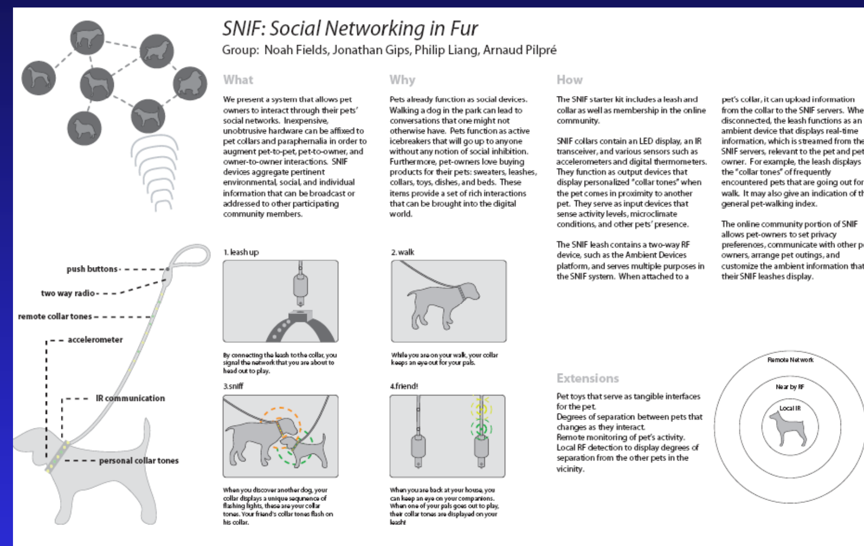
Discovery, Diagnosis, and Design of Team Networks



Noshir Contractor

Jane S. & William J. White Professor of Behavioral Sciences

Professor of Ind. Engg & Mgmt Sciences, McCormick School of Engineering

Professor of Communication Studies, School of Communication &

Professor of Management & Organizations, Kellogg School of Management,

Director, Science of Networks in Communities (SONIC) Research Laboratory

nosh@northwestern.edu

Supported by NSF : IIS-0838564, OCI-0753047, IIS-0729505, IIS-0535214

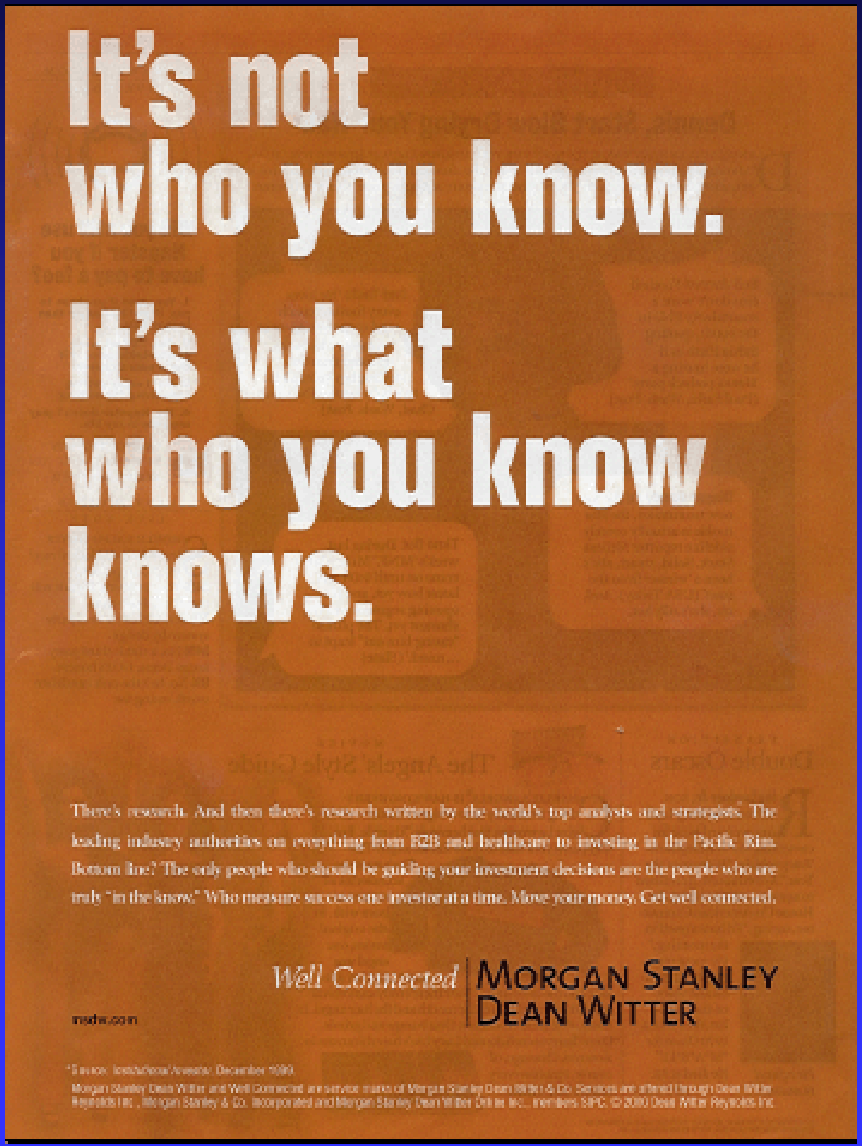
SONIC

Advancing the Science of Networks in Communities

Aphorisms about Networks

- Social Networks:
  - Its not what you know, its **who** you know.
- Cognitive Social Networks:
  - Its not who you know, its **who they think** you know.
- Knowledge Networks:
  - Its not who you know, its **what they think** you know.

Cognitive Knowledge Networks



SONIC

Advancing the Science of Networks in Communities

Multidimensional Networks in Team Science

Multiple types of Nodes and Multiple Types of Relationships

The Hubble telescope: $2.5 billion

Source: David Lazer

SONIC

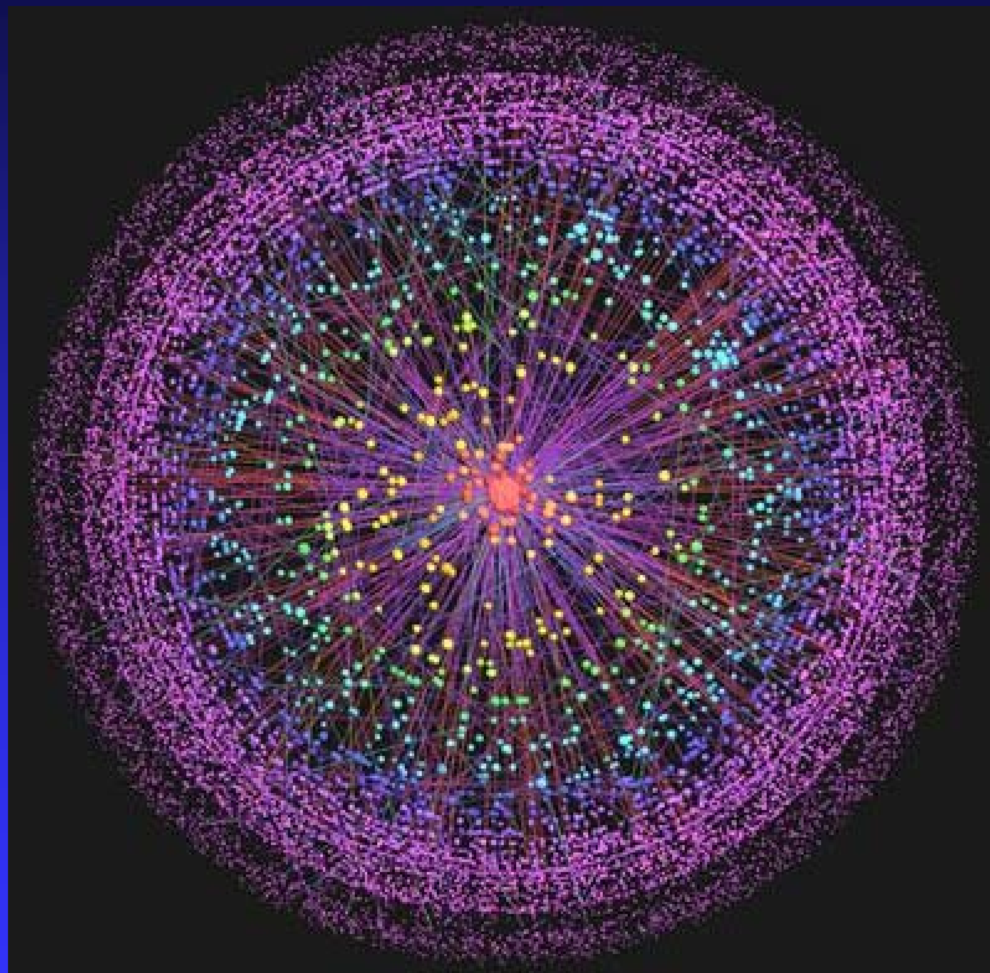
Advancing the Science of Networks in Communities

