structural information about the network and associated documents. Unlike this previous work, we focus on latent space embeddings of email networks for the purpose of producing network visualizations that are meaningful, precise, and accessible to users who wish to perform exploratory or descriptive analyses. There are a large number of existing techniques for network visualization; however, rather than viewing an email network as a single communication network, we treat an email network as a composition of multiple topic-specific networks and visualize each one individually. To date, identifying and visualizing topic-specific networks has been almost completely unexplored.

Our model builds upon three previous approaches. Our strategy is to use a joint generative model of the text and network data with a structure similar to that of Correspondence-LDA (Corr-LDA) [5]. Our model has a set of admixture components for explaining the words of the emails and another set for explaining the recipients of those emails. As in latent Dirichlet allocation (LDA) [6] and Corr-LDA we use multinomial distributions over words as the admixture components for the text data. We use latent spaces in the sense of Hoff's latent space model (LSM) [7] as the admixture components for the network data. LSM models the probability of a connection between two actors as a function of the distance between those actors in a latent (i.e., unobserved) Euclidean space. Latent spaces of this sort have not previously been used in the context of mixture or admixture models. Our model therefore provides the ability to identify topics, as in LDA, and to identify topic-specific subnetworks while embedding those networks of actors in two-dimensional Euclidean space for visualization, as in LSM.

Our model associates network edges with topics. We view an email data set as a multinetwork consisting of observations of edges between pairs of actors where each edge is an author--recipient pair within an email. According to our model, a word in an email is well-described by a particular topic if that word has high probability in the topic and the topic has high probability in that document. Thus, our model infers groups of words (i.e., topics) that commonly occur together in documents. Furthermore, a group of edges (i.e., a subnetwork) is well-described by a latent space if actors who communicate more frequently in that subnetwork are closer together in that space. Our model identifies each topic's subnetwork by finding edges that are well-described by that topic's latent space and which belong to documents containing that topic. A summary of our model in the form of its graphical model and its generative process is shown in Figure 1.

We also introduce and analyze a new email data set. This data set is part of the public record and represents a yet-untapped source of email data for academic researchers in the computer, social, and organizational sciences. The data set consists of emails between department managers in New Hanover County, North Carolina. These departments comprise the executive arm of the government at the county level. In this autonomous local government, the county manager acts as the executive and the individual departments are analogous to the individual departments and agencies in, for instance, the U.S. Federal government. In addition to offering a view into the communication patterns of the department managers of New Hanover County, our analysis serves as a case study in modeling interagency communications in government administration.



α	hyperparameter for a symmetric Dirichlet
$\theta_d \sim \mathrm{Dir}(\alpha)$	topic proportions for doc. d
β	hyperparameter for a symmetric Dirichlet
$\phi_t \sim \mathrm{Dir}(\beta)$	distribution over word types for topic t
$z_{dn} \sim \mathrm{Mult}(\theta_d)$	topic assignment of token n in doc. d
$w_{dn} \sim \mathrm{Mult}(\phi_{z_{dn}})$	type of token n in doc. d
μ, σ^2	Gaussian hyperparameters
$s_i^{(t)} \sim \mathcal{N}(0, \sigma^2)$	position of actor i in latent space t
$b_t \sim \mathcal{N}(\mu, \sigma^2)$	intercept of latent space t
a_d	author of doc. d
$x_{dj} \sim U(1, \ldots, N_d^*)$	the token assignment of edge j in doc. d
$y_{dj} \sim \mathrm{Bern}(p_{a_d j}^{(z_{d x_{dj}})})$	edge between actors a_d and j in doc. d

Figure 1: The directed graphical model for our new model and a table summarizing the corresponding probabilistic generative process, where $N_d^* = \max(1, N_d)$, and we use σ, the logistic sigmoid function, to map offset Euclidean distances to probabilities. The probability of actor i communicating with actor j in space t is defined as $p_{ij}^{(t)} = \sigma(b_t - \|s_i^{(t)} - s_j^{(t)}\|)$.

More specifically, our new data set consists of the complete email inboxes and outboxes of 30 department managers from February, 2011. There are 1,739 emails between the managers themselves, 8,097 emails authored by the managers, and 30,909 emails authored or received by the managers. In all our experiments, we use the fully observed subset of the data set---i.e., emails between the managers themselves. In addition to analyzing our new data set, we also validate our model using the 50 most active actors from the Enron email corpus [8], a widely-used, publicly-available email data set.

The primary goal of our experiments is to illustrate the utility of our model as an exploratory and descriptive tool; however, we first validate our model using a link prediction task and a topic coherence task. These experiments are designed to show that the model is not overfitting and that the topics learned during joint inference are at least as meaningful as the topics learned by LDA. Our model outperforms all our comparison methods on link prediction and achieves topic coherence comparable to LDA and our other baselines. These results suggest that our model can achieve state-of-the-art predictive performance without harming the coherence of the topics it infers, and that our modeling assumptions are reasonable.

After validating our model, we focus on the New Hanover County (NHC) data set. We first show using a simulation study that our model can represent network statistics of the NHC corpus. We then conduct an in-depth analysis to showcase the novel capability of our model to visualize topic-specific communication patterns. Examples of four latent spaces inferred by our model are shown in Figure 2. Through these plots of the topic-specific latent spaces, we are able to (1) analyze the degree to which communication patterns are consistent with the NHC organizational chart and (2) identify groups of government departments who communicate within groups on the same topics, but not across groups. Our methodology is powerful, intuitive, and applicable to any email corpus as well as to general multinetworks with text associated with their edges. Any organization that conducts a substantial proportion of its communications via email would find our model useful for summarizing the topics of its internal communications as well as the communication structure within those topics.

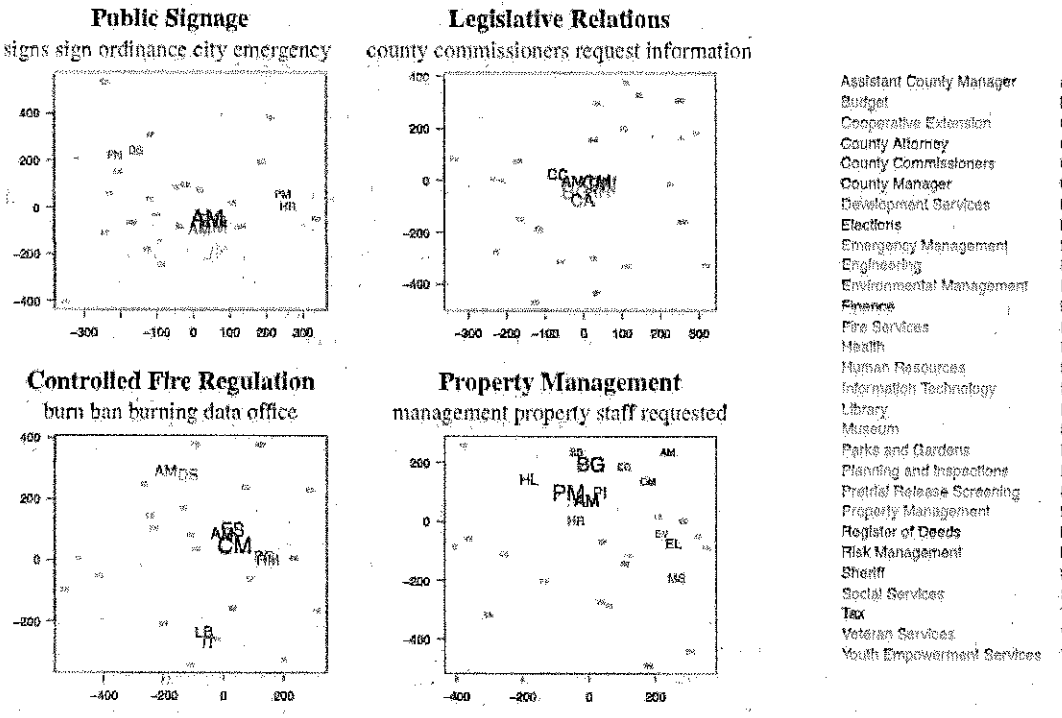


Figure 2: Four example topic-specific latent spaces. The plot for each space is titled with a human-selected label (in bold) for the corresponding topic. Below each label are four or five of the most probable words in the corresponding topic, ordered from most to least probable. The size of each department's acronym indicates how often that department communicates in that space. Acronyms are colored according to their respective 'division' in the New Hanover County government organizational chart, which can be found at http://www.nhcgov.com/Budget/Documents/FY10-11%20Adopted%20Budget.pdf. Notice that the acronym "AM" appears twice in all plots because there are two assistant county managers.

References

[1] W. Mason and D.J. Watts. Collaborative learning in networks. *Proceedings of the National Academy of Sciences*, 109(3):764–769, 2012.

[2] A. McCallum, X. Wang, and A. Corrada-Emmanuel. Topic and role discovery in social networks with experiments on Enron and academic email. *Journal of Artificial Intelligence Research*, 30(1):249–272, 2007.

[3] J. Chang and D. Blei. Relational topic models for document networks. In *Artificial Intelligence and Statistics*, 2009.

[4] E. Erosheva, S. Fienberg, and J. Lafferty. Mixed-membership models of scientific publications. *Proceedings of the National Academy of Sciences*, 101(Suppl 1):5220, 2004.

[5] D.M. Blei and M.I. Jordan. Modeling annotated data. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in Information Retrieval*, pages 127–134. ACM, 2003.

[6] D.M. Blei, A.Y. Ng, and M.I. Jordan. Latent Dirichlet allocation. *The Journal of Machine Learning Research*, 3:993–1022, 2003.

[7] P.D. Hoff, A.E. Raftery, and M.S. Handcock. Latent space approaches to social network analysis. *Journal of the American Statistical Association*, 97(460):1090–1098, 2002.

[8] B. Klimt and Y. Yang. Introducing the Enron corpus. In *First Conference on Email and Anti-spam* (CEAS), 2004.

Directed Acyclic Motifs for Conversation Analytics

Lav R. Varshney
IBM Thomas J. Watson Research Center
1101 Kitchawan Road, Yorktown Heights, New York, USA
lrvarshn@us.ibm.com

Conversations bind people together in social life and allow groups to productively work together to solve problems. Emerging social media platforms such as blogs with commenting, web forums, and discussion dashboards purpose-built to support idea generation and problem solving [1, 2] allow people from around the world to come together and be collectively intelligent. These platforms also capture the detailed structure of technology-mediated conversations [3], which can then be analyzed using techniques from network theory to gain insight into the nature of conversations. Insights into how people interact can lead to better structuring of platforms and improve policies on seeding/moderating productive conversations.

Previous analyses of web conversations have projected the structure onto a representation where nodes are people and directed edges occur when one person directly responds to another [4]. Here we study the original structure where comments are the nodes and directed edges occur when one comment directly references another. Since conversations unfold sequentially in time and therefore lack loops, their structures are directed acyclic graphs (DAGs). Although many networks that capture the structure of human interaction, such as citation networks [5], information spread networks [6], and hierarchical exchange networks [7] are DAGs, there is a paucity of network-theoretic analysis techniques specialized to this class of networks, cf. [8].

The powerful method of network motifs has been used to discover overrepresented subgraphs that form the basic building blocks of many networks [9] and to define superfamilies of networks [10], but has not been used with DAGs. Here, we extend the method of motifs to settings where networks are DAGs and apply it to understand the basic building blocks of two web conversations.

To identify network motifs with statistical significance, the basic approach is to compare the network under study to an ensemble of random networks and find subgraphs that occur significantly more often. We use the ensemble of random DAGs that preserve in-degree and out-degree distributions, generated using a rewiring procedure [8, Sec. III.C]. One could also have considered other random DAG ensembles [11]. To test statistical significance, we use the step-down min-P-based algorithm for multiple-hypothesis correction [12].

Since we are primarily concerned with conversations, we assign names inspired by logic to the several subgraphs that may occur in DAGs on three nodes, Table 1. Note the ordering constraint in DAGs reduces the number of possible subgraphs from sixteen to six. The table also gives closed-form expressions used to perform the triad census. These formulae are in terms of the adjacency matrix A, the adjacency matrix E of a graph formed by converting any directed edge into an undirected edge, the adjacency matrix M of a graph formed by only considering bidirectional links, and the matrix $C = A - M$. The complement of E, $\bar{E}$, is created when all zeros (except on the diagonal) are changed to ones and vice versa. Pointwise matrix multiplication is denoted by $\times$.

To demonstrate the method of motifs for DAGs, we consider the conversation structure in the Polymath Project, a collaboration that proved a new result in combinatorics [13, 14] and a weblog conversation about blogs driving action [15, Fig. 2]. These DAGs are shown in Figure 1. The results of our motif analyses are shown in Figure 2. Statistical analysis shows that the implication ($p = 0.005$) and the chained syllogism ($p = 0.003$) are network motifs for the Polymath Project graph. On the other hand, the weblog conversation graph has the empty triad as a motif ($p = 0.01$). That the motif profiles of the two examples are completely different implies they are built from different building blocks. Going forward, it is of interest to survey a large number of conversations and see whether superfamilies of conversations arise and if so, what functional explanations emerge.

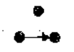
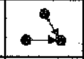
1		empty	$\frac{1}{6}tr(\bar{E}^3)$
2		implication	$\frac{1}{2}\sum \bar{E}^2\times(C+C^T)$
3		fan	$\frac{1}{2}\sum \bar{E}\times(C^TC)$
4		syllogism	$\frac{1}{2}\sum \bar{E}\times(CC^T)$
5		chain	$\sum \bar{E}\times C^2$
6		chained syllogism	$\sum C\times C^2$

Table 1. Subgraphs in Directed Acyclic Graphs.

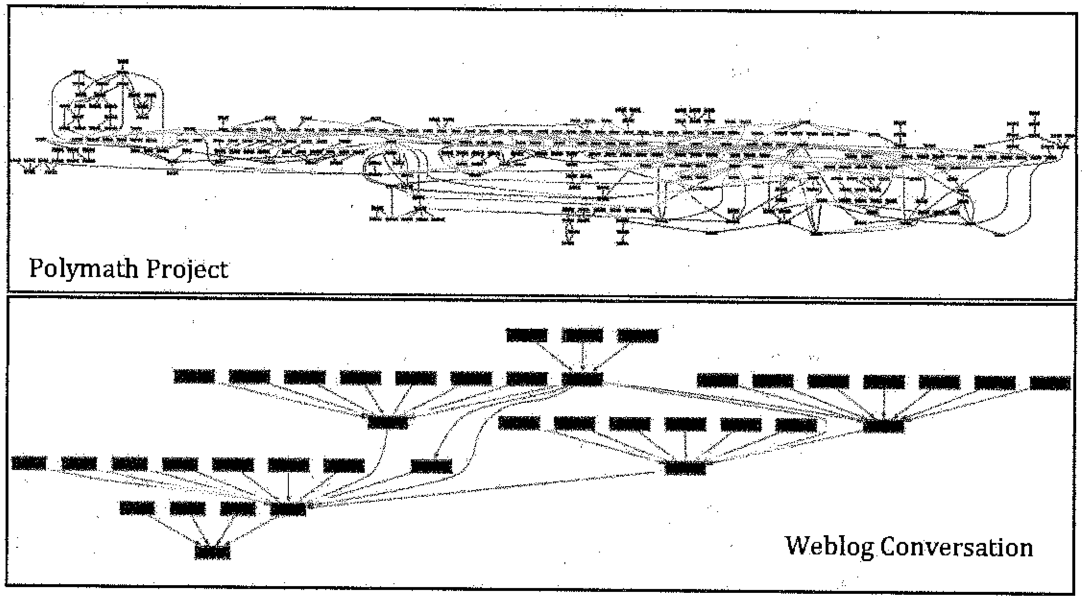


Figure 1. Network Structure of the Polymath Project and a Weblog Conversation

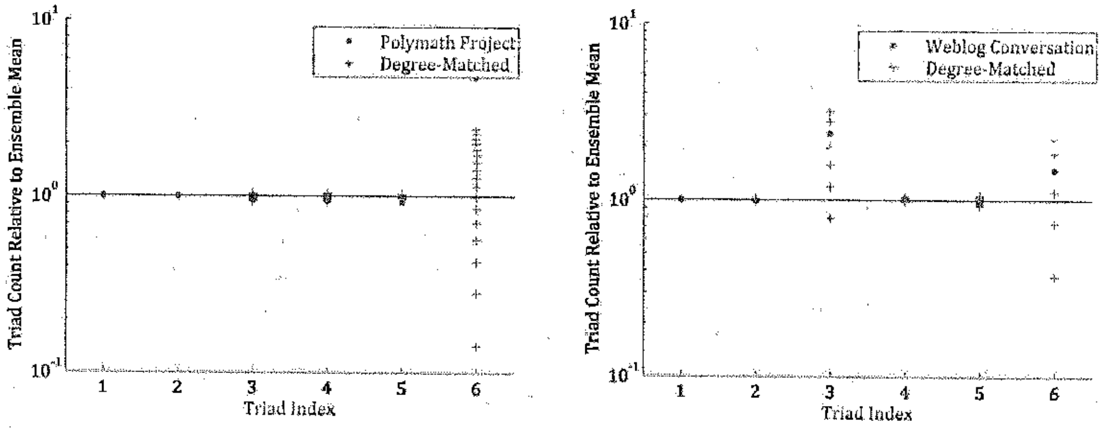


Figure 2. Network Motif Analysis of the Polymath Project and a Weblog Conversation

Acknowledgment
Structural data on the Polymath Project was graciously provided by J. Cranshaw and A. Kittur.