CERN particle accelerator: $1 billion/year

Source: David Lazer

SONIC
Advancing the Science of
Networks in Communities

The Web: priceless*



* *Apologies to MasterCard*

Source: David Lazer

SONIC

Advancing the Science of
Networks in Communities

Digital Harvesting of Relational Metadata

Bios, titles & descriptions

Personal Web sites Google search results

Web of Science Citation

CRAWDAD TECHNOLOGIES LLC

CATPAC

AutoMap

D2K data to knowledge

UBERLINK

issuecrawler

ISI Web of Knowledge
We < Deliver > UNMATCHED QUALITY

IKNOW
Inquiring Knowledge Networks on Web

CI-KNOW Analyses and Visualizations

SONIC
Advancing the Science of Networks in Communities

SOCIAL SCIENCE

Computational Social Science

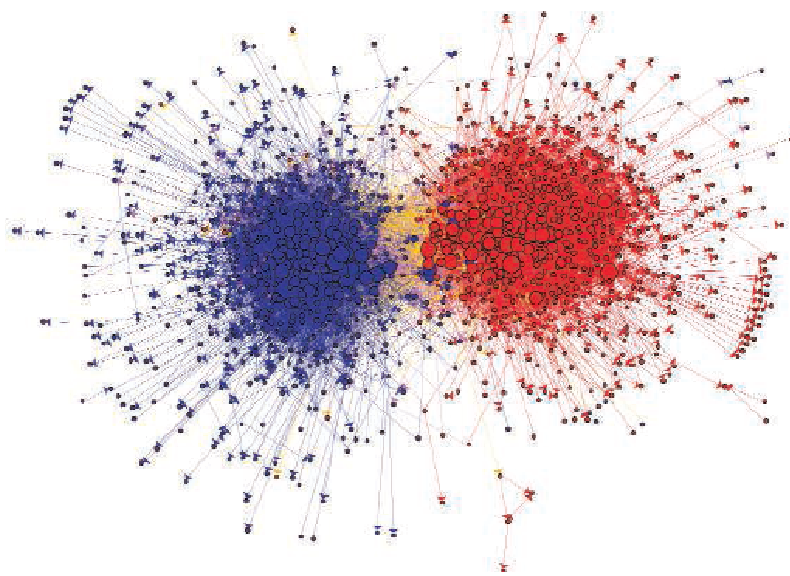
A field is emerging that leverages the capacity to collect and analyze data at a scale that may reveal patterns of individual and group behaviors.

David Lazer,[1] Alex Pentland,[2] Lada Adamic,[3] Sinan Aral,[2,4] Albert-László Barabási,[5] Devon Brewer,[6] Nicholas Christakis,[1] Noshir Contractor,[7] James Fowler,[8] Myron Gutmann,[3] Tony Jebara,[9] Gary King,[1] Michael Macy,[10] Deb Roy,[2] Marshall Van Alstyne[2,11]

We live life in the network. We check our e-mails regularly, make mobile phone calls from almost any location, swipe transit cards to use public transportation, and make purchases with credit cards. Our movements in public places may be captured by video cameras, and our medical records stored as digital files. We may post blog entries accessible to anyone, or maintain friendships through online social networks. Each of these transactions leaves digital traces that can be compiled into comprehensive pictures of both individual and group behavior, with the potential to transform our understanding of our lives, organizations, and societies.

The capacity to collect and analyze massive amounts of data has transformed such fields as biology and physics. But the emergence of a data-driven "computational social science" has been much slower. Leading journals in economics, sociology, and political science show little evidence of this field. But computational social science is occurring—in Internet companies such as Google and Yahoo, and in government agencies such as the U.S. National Security Agency. Computational social science could become the exclusive domain of private companies and government agencies. Alternatively, there might emerge a privileged set of academic researchers presiding over private data from which they produce papers that cannot be critiqued or replicated. Neither scenario will serve the long-term public interest of accumulating, verifying, and disseminating knowledge.

What value might a computational social science—based in an open academic environment—offer society, by enhancing understanding of individuals and collectives? What are the



Data from the blogosphere. Shown is a link structure within a community of political blogs (from 2004), where red nodes indicate conservative blogs, and blue liberal. Orange links go from liberal to conservative, and purple ones from conservative to liberal. The size of each blog reflects the number of other blogs that link to it. [Reproduced from (*8*) with permission from the Association for Computing Machinery]

[1]Harvard University, Cambridge, MA, USA. [2]Massachusetts Institute of Technology, Cambridge, MA, USA. [3]University of Michigan, Ann Arbor, MI, USA. [4]New York University, New York, NY, USA. [5]Northeastern University, Boston, MA, USA. [6]Interdisciplinary Scientific Research, Seattle, WA, USA. [7]Northwestern University, Evanston, IL, USA. [8]University of California–San Diego, La Jolla, CA, USA. [9]Columbia University, New York, NY, USA [10]Cornell University, Ithaca, NY, USA. [11]Boston University, Boston, MA, USA. E-mail: david_lazer@harvard.edu. Complete affiliations are listed in the supporting online material.

Published by AAAS

SONIC

Advancing the Science of Networks in Communities

Projects Investigating Social Drivers for Teams

Science Applications

CI-Scope: Understanding & Enabling CI in Virtual Communities *(NSF)*

NUCATS: Clinical & Translational Science *(NIH)*

VOSS: NanoHub (NSF)

TSEEN: Tobacco Surveillance Evaluation & Epidemiology Network *(NSF, NIH, CDC)*

Business Applications

PackEdge Community of Practice *(P&G)*

Kraft Design Teams

Core Research

Socio-technical Drivers for Understanding & Enabling Teams

Societal Justice Applications

Mapping Climate Change Networks In Low Income Communities *(City of Chicago)*

Mapping Digital Media and Learning Networks *(MacArthur Foundation)*

Entertainment Applications

Second Life (*NSF, Army Research Institute,* Linden Labs)

EverQuest II (NSF, Army Research Institute, Linden Labs)

SONIC

Advancing the Science of Networks in Communities

The Assembly of Task-oriented Groups

Yun Huang, Mengxiao Zhu, Jing Wang, Brian Keegan & Noshir Contractor,
Northwestern University

Nishith Pathak
University of Minnesota

Cuihua Shen, Dmitri Williams
University of Southern California

Supported by NSF IIS-0729505, Army Research Institute (W91WAW-08-C-0106), and Sony Online Entertainment

SONIC
Advancing the Science of
Networks in Communities

Using Digital Traces to Understand Team Assembly

- Massively-multiplayer online games (MMOGs) have over 45 million users worldwide and over $3 billion in revenue in 2008