References

[1] K. Ehrlich, S. Rohall, S. Ross, D. Gruen, T. Ratchford, J. Patterson, H. Takagi, T. Ishihara, and A. Kosugi, "Beyond Discussions: Designing for sociability and structure," in *Proceedings of the Collective Intelligence as Community Discourse and Action Workshop*, pp. 34–38, Feb. 2012.

[2] J. Schneider and A. Passant, "Envisioning a discussion dashboard for collective intelligence of web conversations," in *Proceedings of the Collective Intelligence as Community Discourse and Action Workshop*, pp. 60–64, Feb. 2012.

[3] V. Pallotta and R. Delmonte, "Automatic argumentative analysis for interaction mining," *Argument & Computation*, vol. 2, no. 2–3, pp. 77–106, 2011.

[4] M. Helander, R. Lawrence, Y. Liu, C. Perlich, C. Reddy, and S. Rosset, "Looking for great ideas: Analyzing the innovation jam," in *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*, pp. 66–73, Aug. 2007.

[5] E. Garfield, *Citation Indexing—Its Theory and Application in Science, Technology, and Humanities*, John Wiley & Sons, Inc., 1979.

[6] D. Liben-Nowell and J. Kleinberg, "Tracing information flow on a global scale using Internet chain-letter data," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 12, pp. 4633–4638, Mar. 2008.

[7] M. Marriott, "Caste ranking and food transactions: A matrix analysis," in *Structure and Change in Indian Society*, M. Singer and B. S. Cohn, Eds. Wenner-Fren Foundation, 1968, pp. 133–171.

[8] B. Karrer and M. E. J. Newman, "Random graph models for directed acyclic networks," *Physical Review E*, vol. 80, no. 4, 046110, Oct. 2009.

[9] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, "Network motifs: Simple building blocks of complex networks," *Science*, vol. 298, no. 5594, pp. 824–827, Oct. 2002.

[10] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, and U. Alon, "Superfamilies of evolved and designed networks," *Science*, vol. 303, no. 5663, pp. 1538–1542, Mar. 2004.

[11] J. Goñi, B. Corominas-Murtra, R. V. Solé, and C. Rodríguez-Caso, "Exploring the randomness of directed acyclic networks," *Physical Review E*, vol. 82, no. 6, 066115, Dec. 2010.

[12] M. Reigl, U. Alon, and D. B. Chklovskii, "Search for computational modules in the *C. elegans* brain," *BMC Biology*, vol. 2, 25, Dec. 2004.

[13] J. Cranshaw and A. Kittur, "The Polymath Project: Lessons from a successful online collaboration in mathematics," in *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems (CHI 2011)*, pp. 1865–1874, May 2011.

[14] L. R. Varshney, "Toward a comparative cognitive history: Archimedes and D. H. J. Polymath," in *Proceedings of the Collective Intelligence Conference 2012*, Apr. 2012.

[15] A. de Moor and L. Efimova, "An argumentation analysis of weblog conversations," in *Proceedings of the 9th International Working Conference on the Language-Action Perspective on Communication Modelling*, Jun. 2004.

Artificial Social Scientist: an Application to Marriage Networks

Telmo Menezes*[1] and Camille Roth†[1,2]

[1]CAMS, CNRS/EHESS, 190 avenue de France, Paris, F-75013, France
[2]CMB, CNRS/HU/MAEE, Friedrichstrasse 191, Berlin, D-10117, Germany

1 Introduction

The current availability of both computing power and diverse corpora of large social network datasets creates an unprecedented opportunity for exploration/discovery of human strategies. However, creating plausible models then explain the emergence of such networks is not trivial and requires *a priori* intuitions on the key drivers of network formation. In this ongoing work we employ a machine learning approach to automatically discover generative models for complex social networks. This work is aligned with the idea of creating *artificial scientists*. For example, in a work with some parallels to the ideas presented in this paper, scientific laws are extracted from experimental data using genetic programming (Schmidt & Lipson, 2009).

We apply our technique to the study of matrimonial networks. More specifically, alliance networks(Roth *et al.*, n.d.), which are a simple representation for the structure of marriage relationships in a population. In an alliance network, a node represents a group of individuals and an edge represents a marriage, where the origin is the wife-giving group and the target is the wife-taking group. This representation has the advantage of removing some of the complexities present in more traditional kinship networks Hamberger *et al.* (2011), namely the existence of generations, labeled nodes (for different genders) and labels edges (for different types of relationships).

We present one of our experimental results, where we generated two theories that explain the behaviours underlying a real corpus obtained by field researchers. One of the theories was informally validated by an anthropologist who studied the population *in loco*.

2 Method

2.1 Linking behaviours as computer programs

Our network generators are simple computer programs represented as a tree structure (Koza, 1992; Poli *et al.*, 2008). Tree nodes represent functions that take the value of their child nodes as parameters. Tree leafs are either variables of constant values. A tree can be recursively evaluated all the way to the top, eventually returning a single value. The role of a program in our method we propose is to quantify the plausibility of two nodes establishing a connection at a given moment.

The input variables available to the program are: a numerical node identifier; in and out degrees; in and out strength; link strength; directed, undirected and reverse distance. The function set consists of simple arithmetic operators, general-purpose functions and comparators.

A simulation run is performed to generate the synthetic network that will be used for comparison against the real network. The number of nodes in the real network (m) and the number of edges (n) are taken as reference for the simulation. The simulation is initialised with m disconnected nodes and runs for n cycles, with a new edge being generated at each cycle.

During each simulation cycle, a weight is assigned to every possible directed edge. This weight is the result of the evaluation of the generator program. The edge to create is then selected by a stochastic process that assigns to each possible edge a probability of selection proportional to it's weight. If all probabilities are 0, one edge is selected amongst all possible pairs according to an uniform distribution.

---

* telmo@telmomenezes.com
†roth@ehess.fr

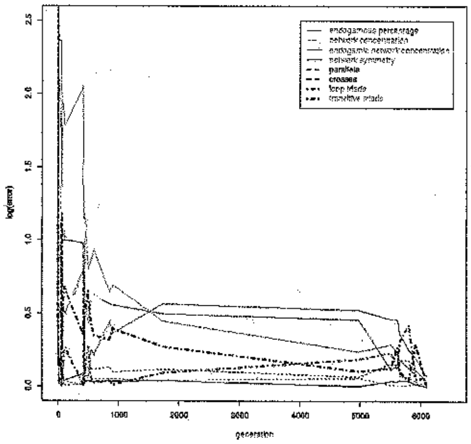


Figure 1: Convergence of the several network metrics during an evolutionary run.

	ϵ_0	$c_{\mathbf{x}}$	$c_{\mathbf{x}_0}$	$s_{\mathbf{x}}$	Π	χ	Δ	Δ'
real	0.0	0.0017	0.0	0.4673	335	324	2200	2270
synthetic A	0.0	0.0017	0.0	0.4658	328	332	2194	2289
synthetic B	0.0	0.0018	0.0	0.4913	335	300	2028	2438

Table 1: Metrics for real and synthetic networks.

2.2 Evolutionary search for behaviours

The search algorithm is initialised with a population of randomly generated programs. The algorithm runs for a number of generations, where successive populations are generated from the previous ones, by mutating and recombining programs from the previous generation. Individuals are stochastically selected to generate offspring according to a measure of their quality, usually called the *fitness function*. Evolutionary algorithms are not guaranteed to converge on a solution, but will tend to approximate an optimum. The specific type of evolutionary algorithm we use in the work is called *genetic programming*, because a solution is represented as a computer program.