- What does social behavior in online worlds tell us about the “real” world and vice versa?
  - Online games exhibit features that map onto real world processes:
    - Social networks, economics, groups, communication, conflict, expertise, leadership, crime, innovation, epidemics, etc.
  - Online games already capture the signatures of these behaviors in huge databases, just waiting to be analyzed

Hypotheses

- Team formation mechanisms

H1: Players who have low combat ability are more likely to participate in teams than those who have high combat ability. *(Self-interest)*

H2: Players are more likely to join the same set of players multiple times. *(Reduce Coordination cost)*

H3a: Players are more likely to join teams of high expertise diversity. *(Transactive Memory)*

H3b: Players are more likely to join teams in which they can provide unique expertise. *(Transactive Memory)*

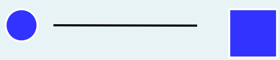
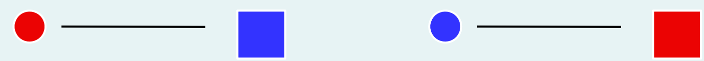
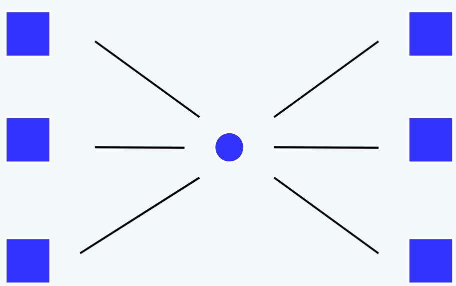
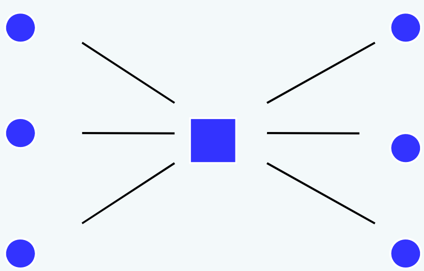
Hypotheses (cont.)

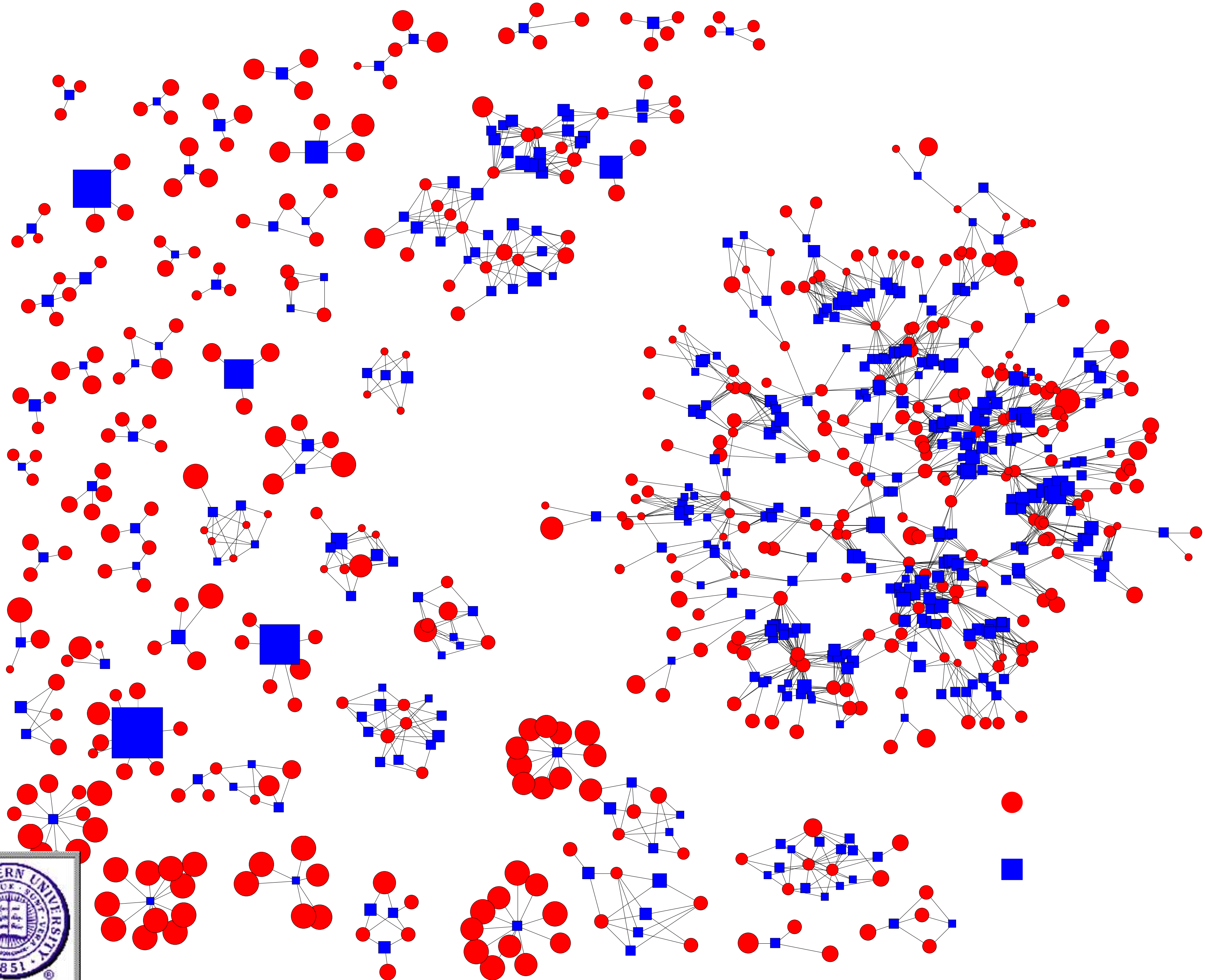
- Group outcome
  - H4: Teams with many players are more likely to have member death. *(Higher Coordination cost)*
  - H5: Teams with many players tend to have higher performance. *(Mutual interest)*
  - H6: Teams with many players have shorter duration. *(Higher Coordination Cost)*

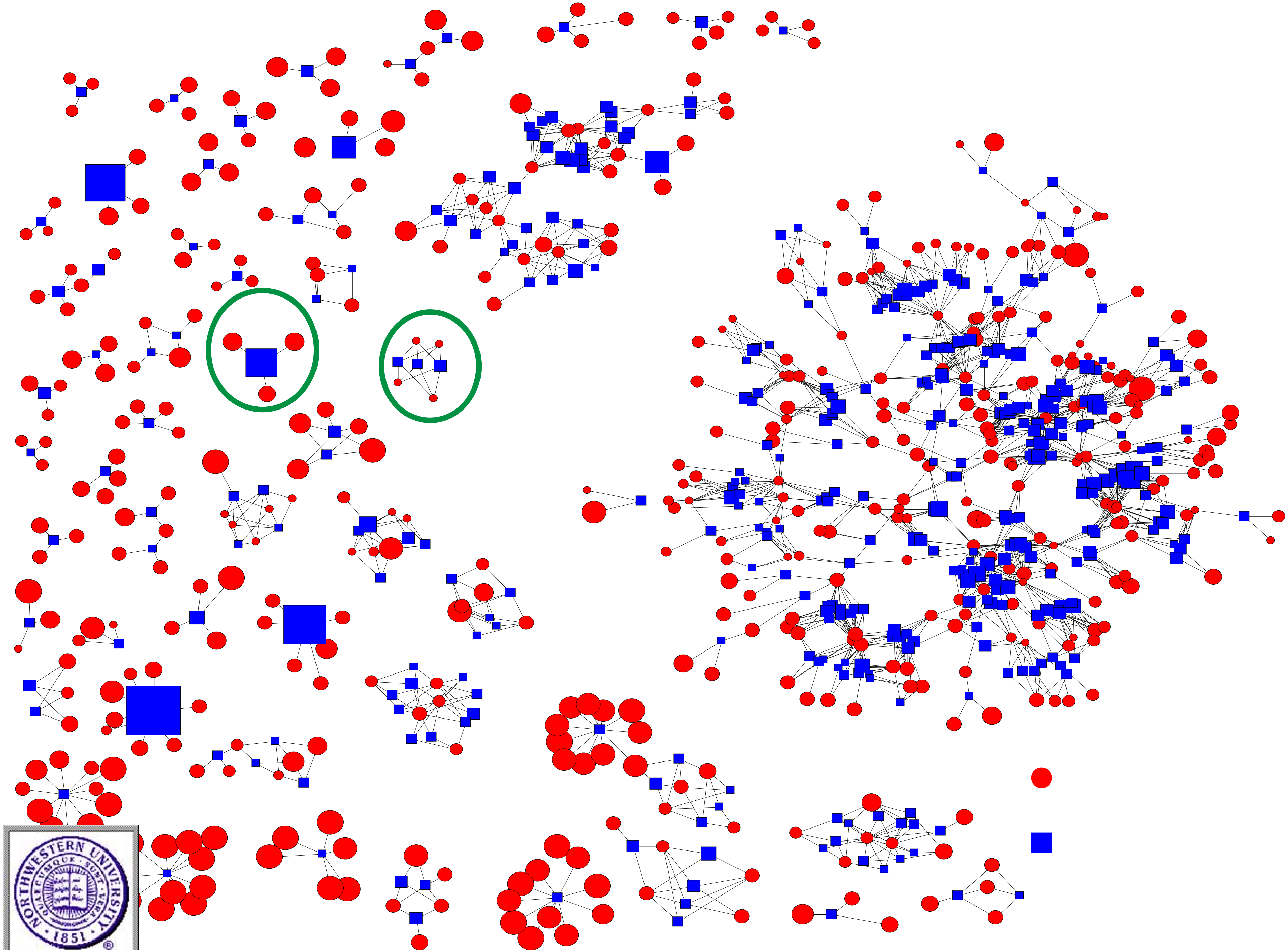
SONIC

Advancing the Science of Networks in Communities

Team Formation Structures

Team membership 	Selective team membership 
Membership in multiple teams 	Teams with many members 
Notes: ● Characters whose attributes are <u>not</u> taken into consideration ● Characters whose attributes are taken into consideration	—— Team membership ■ Teams whose attributes are <u>not</u> taken into consideration ■ Teams whose attributes are taken into consideration





Descriptive Statistics for the Zones

- In the whole dataset, there are 2,774 characters, 3,547 group events; 15,152 group membership links.
- divide it into 11 zones based on the game map

Zone #	Zone Name	# of char	# of group	Median Level	Mean Group Size
1	Thundering Steppes	639	591	29	4.15
2	Kingdom of Sky	625	436	65	4.80
3	The Enchanted Lands	530	537	38	4.48
4	Desert of Flames	499	518	53	4.36
5	Antonica	465	396	21	4.04
6	Commonlands	380	315	24	4.01
7	Nektulos Forest	287	161	36	3.92
8	Feerrott	269	206	45	4.45
9	Everfrost	211	165	45	4.36
10	Lavastorm	198	141	49	4.51
11	Zek	170	81	40	3.95

Results: Antonica as An Example

Findings	Coefficient
Low level players are more likely to join groups. (H1: Supported)	**-0.01***
Players are more likely to join the same set of players for multiple times. (H2: Not supported)	-0.11
Players are more likely to join groups of high expertise diversity. (H3a: Supported)	**4.24***
Players are more likely to join groups in which they can provide unique expertise (H3b: Partially supported) Supported for priests but the other character classes do not show such a tendency.	**-1.27* (Priest)** 0.03 (Mage) -0.07 (Scout)
Groups of larger size are more likely to have member death. (H4: Supported)	**0.60***

Notes:
* indicate twice of standard deviation
Green indicates results supporting the hypotheses; **black** indicates non-significant results; red indicates results in the opposite direction.

Results: Antonica as An Example

Findings	Coefficient
Groups with many players gain higher performance. (H5: Supported)	0.005*
Groups with many players have shorter duration. (H6: Supported)	-0.33*
Players are active in joining groups.	5.00*
Players tend to join multiple groups (or group events).	0.79*
Combat groups tend to be small.	-7.44*
Compared to fighters, priests are more likely to join a group, but mages or scouts are not.	0.92* (priest) -0.05 (mage) 0.004 (scout)

Notes: * indicate twice of standard deviation

Green indicates results supporting the hypotheses;

Black indicates non-significant results;

Red indicates results in the opposite direction.

SONIC

Advancing the Science of Networks in Communities

Results Summary

- Players are active in joining groups, especially those at lower levels.
- Players are more likely to join the groups that 1) have higher expertise diversity and 2) to which they can provide unique expertise (especially for priest and mage).
- Groups with more members tend to 1) have higher performance, 2) last a shorter time, and 3) be more likely to have member death during the combat.
- Players tend to join multiple groups, and most groups are of small size.

Projects Investigating Social Drivers for Teams

Science Applications

CI-Scope: Understanding & Enabling CI in Virtual Communities *(NSF)*

NUCATS: Clinical & Translational Science *(NIH)*

VOSS: NanoHub (NSF)

TSEEN: Tobacco Surveillance Evaluation & Epidemiology Network *(NSF, NIH, CDC)*

Business Applications

PackEdge Community of Practice *(P&G)*

Kraft Design Teams

Core Research

Socio-technical Drivers for Understanding & Enabling Teams

Societal Justice Applications

Mapping Climate Change Networks In Low Income Communities *(City of Chicago)*

Mapping Digital Media and Learning Networks *(MacArthur Foundation)*

Entertainment Applications

Second Life (*NSF, Army Research Institute,* Linden Labs)

EverQuest II (NSF, Army Research Institute, Linden Labs)

SONIC

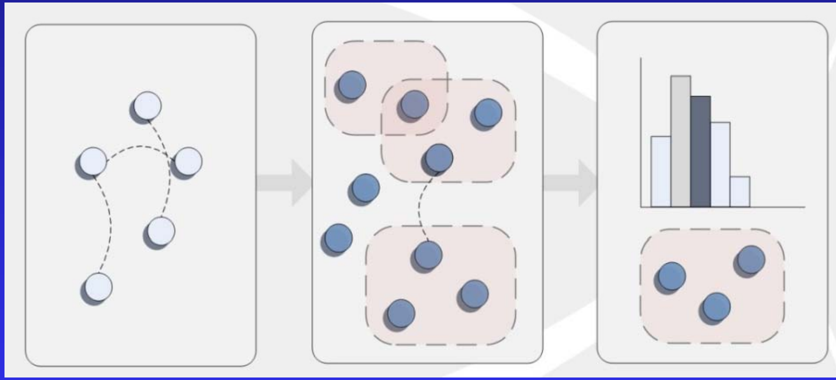
Advancing the Science of Networks in Communities

The Impacts of Co-authorship Networks and Citation Networks in “Team Science”*

By Meikuan Huang, Jordan Liu, Annie Wang, & Noshir Contractor

- **“Group-staffing riddle”** (Huber & Lewis, 2010):

  How to assembly a group to obtain both
  (1) high productivity based on diversity of expertise and cognitive models &
  (2) smooth coordination and communication among group members with shared cognitive models



- Our goal: To discover how prior co-authorship and citation network configurations influence team formation and success in scientific research groups.

*Funded by NIH/NCRR grant for Northwestern University Clinical and Translational Sciences Institute (NU-CATS) (2008-2013).