The fitness function we employ is the length of the vector of the ratio between eight metrics for the real and simulated network. These metrics were selected for being key in describing the most fundamental aspects of an alliance network, and because of their straightforward interpretation in anthropological terms:

- ϵ_0 denotes the percentage of endogamous marriages;
- $c_{\mathbf{x}}$ is an index of **network concentration** which indicates the degree to which marriages are concentrated among a few pairs of marriage partners;
- $c_{\mathbf{x}_0}$ is an index of **endogamic network concentration**, and expresses how much endogamous marriages are concentrated over a few groups;
- $s_{\mathbf{x}}$ is an index of **network symmetry**, which indicates the degree to which marriages tend to be reciprocated and groups are likely to be the wife-givers of their own wife-takers;
- Π is the number of **parallel circuits** of length 2;
- χ is the number of **crossed circuits** of length 2;
- Δ is the number of **transitive triads**;
- Δ' is the number of **non-transitive triads**.

3 Results: Samo Corpus

From a set of five experiments, two resulted in high quality approximations as well as relatively simple programs. Program A approximates the real network closer than B, but is also more complex. Program B does not use specific node identifiers (e.g: OrigId), depending solely on endogenous network phenomena.

3.1 Program A

Listing 1: Program A

```
if max(UndirectedDistance, TargetOutputDegree) < (OrigId * ReverseDistance)
    return 0.07
else
    return DirectedDistance
```

This program configures two modes of operation, the first is a constant (low) value while the second makes the probability of connection proportional to the directed distance between two nodes. The first mode can be qualitatively defined as a low probability mode, given that DirectedDistance is at least 1 (except for endogamous connections). Roughly, the probability of selecting the second (stronger) mode increases with the reverse distance between the nodes and decreases with the target output degree. The origin node id is also considered, introducing an *a priori* preference for certain groups over others. Some nodes are intrinsically more desirable as wife givers. Overall, groups tend to form connections with groups that are currently distant from them.

3.2 Program B

Listing 2: Program B

```
if (ReverseDistance > 0)
    return (2.2 / log(TargetInputStrength)) + ReverseStrength
else
    return 0
```

This program appears to also contains two modes of operation, but that's not the case. In fact, the condition just prevents the creation of endogamous links, given that a distance of 0 is only possible from a group to itself. Connections only take place between two distinct groups, with the probability depending on the sum of the current reverse strength (a reciprocity behaviour) with a constant divided by the natural logarithm of the current target input strength. This program explicitly forbids endogamous links while promoting reciprocity and preferring the target groups that have received less wives.

4 Discussion

We found two models that are quite different, but both able to approximate the real data very closely. We do not find this surprising. Even human scientists are capable of generating a multitude of theories that predict equally well different phenomena.

The two models have one thing in common: both promote balance in terms of preferring creation of links with currently less popular targets. In the case of program A this is implicit, given that the less connected a node is the more likely it is to be distant to any other given node. In fact we suggest that both models propose a dynamic of *anti-preferential attachment*.

Both the quality of the approximations and the legibility of the generated models suggest that this is a promising technique to model complex social networks in general.

References

K. Hamberger, M. Houseman, and D. White (2011). Kinship network analysis. *Pages 533–549 of*: P. Carrington and J. Scott (eds), *The sage handbook of social network analysis*. London: Sage.

J. R. Koza (1992). *Genetic programming: On the programming of computers by means of natural selection*. Cambridge, MA, USA: MIT Press.

T. Menezes (2012). *Synthetic*. `http://github.com/telmomenezes/synthetic`.

R. Poli, W. B. Langdon, and N. F. McPhee (2008). *A field guide to genetic programming*. Published via `http://lulu.com` and freely available at `http://www.gp-field-guide.org.uk`. (With contributions by J. R. Koza).

C. Roth, F. Gargiulo, A. Bringé, and K. Hamberger. *Random alliance networks*. In preparation.

M. Schmidt and H. Lipson (2009). Distilling free-form natural laws from experimental data. *Science*, **324**(5923), 81–85.

Urban characteristics attributable to density-driven tie formation

Wei Pan,[1] Gourab Ghoshal,[2] Coco Krumme,[1] Alex Pentland,[1] Manuel Cebrian[3]

[1]MIT Media Lab, Cambridge, MA 02139
[2]Rowland Lab, Harvard University, Cambridge, MA 02115
[3]Department of Computer Science and Engineering, University of California, San Diego, CA 92093
E-mail:{panwei,coco,pentland, cebrian}@media.mit.edu, gghoshal@fas.harvard.edu

Empirical evidence suggests that urban features ranging from individual-level dyadic interactions (number of acquaintances, communication time) to population-level variables (contagious disease rates, patenting activity, productivity and crime) grow predictably with respect to city population, and that they do it so with surprisingly similar scaling exponents, which implies a universal scaling mechanism behind them. However there is no widely accepted generative theory that can explain these observations at both scales. We believe that the many scaling components can be explained as a result of social interactions and information flow, and we propose a simple model to explain the empirically measured scaling exponent as a function of *social tie density*. This yields a generative model for city growth, linking the effects of geography, population density and social interaction. Our results suggest that scaling behavior of city characteristics can be explained as a consequence of geographical constraints on social structure, without the need to appeal to modularity, specialization, or hierarchy.

Researchers [1, 2] show from both online and phone call datasets the probability that a link will form between two individuals at distance r is in effect a gravity model. Based on these principles on the role of density in driving social interactions, our work shows in closed form that the relationship between the number of social ties in a city and the density of the city is governed by:$T(\rho) = \rho \ln \rho + C\rho$, where C is a constant independent of ρ.

Our model prediction matches well with the number of phone call ties in US counties of different size as illustrated in Fig. 1(a). In addition, the number of ties in our model from both our closed form and simulations shown in Fig. 1(b) are reasonably close to a power-law distribution with $\beta \approx 1.2$, revealing the mechanism behind the mystery of the narrow power-law band in city growth.

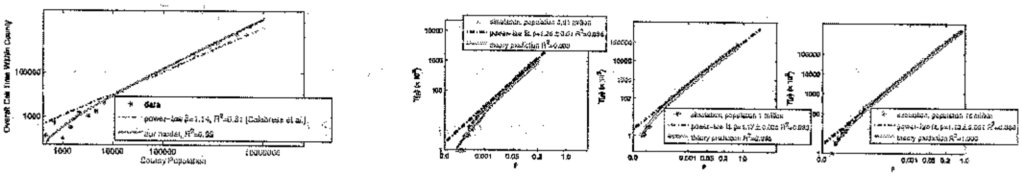


Figure 1: **Left(a):** This figure illustrates real aggregated call tine data on number of friendship ties in counties of different densities from Calabrese *et al.* [3] together with the theoretical prediction from our model which connects density ρ and number of social ties $T(\rho)$ via the $\rho \log \rho$ function. Our model captures both the power-law growth pattern and tilts on both end of the growth curve. **Right(b):**The number of ties $T(\rho)$ plotted as a function of ρ for various simulation grid sizes N and theory prediction.The best power-law fit to the scaling exponent yields a value of $\beta \approx 1.15$ independent of N similar to exponents in empirical data ($1.1 \leq \beta \leq 1.3$).

To understand how the growth processes of cities work to create observed super-linear scaling, it is insufficient to state the expected level of link formation. After all the pattern by which link spread information is at the heart of value-creation and productivity. Because it is well known that the structure of the network has a dramatic effect on the access of information and ideas [4, 5] as well as on epidemic spreading [6, 7], we believe that higher social tie density would generate a more tightly integrated society and thus engender greater levels of information flow and interaction, which lead to the productivity and innovation related as well as disease related scaling urban features.

To test the hypothesis that a city's productivity is related to how fast information travels and how fast its citizens gain access to innovations or information, it is natural to examine how this speed scales with population density under our model. The same analogy motivates the investigation into disease spreading: with greater connectivity, pathogens spread more quickly, and thus it is of interest to quantify the functional relationship between link topology and spread. We run both the SI model and the complex contagion model of diffusion [8, 9, 10] on networks generated by our model, and we notice in Fig. 2 that both diffusion models generate the same superlinear scaling spreading rate. As a consequence we propose that an explanation for both the super-linear scaling in productivity and disease is the super-linear speed at which both information flow and pathogens traverse our social network.

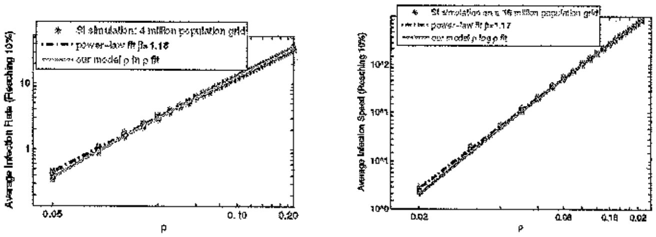


Figure 2: **a)** The mean spreading rate as a function of density ρ . The points correspond to a average over 30 realizations of simulations of the SI model on a 200 × 200 grid. The dashed line corresponds to a fit of the form $R(\rho) \sim \rho^{1+\alpha}$ with $\alpha = 0.18$. The solid line is a fit to our social-tie density model. **b)** The mean spreading rate as a function of ρ under the complex contagion diffusion model. The dashed line corresponds to the power-law fit of the form $R(\rho) \sim \rho^{1+\alpha}$ with $\alpha = 0.17$. Solid line is our model fit. In both diffusion models, our social-tie density model fits better with mean square error 29% and 41% less than the power-law fit respectively.

References and Notes

[1] D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, and A. Tomkins. Geographic routing in social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102(33):11623, 2005.

[2] G. Krings, F. Calabrese, C. Ratti, and VD Blondel. A gravity model for inter-city telephone communication networks. *Journal of Statistical Mechanics: Theory and Experiment*, L07003, 2009.

[3] F. Calabrese, D. Dahlem, A. Gerber, D.D. Paul, X. Chen, J. Rowland, C. Rath, and C. Ratti. The connected states of america: Quantifying social radii of influence. in Proc. of IEEE International Conference on Social Computing (SocialCom), 2011.

[4] M. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380, 1973.

[5] N. Eagle, M. Macy, and R. Claxton. Network diversity and economic development. *Science*, 328(5981):1029, 2010.

[6] V. Colizza, A. Barrat, M. Barthélemy, and A. Vespignani. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7):2015, 2006.

[7] P. Wang, M.C. González, C.A. Hidalgo, and A.L. Barabási. Understanding the spreading patterns of mobile phone viruses. *Science*, 324(5930):1071, 2009.

[8] W.O. Kermack and A.G. McKendrick. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A*, 115:700–721, 1927.

[9] R.M. Anderson and R.M. May. *Infectious diseases of humans: dynamics and control.* Wiley Online Library, New York, USA, 1991.

[10] D. Centola and M. Macy. Complex contagions and the weakness of long ties1. *American Journal of Sociology*, 113(3):702–734, 2007.

Structural Analysis of Viral Spreading Processes in Social and Communication Networks Using Egonets

Victor M. Preciado
Department of Electrical and Systems Engineering
University of Pennsylvania, Philadelphia PA 19104
E-mail: preciado@seas.upenn.edu

Understanding the behavior of viral spreading processes taking place in large complex networks is of critical interest in mathematical epidemiology. Spreading processes are relevant in many real scenarios, such as disease spreading in human populations, malware propagation in computer networks, or information dissemination in online social networks. To study viral spreading processes, a variety of stochastic dynamical models has been proposed in the literature. In these models, the steady-state infection of the network presents two different regimes depending on the virulence of the infection and the structure of the network of contacts. In one of the regimes, an initial infection dies out at a fast (usually exponential) rate. In the other regime, an initial infection becomes an epidemic. Both numerical and analytical results show that these two regimes are separated by a phase transition at an epidemic threshold determined by both the virulence of the infection and the topology of the network. One of the most fundamental question in mathematical epidemiology is to find the value of the epidemic threshold in terms of the virus model and the contact network .

In many cases of practical interest it is unfeasible to exactly retrieve the complete structure of a network of contacts. In these cases, it is impossible to exactly compute the epidemic threshold. On the other hand, in most cases one can easily retrieve the structure of egocentric views of the network, also called *egonets*. To estimate the value of the epidemic threshold, researchers have proposed a variety of random network models in which they can prescribe structural properties that can be retrieved from these egonets, such as the degree distribution, local correlations, or clustering.

Although random networks are the primary tool to study the impact of local structural features on the epidemic threshold, this approach presents a major flaw: Random network models implicitly induce many structural properties that are not directly controlled but can have a strong influence on the value of the epidemic threshold. For example, it is possible to find two networks having the same degree distribution, but with opposite dynamical behavior. Therefore, it is difficult (if not impossible) to isolate the role of a particular structural property in the network performance using random network models.

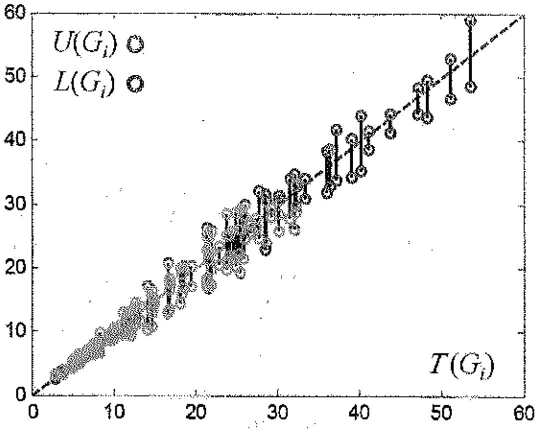


Fig. 1. Scatter plot of the spectral radius, $\lambda_1(G_i)$, versus the lower bound $\beta_2(G_i)$ (blue circles) and the upper bound $\delta_2(G_i)$ (red circles), where each point is associated with one of the 100 social subgraphs considered in our experiments.

In this paper, we develop a mathematical framework, based on algebraic graph theory and convex optimization, to study how local structural properties of the network constrain the interval of possible values in which the epidemic threshold must lie. As a result of our analysis, we present a computationally efficient approach to find this interval from a collection of egonets. To validate our approach, we analyze real data from a regional network of Facebook. In particular, we consider a set of 100 different social subgraphs $\mathbf{G} = \{G_i\}_{i \leq 100}$. For each social subgraph $G_i \in \mathbf{G}$, we compute lower and upper bounds on the epidemic threshold, denoted by $L(G_i)$ and $U(G_i)$, respectively. Since we have access to the complete network topology, we can also numerically compute the exact value of the epidemic threshold, which we denote by $T(G_i)$. Fig. 1 represents a scatter plot where each red circle has coordinates $(T(G_i), U(G_i))$, and each blue circle has coordinates $(T(G_i), L(G_i))$, for all $G_i \in \mathbf{G}$. For all the social subnetworks in $\mathbf{G}$, the epidemic thresholds are remarkably close to the theoretical bounds $L(G_i)$ and $U(G_i)$. In other words, in our collection of social subgraphs, *local structural properties of the network strongly constrain the location of the epidemic threshold, and consequently the ability of a social network to disseminate information virally.*

From a practical point of view, our approach can be used to design immunization strategies to tame the spreading of harmul viruses, or to design network topologies in order to facilitate the dissemination of information.

On the structure of communities in networks

Bruno Abrahao Sucheta Soundarajan John Hopcroft Robert Kleinberg

Department of Computer Science
Cornell University
{abrahao, sucheta, jeh, rdk}@cs.cornell.edu

1 Introduction

Community structure captures the tendency of entities in a network to group together in meaningful subsets whose members have a distinctive relationship to one another. Despite playing a fundamental role in the structure and function of networks, community structure has proved to be frustratingly difficult to define, quantify, and extract. In addition to challenges related to computational tractability, several major factors account for the intricacies of community extraction.

First, the application domain includes a wide variety of networks of fundamentally different natures. Second, the literature offers a multitude of disparate community detection algorithms. Due to differences in concept and design, the output of these procedures exhibits high structural variability across the collection. Next, there is no established consensus on the question of what properties distinguish subgraphs that are communities from those that are not communities. Additionally, it is difficult to obtain negative examples of communities; in theory, we can obtain examples of community structure, e.g., by asking experts to identify communities in a given domain, and then declare that every other subset of nodes in the network is a negative example. However, enumerating all forms of negative examples is obviously intractable, and even if we could enumerate all possible negative examples, we are still faced with the problem that these sets might also be examples of valid communities that the expert failed to identify.

In this work, we address these issues by presenting a framework that enables researchers to assess the structural dissimilarity among the output of multiple community detection algorithms and between the output of algorithms and communities that arise in practice. Our approach analyzes communities by taking account of a broad spectrum of structural properties, and reveals nuances of the structure of communities.

2 Overview

We frame this as a class separability problem, which simultaneously handles many classes of communities and types of structural properties. To this end, we specify a learning problem in which we map the distinct communities into a feature space, where the dimensions represent measures that characterize a community's link structure. The separability of classes provides information on the extent to which different communities come from the same (or different) distributions of feature values. We extract different classes of communities that can be labeled as either intrinsically-defined or extrinsically-defined communities.