Theoretical Background

(1) Transactive memory (TM)

- Shared cognitive models or directories of “who knows what” among group members (Hollingshead, 1997, 1998; Wegner, 1995).
- A key TM dimension: Sharedness of knowledge at the group level, or the extent to which all members have similar perceptions of each other’s task responsibilities and expertise level in different knowledge areas (Brandon & Hollingshead, 2004; Huber &Lewis, 2010)

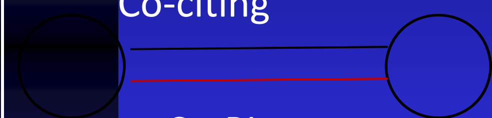
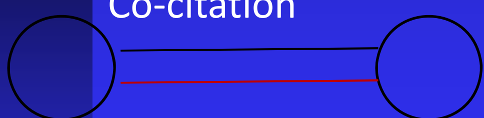
(2) Prior collaboration

- People are likely to prefer partners with whom they are already familiar from prior work on joint projects (Hinds, Carley, Krackhardt, & Wholey, 2000)

(3) Homophily

- The tendency of individuals to interact more with those to whom they are more similar (Ibarra, 1992; McPherson & Smith-Lovin, 1987)
- Reasons: Ease of communication, shared understandings and comfort (Carley, 2002).

Hypotheses & Analysis

H1	Co-authorship  Co-PI	Researchers tend to collaborate on proposal teams with those with whom they have a co-authorship relationship.
H2	Co-citing  Co-PI	Researchers tend to collaborate on proposal teams with those with whom they have a citation relationship.
H3	Co-citation  Co-PI	Researchers who cite similar publications are more likely to collaborate on proposal teams.

Analysis:

•ERGM models (Exponential Random Graph Modeling) (Frank & Strauss, 1986; Robins & Pattison, 2005; Wasserman & Pattison, 1996)

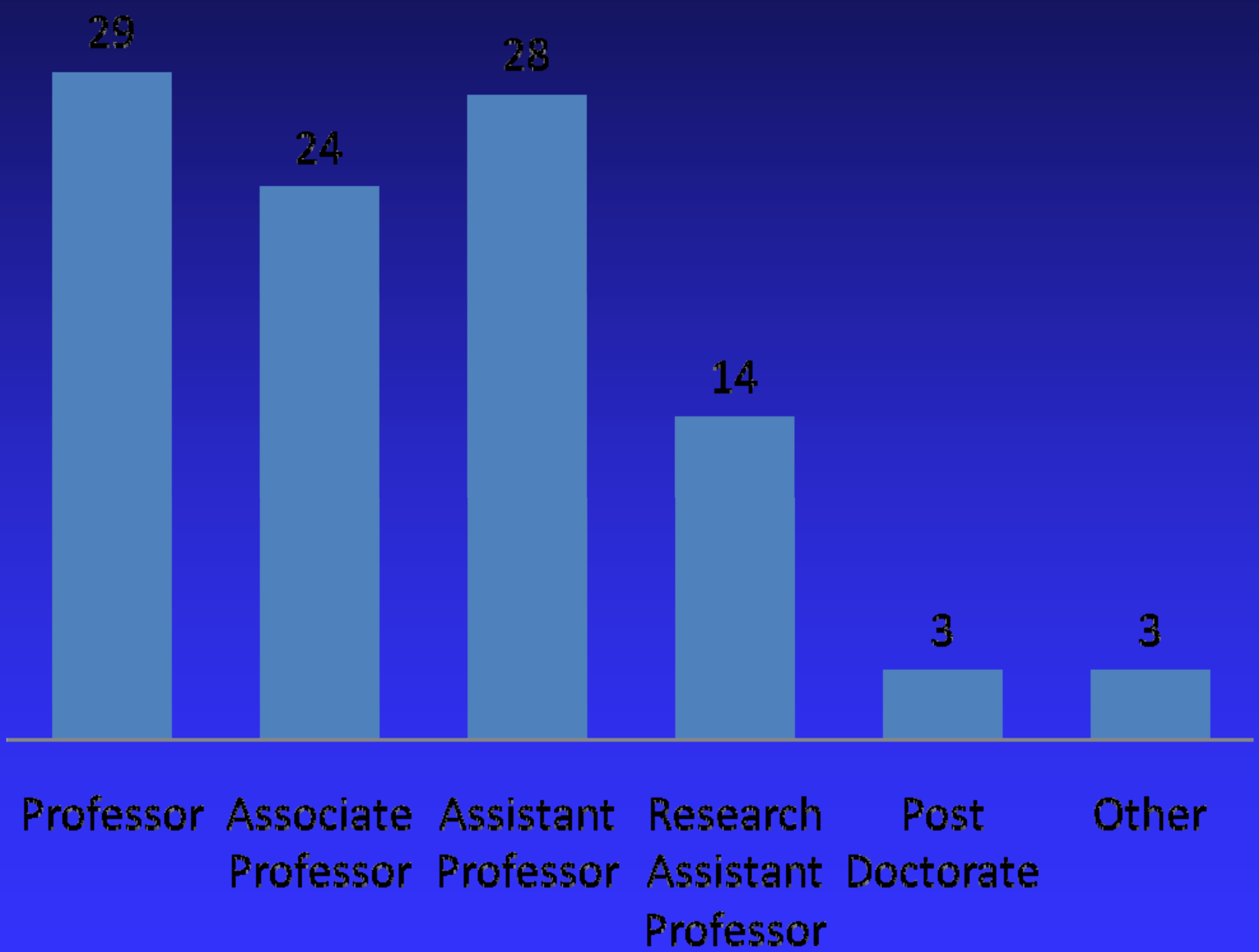
• PNet (Wang, Robins, & Pattison, 2006).

Data

- 60 Proposals
- 117 applicants, with 60 PIs and 57 Co-PIs, totally
- 37 departments in total

Tenure Distribution

- 29 Professors
- 24 Associate Professors
- 28 Assistant Professors
- 14 Research Assistant Professors
- 3 Post Docs
- 3 Others: student, research scientist, adjunct assistant professor
- 101 in data

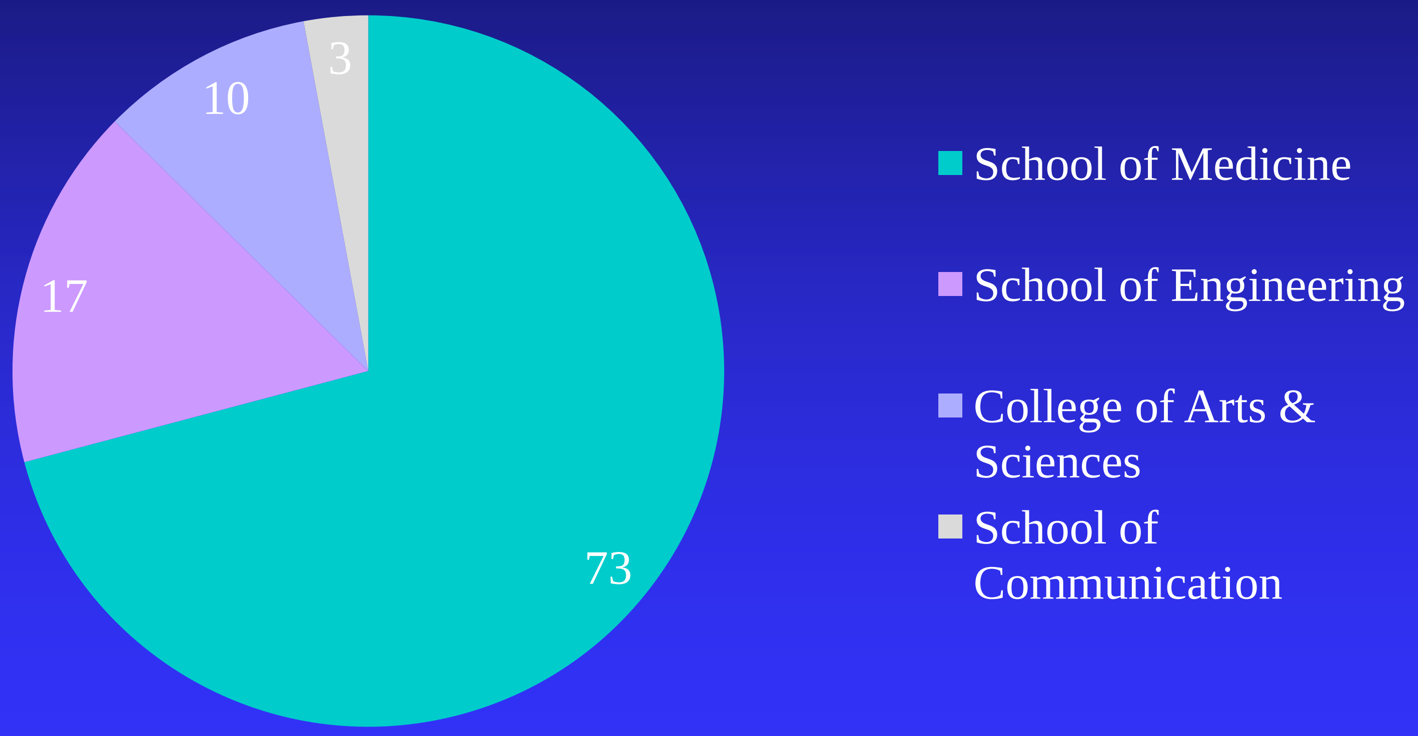


Departments

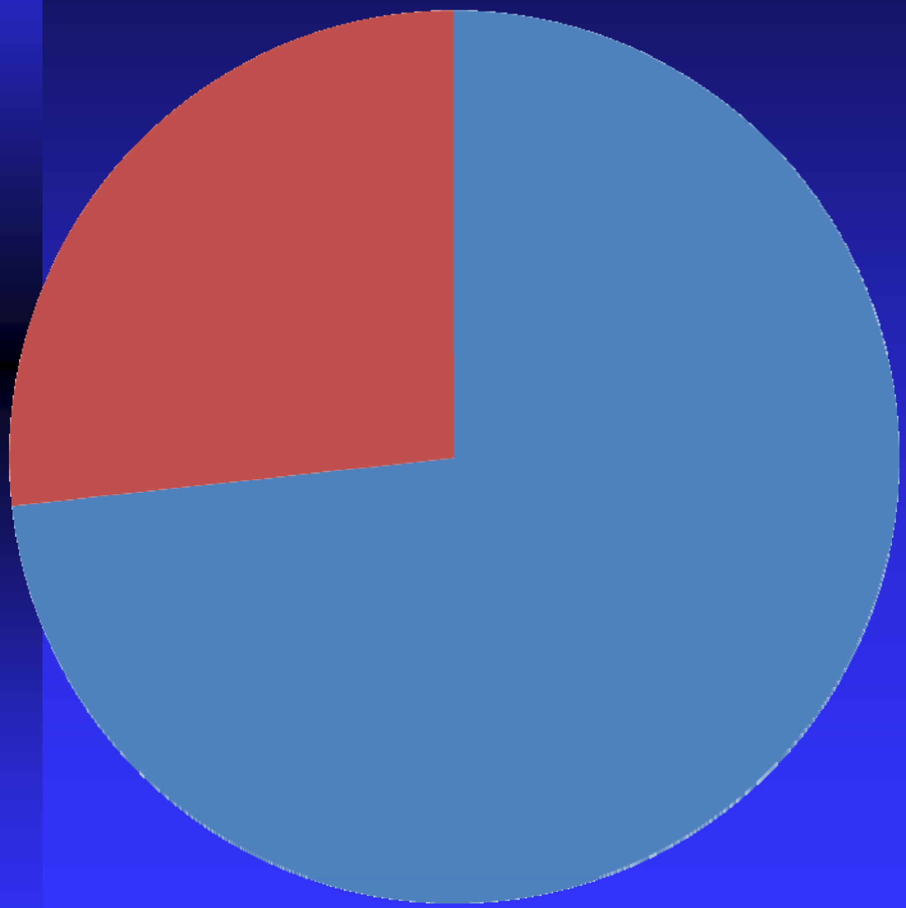
Department	Number of applicants in the department
Physical Medicine and Rehabilitation	9
Surgery	8
Biomedical Engineering *	6
Cardiology	6
Pediatrics	5
Chemistry *	4
Hematology Oncology	4
Infectious Disease	4
Molecular Pharmacology	4
All others	3 or less

* Indicates that the department is outside the medical school.