We define the first set of communities by properties intrinsic to their link structure. For our purposes, these are the sets that community detection algorithms output. Each class of intrinsically defined communities comprises a set of examples that a specific algorithm extracts. We also define communities by meaningful annotations provided with the datasets, such as explicit declaration of community membership, product categories, grouping by protein function, and so on. In this fashion, for each network, we form a class of extrinsically-defined communities.

To demonstrate our approach, we furnish our framework with a large set of structural properties and consider ten different community detection procedures, representative of various categories of popular algorithms, to produce examples of different structural classes. We consider a diverse collection of large scale real networks whose domains span biology, on-line shopping, and social systems. We then assess separability using supervised classifiers both parametric, namely Support Vector Machines [10], and nonparametric, namely k-Nearest Neighbors [1], together with a feature selection analysis using correlation-based methods [5].

3 Methodology

We analyze nine large datasets, namely DBLP, two portions of the LiveJournal social network (denoted as LJ1 and LJ2), two portions of the Facebook network (denoted as Grad and Ugrad), Amazon, and three genetic networks denoted by HS, SC, and Fly. These datasets range in size from 503 to 500,000 nodes.

The networks we analyze contain annotations, which we use to identify extrinsically defined (or annotated) communities. Some of these sets are user-defined, i.e., users explicitly declare their participation in the community, while others reflect contextual information of the underlying process or organization, e.g., university department, protein function, product category, etc.

To study classes of intrinsically defined communities, we selected a collection of 10 community detection procedures, which are representative of strategies employed by a broad range of algorithms in the literature. We applied these procedures to each of the nine network datasets, and labeled the resulting sets with the identity of the community detection procedure that produced them. In total, for each network, we created 11 structural classes of communities: one class of extrinsically-defined communities, which comprises examples of annotated communities, and each of the other 10 classes corresponding to intrinsically-defined communities, which comprise examples extracted by each of the 10 community detection algorithms respectively.

The algorithms that we consider are breadth-first search (BFS), random walk with and without restart (RW15 and RW0), an algorithm to identify $\alpha-\beta$ communities [7] (AB), InfoMap [9], Link Communities (LC) [2], Louvain Modularity [3], Newman-Clauset-Moore Modularity [8], Markov Clustering Algorithm (MCL) [4], and Metis [6].

In the next phase, we measure the subgraphs induced by the communities produced in the previous step and those induced by annotated communities. We use a large spectrum of measurements that cover many properties of both the internal link structure and the external interaction of the community with the rest of the network. These measurements include features such as size, conductance, edge density, and distributions of various centrality measures (such as node betweenness, edge betweenness and information centrality).

By measuring these structural properties for each example of a community derived in the previous phase, we obtain 11 classes of labeled examples in feature space, which constitute the input in our framework.

In this work, we treat the research question of discriminating the structure of different communities as a class separability problem. This analysis is informative of the extent to which different algorithms produce structural differences and the extent to which community detection algorithms succeed in producing sets that resemble annotated communities. More specifically, we employ the Support Vector Machine (SVM) and k-Nearest Neighbor methods to confirm each other's outcomes while ruling out variability due to the specifics of each algorithm.

We perform two experiments. In the first, we are interested in analyzing structural consistency within the 11 classes of communities. For example, do the communities generated by BFS tend to resemble one another, or is there confusion between different classes? For each network, we use cross-validation to train a multi-class classifier on elements from each of the 11 classes of communities from that network. We then evaluate the classifier on the remaining elements from each of the 11 classes. Table 1 contains the results of this experiment. We see that for most networks and classes, a plurality of the probability mass from that class was properly classified. This experiment shows that the classes of communities tend to be internally consistent, demonstrating that different community detection methods produce results that are fundamentally different from each other, and suggesting that for a practitioner who wishes to find a specific type of community, the choice of algorithm is crucial.

While the first experiment showed that the various classes, including the class of annotated communities, tend to be internally consistent, we are also interested in identifying which of the 10 intrinsically defined classes the set of annotated communities most resembles. That is, although no algorithm fully captures the structure of annotated communities, which comes the closest? We train a classifier on elements from the 10 classes of communities identified through algorithms, and then apply this classifier to the set of annotated communities. We see that for nearly every network, the annotated communities are most similar to the two classes of random walk communities.

Finally, we apply the Correlation-based Feature Selection method [5] to each of the 11 classes of communities in order to determine which features are most valuable in discriminating between classes. We see that for most

	Grad	**Ugrad**	**HS**	**SC**	**Fly**	**DBLP**	**Amaz**	**LJ1**	**LJ2**
BFS	60%	88%	73%	70%	(40%)	63%	55%	86%	81%
RW0	44%	55%	43%	(39%)	(27%)	52%	43%	61%	63%
RW15	40%	(29%)	44%	42%	34%	46%	39%	57%	57%
AB	83%	91%	90%	71%	60%	70%	74%	90%	89%
IM	27%	(23%)	72%	73%	(2%)	62%	51%	82%	66%
LC	68%	96%	83%	85%	83%	67%	56%	90%	89%
Louv.	24%	(3%)	49%	(1%)	(0%)	45%	58%	38%	49%
Newm.	(14%)	(25%)	(15%)	(0%)	90%	26%	39%	45%	56%
MCL	19%	(22%)	57%	28%	(34%)	59%	46%	80%	74%
Metis	61%	73%	81%	90%	(42%)	88%	66%	92%	86%
Annot.	37%	33%	50%	46%	(8%)	47%	40%	72%	71%

Table 1: Percentage of the probability mass of classification of elements in the test set into the correct class, using SVM, for all networks. Values in parentheses indicate that a plurality of the probability mass from that was classified as some other class.

networks, conductance and diameter are especially valuable.

As illustrated by our experiments, by producing artificial or real examples of communities that possess the structure we wish to find, we can use our framework to enable an informed choice of the most suitable community detection method for a given network. In addition, it allows for a comparison of existing community detection algorithms and may guide the design of new ones.

References

[1] D. W. Aha, D. Kibler, and M. K. Albert. Instance-based learning algorithms. *Machine Learning*, 6:37–66, January 1991.

[2] Y. Ahn, J. P. Bagrow, and S. Lehmann. Link communities reveal multiscale complexity in networks. *Nature*, 466(7307):761–764, 2010.

[3] V. D. Blondel, J. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. Mar. 2008. Journal of Statistical Mechanics.

[4] S. V. Dongen. Graph clustering via a discrete uncoupling process. *SIAM Journal on Matrix Analysis and Applications*, 30(1):121–141, 2008.

[5] M. A. Hall. *Correlation-based Feature Subset Selection for Machine Learning.* PhD thesis, Department of Computer Science, University of Waikato, 1999.

[6] G. Karypis and V. Kumar. A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM J. Sci. Comput.*, 20:359–392, December 1998.

[7] N. Mishra, R. Schreiber, I. Stanton, and R. Tarjan. Finding strongly knit clusters in social networks. *Internet Mathematics*, 5(1):155–174, Jan. 2008.

[8] M. Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577 –8582, June 2006.

[9] M. Rosvall and C. Bergstrom. Multilevel compression of random walks on networks reveals hierarchical organization in large integrated systems. *PLoS ONE*, 6(4):e18209, 04 2011.

[10] V. N. Vapnik. *Statistical Learning Theory.* Wiley-Interscience, 1st edition, Sept. 1998.

[11] E. Weinan, T. Li, and E. Vanden-Eijnden. Optimal partition and effective dynamics of complex networks. *Proceedings of the National Academy of Sciences*, 105(23):7907–7912, June 2008.

On the Convergence of the Hegselmann-Krause System

Arnab Bhattacharyya* Mark Braverman* Bernard Chazelle* Huy L. Nguyen*

August 27, 2012

1 Overview

Lately, there has been a surge of attention given to network-based dynamical systems, in which agents interact according to local rules via a dynamic communication graph [7]. Much of the previous work has focused on the exogenous case, where the communication topology is decoupled from the evolution of the system. We refer interested readers to [3] for a good overview of research in this area. In the more common, endogenous version, the communication graph changes over time according to the current states. The feedback loop between dynamics and topology creates considerable difficulties, and efforts have been underway to build an algorithmic calculus within the broad framework of *influence systems* [5]. This work investigates the complexity of *Hegselmann-Krause systems* (abbreviated as HK system from now on), a popular model of opinion dynamics that has proven highly influential over the years [1, 8, 9, 11, 13] and stands as the archetype of a diffusive influence system [6]. In the d-dimensional version of the model, each agent has an opinion represented as a point in $\mathbb{R}^d$. Two agents are neighbors if they are within unit distance from each other. At every time step, each agent moves synchronously to the mass center of its neighbors.