Applicant Distribution Across Schools



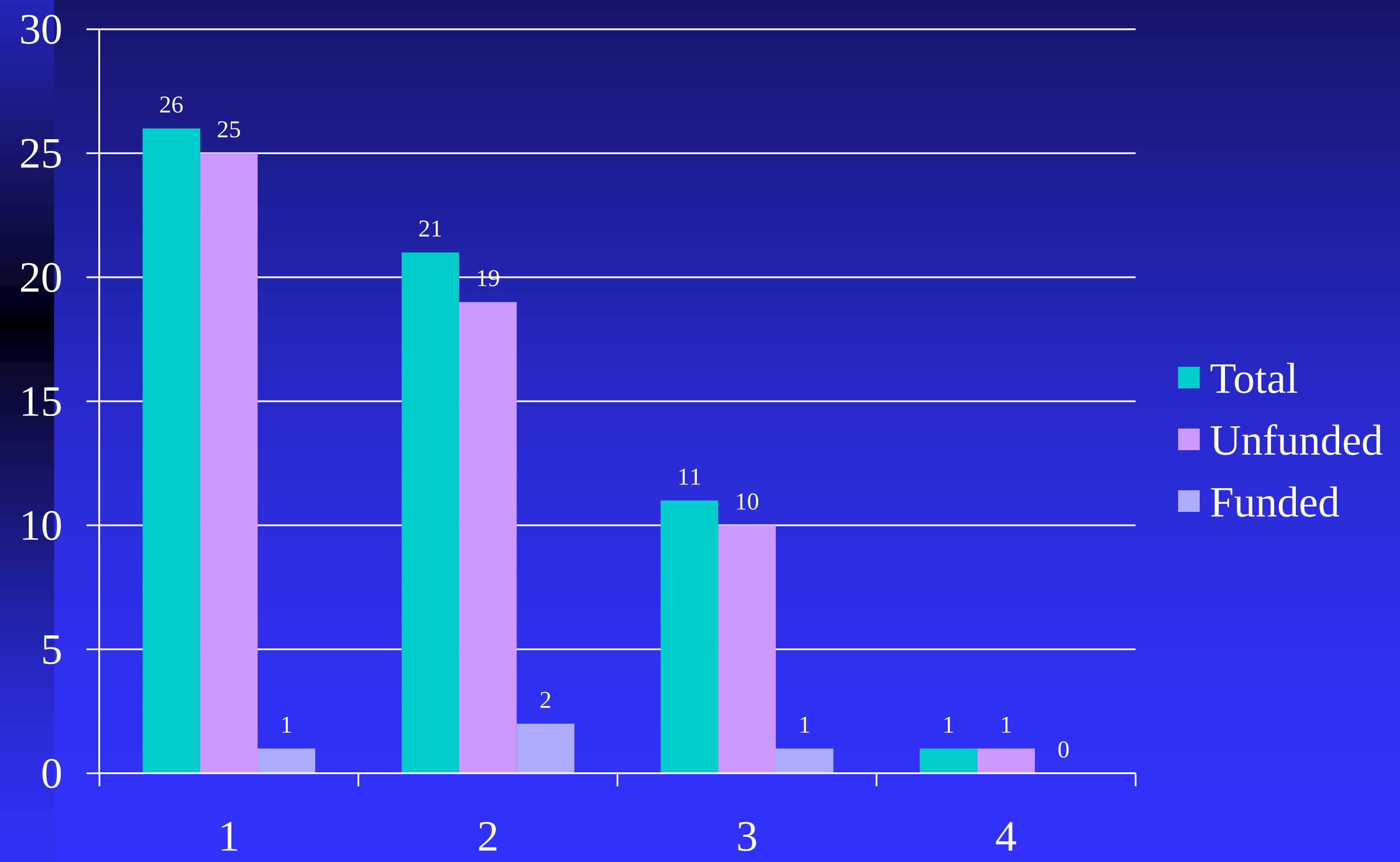
Gender Distribution



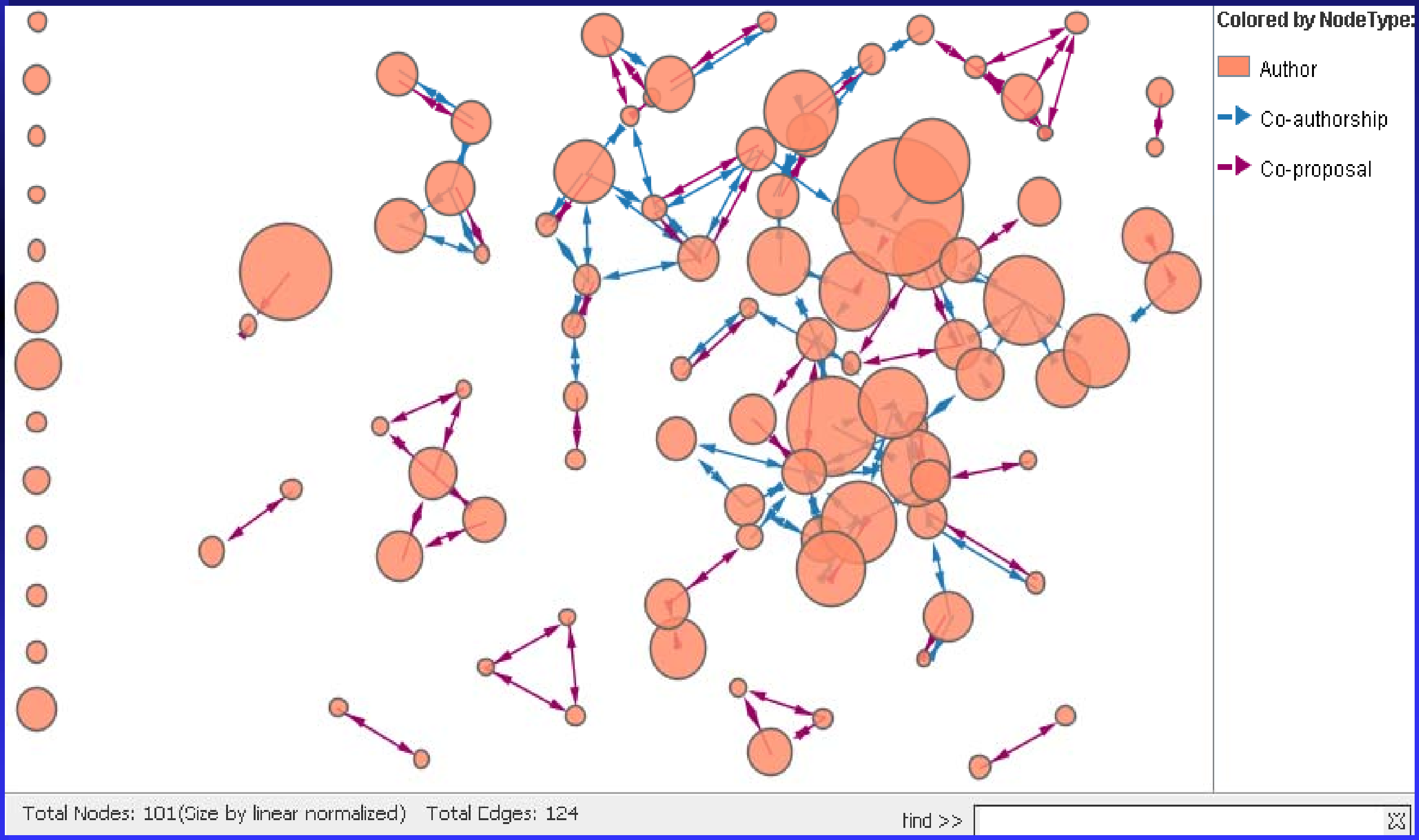
74 males (72%)

27 females (28%)

Number of Applicants in the Proposal

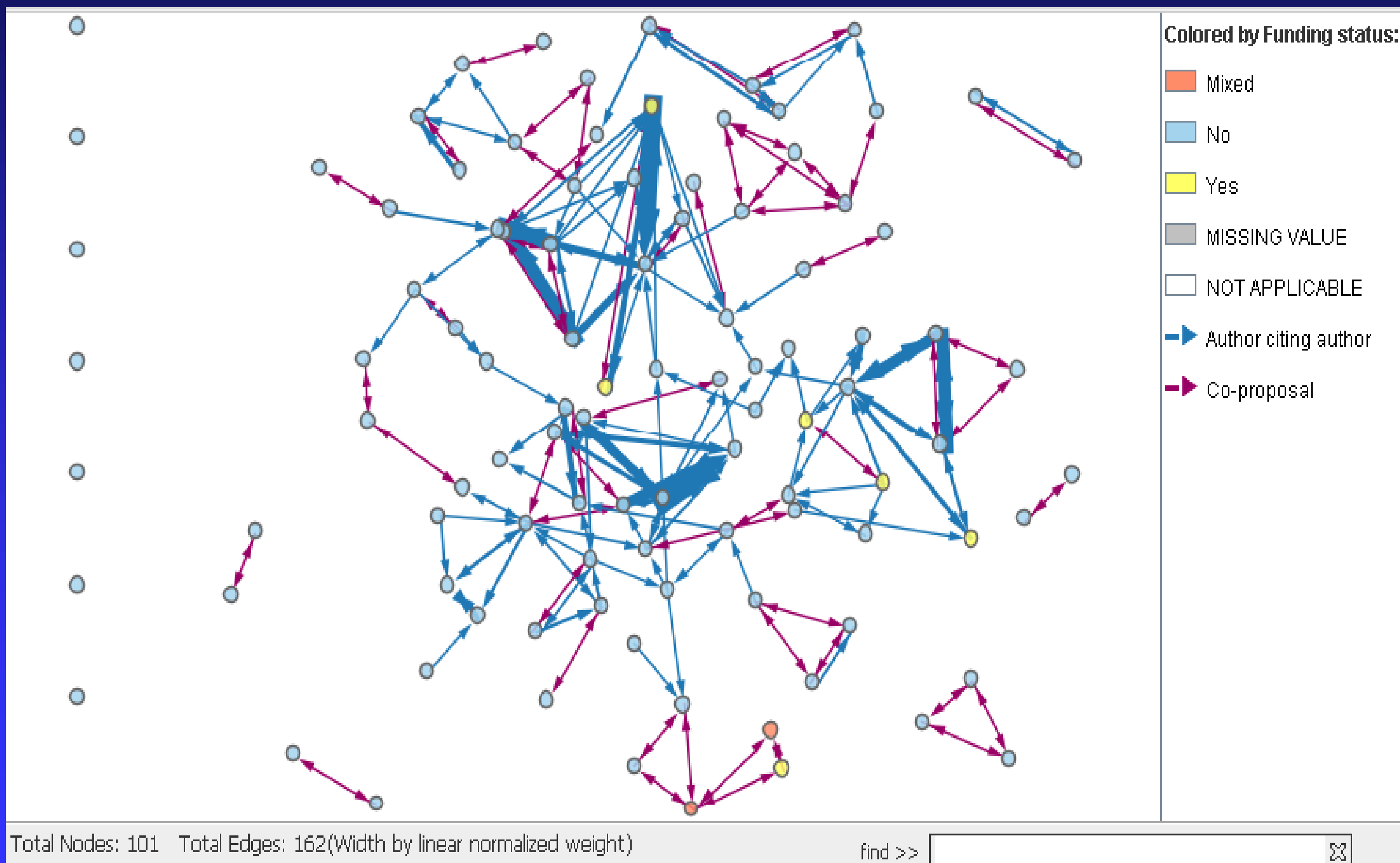


H1: Co-proposal & Co-authorship Network

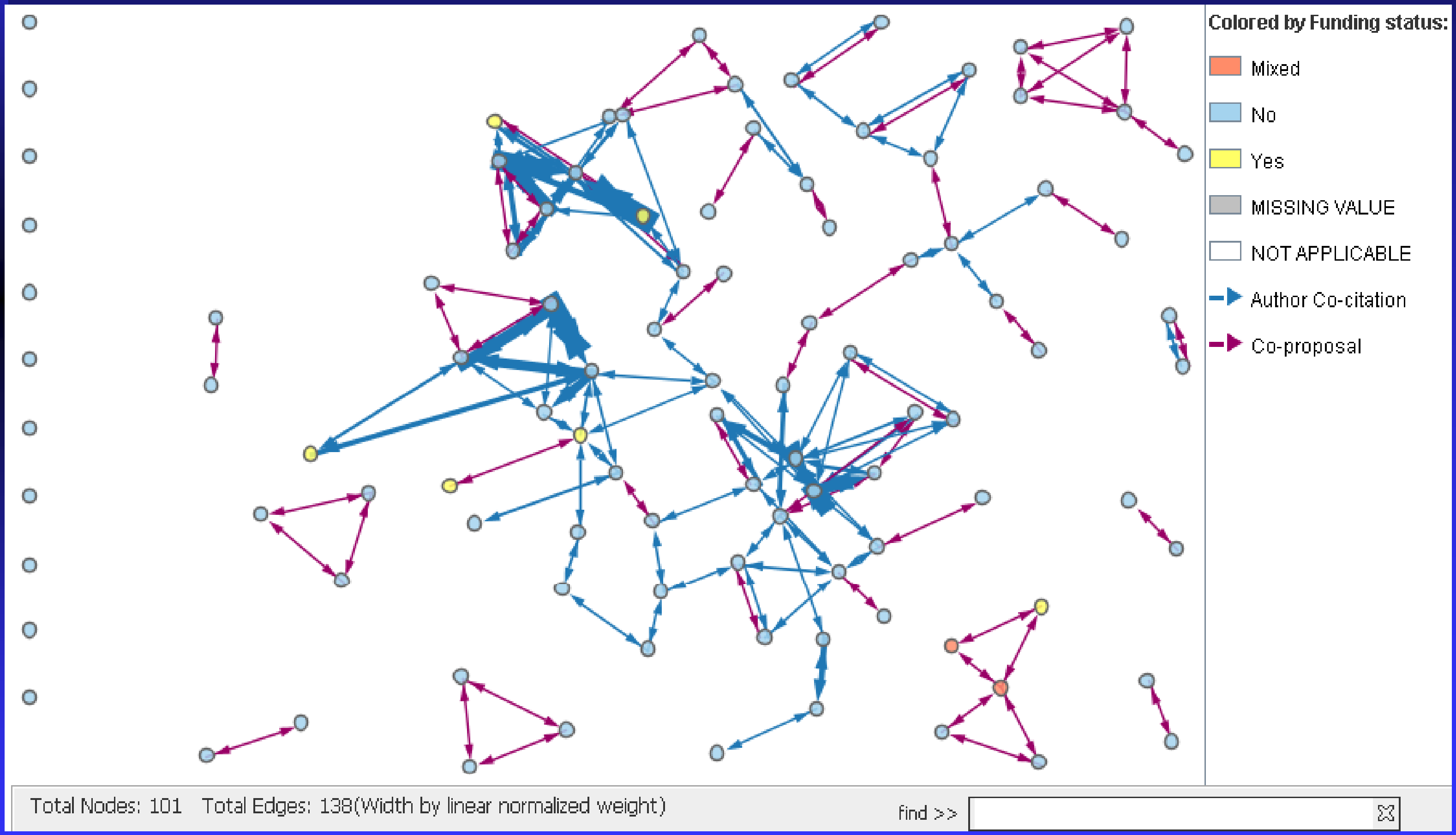


Node size indicates the # of publications

H2: Co-proposal & co-citation network



H3: Co-proposal & Citing network



	Analysis	Effects	Estimates	SD		
Control	PNet	Edge (DV: co-proposal)	-4.89	0.17	-0.02	*
Control	PNet	2-star (with co-citation as covariate)	-0.46	0.21	[illegible]	[illegible]
Control	PNet	2-star (with citing as covariate)	-0.45	0[illegible]	[illegible]	[illegible]
H1	PNet	coauthorship_edge	3.67	0.32	0.04	*
H2	PNet	citing_edge	2.78	0.33	-0.05	*
H3	PNet	co-citation_edge	2.96	0.37	0.03	*

Researchers are not likely to randomly form a project collaboration relationship with each other.

Researchers are more likely to have better familiarity of and collaborate again with those they share a collaboration history (co-authorship or citing each other).

Researchers are also more likely to collaborate with those who cited similar articles in their publications.

Funded vs. Unfunded

	Funded (N = 8)		**Unfunded** (N = 93)	
Effects	Estimates	SD	Estimates	SD
Edge (co-proposal)	-3.28	1.07	-4.33	0.13
Co-author	6.95	7.14	0.34	1.06
Cite one another	7.32	4.61	-2.93	4.37
Cite same sources	6.61	7.99	-4.17	15.83

3D Strategy for Enabling Team Science

- **D iscovery**: Effectively and efficiently foster network links from people to other people, knowledge, and artifacts (data sets/streams, analytic tools, visualization tools, documents, etc.)
  - ***"If only NSF knew what NSF knows".***
- **Diagnosis**: Assess the "health" of internal and external networks - in terms of scanning, absorptive capacity, diffusion, robustness, and vulnerability to external environment
- **Design**: Model or re-wire networks using social and organizational incentives (based on social network research) and network referral systems to enhance emergent and mature teams

SONIC

Advancing the Science of Networks in Communities

Design Examples: Mapping & Enabling Networks in …

Tobacco Research: TobIG Demo

Computational Nanotechnology: nanoHUB Demo

Cyberinfrastructure: CI-Scope Demo

Oncofertility: Onco-IKNOW

SONIC

Advancing the Science of Networks in Communities

Summary

- The Science of Team Science is well poised to make a quantum intellectual leap by facilitating collaboration that leverages recent advances in:
  - Theories about the social motivations for creating, maintaining, dissolving and re-creating network ties within teams
  - Developments in cyberinfrastructure and Web 2.0 that provide the technological capability to capture and analyze relational metadata needed to more effectively understand and enable teams.
  - Statistical techniques to make theoretically grounded team assembly recommendations that go beyond the Lovegety and SNIF
  - Petascale computational infrastructure to execute the statistical and optimization algorithms

Acknowledgements

National Science Foundation
WHERE DISCOVERIES BEGIN

NATIONAL CANCER INSTITUTE
National Cancer Institute
U.S. National Institutes of Health | www.cancer.gov

AIR FORCE RESEARCH LABORATORY

ARMY RESEARCH INSTITUTE

ARL

NORTHWESTERN UNIVERSITY 1851

MACARTHUR
The John D. and Catherine T. MacArthur Foundation

SONIC
Advancing the Science of
Networks in Communities

Research Team @ SONIC

Yun Huang
Post-doc

Annie Wang
Post-doc

Mengxiao Zhu
IEMS Doctoral candidate

David Huffaker
SoC Doctoral candidate

Brian Keegan
Doctoral Candidate

Jingling Li
Programmer

Jeffrey Treem
Doctoral candidate

York Yao
Programmer

Zack Johnson
SoC Undergraduate