HK systems are known to converge [10, 12, 16], meaning that they eventually come to a full stop. The convergence time has been bounded by $n^{O(n)}$ time and conjectured to be polynomial [4]. It was shown to be $O(n^5)$ in the case $d = 1$ [14]. The contribution of this work is threefold: first, we prove that the convergence time is indeed polynomial in n, regardless of the dimension; second, we lower the one-dimensional bound from $O(n^5)$ to $O(n^3)$; third, we establish a quadratic lower bound, which improves the known bound of $\Omega(n)$ [14]. We also consider noisy variants of the model.

The bidirectionality of the system plays a crucial role in the proof, as it should. Indeed, it is known from [5] that allowing different radii and averaging weights for the agents can prevent the convergence of the communication graph. Our proofs are an elementary mix of geometric and algebraic techniques. Much of the current technology for HK systems centers around products of stochastic matrices. This work injects a geometric perspective that, we believe, will be necessary for further progress on the more difficult directional case.

2 Future Work

In this work, we analyze the convergence rate of the homogeneous HK system in arbitrary dimensions. The system is shown to converge in polynomial time, but can take at least a quadratic

---

*Princeton University, Computer Science Department and Center for Computational Intractability. Email: {`arnabb, mbraverm, chazelle, hlnguyen`}`@cs.princeton.edu`

number of steps in the worst case. Getting tight bounds on the convergence time of the system, even in just one dimension, remains an interesting open problem.

A particularly interesting challenge is to analyze the heterogeneous version of the HK system, i.e. when the neighborhood radii of all the agents are not necessarily the same [15]. New ideas are needed to understand the behavior of this system in particular, and directional systems in general. Our analysis of a noisy variant of homogeneous HK system is a step towards studying more complicated directional systems.

Beyond convergence rate, the behavior of the homogeneous HK system is still full of mysteries. Most notable perhaps is the 2R conjecture [2]: when agents are drawn uniformly at random on an interval, they converge to clusters at distance close to twice the minimum possible inter-cluster distance. Resolving this conjecture remains well beyond our understanding of the system.

References

[1] R. Axelrod. The dissemination of culture: a model with local convergence and global polarization. *Journal of Conflict Resolution*, 1997. Reprinted in "The complexity of cooperation," Princeton University Press, Princeton, 1997. 1

[2] V. D. Blondel, J. M. Hendricx, and J. N. Tsitsiklis. On the 2R conjecture for multi-agent systems. In *European Control Conference*, pages 2996–3000, July 2007. 2

[3] V. D. Blondel, J. M. Hendricx, and J. N. Tsitsiklis. On Krause's multi-agent consensus model with state-dependent connectivity. *IEEE Transactions on Automatic Control*, 54(11), Nov. 2009. 1

[4] B. Chazelle. The total s-energy of a multiagent system. *SIAM J. Control and Optimization*, 49(4):1680–1706, 2011. 1

[5] B. Chazelle. The dynamics of influence systems. *arXiv:1204.3946v2*, 2012. Prelim. version in Proc. 53rd FOCS, 2012. 1

[6] B. Chazelle. Natural algorithms and influence systems. *CACM Research Highlights*, 2012. 1

[7] D. A. Easley and J. M. Kleinberg. *Networks, Crowds, and Markets - Reasoning About a Highly Connected World.* Cambridge University Press, 2010. 1

[8] S. Fortunato. On the consensus threshold for the opinion dynamics of krause-hegselmann. *International Journal of Modern Physics C*, 16(2):259–270, 2005. 1

[9] R. Hegselmann and U. Krause. Opinion dynamics and bounded confidence: models, analysis and simulation. *J. Artificial Societies and Social Simulation*, 5(3), 2002. 1

[10] J. M. Hendrickx and V. D. Blondel. Convergence of different linear and non-linear Vicsek models. In *Proc. 17th International Symposium on Mathematical Theory of Networks and Systems (MTNS2006)*, pages 1229–1240, July 2006. 1

[11] U. Krause. A discrete nonlinear and non-autonomous model of consensus formation. In *Proc. Commun. Difference Equations*, pages 227–236, 2000. 1

[12] J. Lorenz. A stabilization theorem for dynamics of continuous opinions. *Physica A: Statistical Mechanics and its Applications*, 355(1):217–223, 2005. 1

[13] J. Lorenz. Continuous opinion dynamics under bounded confidence: a survey. *International Journal of Modern Physics C*, 16(18):1819–1838, 2007. 1

[14] S. Martínez, F. Bullo, J. Cortés, and E. Frazzoli. On synchronous robotic networks – Part ii: Time complexity of rendezvous and deployment algorithms. *IEEE Transactions on Automatic Control*, 52(12):2214–2226, Dec. 2007. 1

[15] A. Mirtabatabaei and F. Bullo. Opinion dynamics in heterogeneous networks: convergence, conjectures and theorems. *arXiv:1103.2829v2*, Mar. 2011. To appear in SIAM J. Control Optim. 2

[16] L. Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Transactions on Automatic Control*, 50(2):169 – 182, Feb. 2005. 1

[17] M. Roozbehani, A. Megretski, and E. Frazzoli. Lyapunov analysis of quadratically symmetric neighborhood consensus algorithms. In *CDC*, pages 2252–2257, 2008.

Variational Inference for Label Aggregation in Crowdsourcing

Qiang Liu[1], Jian Peng[2,3] and Alex Ihler[1]

1. University of Irvine; 2.TTI-C; 3. CSAIL, MIT

Crowdsourcing has become an efficient and inexpensive way to label large datasets in many application domains, including computer vision and natural language processing. Resources such as Amazon Mechanical Turk provide markets where the requestors can post tasks known as HITs (Human Intelligence Tasks) and collect large numbers of labels from hundreds of online workers (or annotators) in a short time and with relatively low cost. A major problem of crowdsoucing is that the qualities of the labels are often unreliable and diverse, mainly because it is difficult to control and monitor the performance of a large collection of workers. In the extreme, there may exist "spammers", who submit random answers rather than good-faith attempts to answer the question, or even "adversaries", who may deliberately give wrong answers, either due to malice or to a misinterpretation of the task. A common strategy to improve reliability is to add redundancy, such as assigning each task to multiple workers, and aggregate the workers' labels. The baseline majority voting heuristic, which simply assigns the label returned by the majority of the workers, is known to be error-prone, because it counts all the annotators equally. In general, efficient aggregation methods should take into account the differences in the workers labeling abilities.

A principled way to address this problem is to build generative probabilistic models for the annotation processes, and assign labels using standard inference tools. A line of early work builds simple models characterizing the annotators using confusion matrices, and infers the labels using the EM algorithm. Recently however, significant efforts have been made to improve performance by incorporating more complicated generative models. However, EM is widely criticized for having local optimality issue; this raises a potential tradeoff between more dedicated exploitation of the simpler models, either by introducing new inference tools or fixing local optimality issues in EM, and the exploration of larger model space, usually with increased computational cost and possibly the risk of over-fitting.

On the other hand, variational approaches, including the popular belief propagation (BP) and mean field (MF) methods, provide powerful inference tools for probabilistic graphical models. These algorithms are known to be efficient, often with provably strong local optimality properties, or even globally optimal guarantees. To our knowledge, no previous attempts have taken advantage of variational tools for the crowdsourcing problem. A closely related approach is a message passing algorithm in Karger et al. (referred to as KOS in the sequel), which the authors asserted to be motivated by but not equivalent to standard belief propagation. KOS was shown to have strong theoretical guarantees on (locally tree-like) random assignment graphs, but does not have an obvious interpretation as a standard inference method on a generative probabilistic model. As one consequence, the lack of a generative model interpretation makes it difficult to either extend KOS to more complicated models or adapt it to improve its performance on real-world datasets.

Contribution. In this work, we approach the crowdsourcing problems using tools and concepts from variational inference methods for graphical models. First, we present a belief-propagation-based method, which we show includes both KOS and majority voting as special cases, in which particular prior distributions are assumed on the workers' abilities. However, unlike KOS our method is derived using generative principles, and can be easily extended to more complicated models. On the other side, we propose a mean field method which we show closely connects to, and provides an important perspective on, EM. For both our BP and MF algorithms (and consequently for EM as well), we show that performance can be significantly improved by using more carefully chosen priors. We test our algorithms on both simulated and real-world datasets, and show that both BP and MF (or EM), with carefully chosen priors, is able to perform competitively with state-of-the-art algorithms that are based on far more complicated models.

Generalized Mean-Field Approximation for Opinion Spreading in Social Networks

Sungsu Lim
KAIST
ssungssu@kaist.ac.kr

Kyomin Jung
KAIST
kyomin@kaist.edu

How political opinion, product adoption, rumors, and trends diffuse through social networks has been a fundamental question for many decades in a wide variety of research disciplines. The spread of opinion on networks can be thought of as a state dynamics, where each node decides its state based on interactions with its neighbors. In many such settings, the dynamics of the states of nodes are described by some *Markov process* on a graph $G = (V, E)$ with a finite state space S. I.e., the state of a node v at each time step t is determined solely by the states of v and its neighbors at a time period $\{t-j, \cdots, t-1\}$ for some nonnegative integer j. Examples of such Markov processes include information diffusion models [7], the voter model [15], the pricing model [6], the naming game [14], and the gossip algorithm [4], some of which we will present in more detail later.

To predict the behavior of opinion spreading under the Markov process, mean-field (MF) approximation with ordinary differential equations (ODEs) has been widely used [2, 3, 5, 11]. Such an approach describes the state change rule as a system of ODEs that reflect the inherent Markovian state dynamics. In essence, the actual state ratio dynamics converges to the solution of the ODEs as the number of nodes goes to infinity. However, the analyses of its convergence are known only for symmetric network structures such as complete or bipartite graphs [8]. In our work, we propose a generalized MF method that relaxes the condition on strong symmetry, and prove the convergence to the ODE solution on any *slightly dense* graphs.

More precisely, consider a network structure $G = (V, E)$ with $n = |V|$ nodes so that each node has degree $\omega(\log n)$ (hence, *slightly dense*). Initially each node determines its state at time 0 randomly according to the state ratios $(s_i)_{i \in S}$ in an i.i.d. manner. Then we prove that the solution of the actual state dynamics of the Markov process converges to the solution of the ODEs. In addition, we also show that the ratios of the states among the neighbors of any given node is close to the actual state ratios of the entire network. As a general framework, our analysis can be applied to many opinion spreading processes in social networks caused by public service advertising, group-targeted marketing, and external influence [1,10]. Note that our result does not take any structural information of the network into account. Hence, surprisingly our results show that the MF approximation holds independently of the network structure, such as the community structure, as long as the initial states are drawn in an i.i.d. manner.

We adopt the standard (continuous) asynchronous time model [11] to express the state dynamics, where on average there are n state updates per unit time. Our result can also be applied to the slotted synchronous time model [4]. The following theorems formalize our main results.

Theorem 1 *Consider a Markov process with a finite state space S on a graph $G = (V, E)$ so that all nodes are of degrees $\omega(\log n)$, and with an initial state ratio $(s_i)_{i\in S} \in [0,1]^{|S|}$. For each time $t \in \mathbb{R}_+$, let $\mathbf{a}(t) = (a_i(t))_{i\in S}$ be the solution of the system of ODEs that corresponds to a given Markov process* [1]. *For $t \in \mathbb{R}_+$, let $\mathbf{b}(t) = (b_i(t))_{i\in S}$ be the random variable for the actual state ratio of V. Then, for any constants $\epsilon > 0$, $\delta > 0$, and $T > 0$,*

$$Pr\left(\sup_{0\le t\le T} \|\mathbf{b}(t) - \mathbf{a}(t)\|_1 < \epsilon\right) = 1 - o(n^{-\delta}). \tag{1}$$

This means that the overall state ratio of V converges to $\mathbf{a}(t)$ uniformly over $t \in [0, T]$ as n goes to infinity. Furthermore, the state ratio of the neighbors of each node at time $t \in [0, T]$ is also very close to $\mathbf{a}(t)$, as stated in the following theorem.

Theorem 2 *For $t \in \mathbb{R}_+$ and $v \in V$, let $\mathbf{b}_v(t) = (b_{v,i}(t))_{i\in S}$ be the random variable for the actual state ratio of the neighbors of v. Under the same condition as in Theorem 1, for any constants $\epsilon > 0$, $\delta > 0$, and $T > 0$,*

$$Pr\left(\sup_{0\le t\le T, v\in V} \|\mathbf{b}_v(t) - \mathbf{a}(t)\|_1 < \epsilon\right) = 1 - o(n^{-\delta}). \tag{2}$$

We conducted experiments on the network datasets consisting of two synthetic networks (Watts-Strogatz (WS) and the Barabási-Albert (BA)) of 10,000 nodes with average degree 100, and two real-world networks (ePinions [12] (75,879 nodes with average degree 13.4) and Slashdot [9] (77,360 nodes with average degree 23.4)) [2]. The opinion spreading models we used were the linear threshold model [7] and the ternary voter model [11]. Our empirical results confirmed that Theorem 1 and Theorem 2 are very accurate for WS and BA networks, and are validated quite well for real-world networks where the slightly dense condition is violated. Finally, we present the details of some of the popular Markovian opinion spreading models where our results can be applied.

- *Information diffusion models* [3,7,10]: In many information diffusion models, including the general threshold model and the independent cascade model, the state space can be expressed by $S = \{0(\text{inactive}), 1(\text{active})\}$. In the general threshold model, each node v becomes active if a function evaluated on the states of its neighbors exceeds the threshold of v. In the independent cascade model, every adoptor v has a single chance to influence its non-adopted neighbor u with a certain probability. Our work can also be applied to non-monotonic information diffusion models, such as SIS model and its recent variants [3], and the information diffusion model with external influence [10].

---

[1] Solve $\mathbf{a}(t)$ by a standard MF method with the system of ODEs.

[2] `http://snap.stanford.edu`

- *Voter model* [5,11]: In the voter model, the state space corresponds to the set of candidates or items to vote, and a node v either updates its state by copying that of its neighbor chosen uniformly at random with some probability, or preserves its state with the remaining probability.

- *Naming game* [13,14]: Naming game was originally intended to model language diffusion in a society. This model has been widely used for describing how a multi-agent system can converge towards a consensus state in a self-organized way. Similarly to the voter model, a node (listener) is selected, and it interacts with a randomly chosen neighbor (speaker). The speaker randomly selects a language from its language list and sends it to the listener. Then, the listener updates its language list.

References

[1] M.-F. Balcan, A. Blum, and Y. Mansour. Improved equilibria via public service advertising. In *Proceedings of SODA*, 2009.

[2] M. Benaïm and J. Weibull. Mean-field approximation of stochastic population processes in games. 2009.

[3] A. Beutel, B. A. Prakash, R. Rosenfeld, and C. Faloutsos. Interacting viruses in networks: Can both survive? In *Proceedings of ACM SIGKDD*, 2012.

[4] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transactions on Networking*, 14, 2005.

[5] K. Jung, B. Y. Kim, and M. Vojnovic. Distributed ranking in networks with limited memory and communication. In *Proceedings of ISIT*, 2012.

[6] M. L. Katz and C. Shapiro. Network Externalities, Competition, and Compatibility. *The American Economic Review*, 75, 1985.

[7] D. Kempe, J. Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of ACM SIGKDD*, 2003.

[8] T. Kurtz. *Approximation of Population Processes.* SIAM, 1981.

[9] J. Leskovec, K. Lang, A. Dasgupta, and M. Mahoney. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*, 6, 2009.

[10] S. Myers, C. Zhu, and J. Leskovec. Information diffusion and external influence in networks. In *Proceedings ACM SIGKDD*, 2012.

[11] E. Perron, D. Vasudevan, and M. Vojnović. Using three states for binary consensus on complete graphs. In *Proceedings of INFOCOM*, 2009.

[12] M. Richardson, R. Agrawal, and P. Domingos. Trust management for the semantic web. In *Proceedings of ISWC*, 2003.

[13] L. Steels. A self-organizing spatial vocabulary. *Artificial Life*, 2, 1995.

[14] J. Xie, S. Sreenivasan, G. Korniss, W. Zhang, C. Lim, and B. K. Szymanski. Social consensus through the influence of committed minorities. *Phys. Rev. E*, 84, 2011.

[15] M. Yildiz, R. Pagliari, A. Ozdaglar, and A. Scaglione. Voting models in random networks. In *Proceedings of ITA*, 2010.