Sic Transit Gloria Mundi Virtuali?
Promise and Peril in the Computational Social Science of Clandestine Organizing

Brian Keegan*, Muhammad Aurangzeb Ahmad†,
Dmitri Williams‡, Jaideep Srivastava†, Noshir Contractor*

* Northwestern U., † U. Minnesota, ‡ U. Southern California

June 17, 2011 – Web Science 2011   @bkeegan

NORTHWESTERN UNIVERSITY

SONIC
advancing the science of networks in communities

MMOGs

- Massively-multiplayer online games (MMOGs) operate similarly to "real world" societies

- MMOGs, unlike the real world, also have large databases storing exhaustive records of cross-sectional and behavioral data

- Use MMOGs to inform theory about offline social behavior **[Williams 2010]**

Sic transit gloria mundi **virtuali**

Thus passes the glory of the virtual world

Whither glory?

- Virtual worlds and MMOGs will…
  - "change the way people work and businesses compete" **[Reeves & Reed 2009]**
  - "[be] a prototype of the future of Western culture" **[Bainbridge 2010]**
  - "[be] the precursor to…a new social, political, and economic order" **[Castronova 2007]**
- Sociotechnical systems like MMOGs don't emancipate us from biological, psychological, social, and cultural forces
- Virtual worlds and MMOGS also reproduce offline social ills
  - Cheating, exploitation, inequality, etc.

Whither promise and peril?

- Computer-mediated environments create exhaustive digital trace logs of behavior and interactions → ***computational social science (CSS)***

- **Promise**: Collecting and analyzing data on types of behavior intractable or impossible to do offline

- **Peril**: Developing "dual use" technologies to better monitor, control, and repress users as many activities become increasingly mediated

Case: Gold farming in MMOGs

o **Gold farming** and **real money trade** involve the exchange of virtual in-game resources for "real world" money

o Laborers in China and S.E. Asia paid to perform repetitive practices ("farming") to accumulate virtual wealth ("gold")

o Western players purchase farmed gold to obtain more powerful items/abilities and open new game areas

o Market for real money trade exceeds $3 billion annually **[Lehdonvirta & Ernkvist 2011]**

Toward a CSS of Clandestine Orgs?

- Clandestine organizations as networks [Morselli 2007; Stohl & Stohl 2007]

  - Multidimensional relations – trust, exchange, communication, authority, affiliation, etc.

  - Networks are temporary, dynamic, emergent, adaptive, flexible

  - Networks structured by micro, meso, and macro processes – different processes lead to different structures [Monge & Contractor 2003]

  - Descriptive approaches do not address underlying processes of how networks emerge, stabilize, dissolve [Monge & Contractor 2003]

- Gold farmers might operate under similar motivations and constraints as other clandestine orgs like drug traffickers

  - Profit motive, distribution challenges, selection pressures

Research Question(s)

- Goal: To identify (or disrupt) clandestine network, one needs to understand (or attack) the **processes** which create, stabilize, reconstitute the network

- Do high-risk transactions in an online setting follow the same organizing logic as high-risk transactions in offline settings?

- Which forms of in-game trade activity are most similar to patterns observed in offline drug trafficking network?

Data

- EQ2 data provided by Sony Online Entertainment
  - Exchanges between players – currency & items
  - List of accounts banned by administrators
  - Huge data: sample from 1 representative week from 1 server

- CAVIAR drug trafficking network [Morselli et al. 2007, 2008]
  - Canadian drug trafficking sting operation
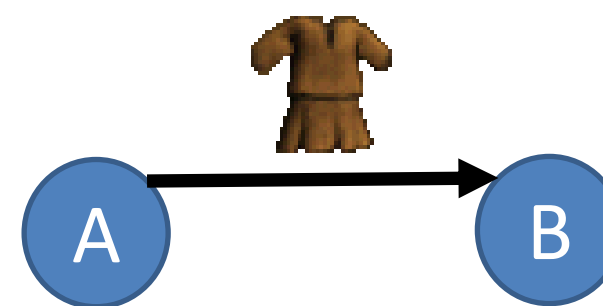 - Ties represent communication and exchange of materials
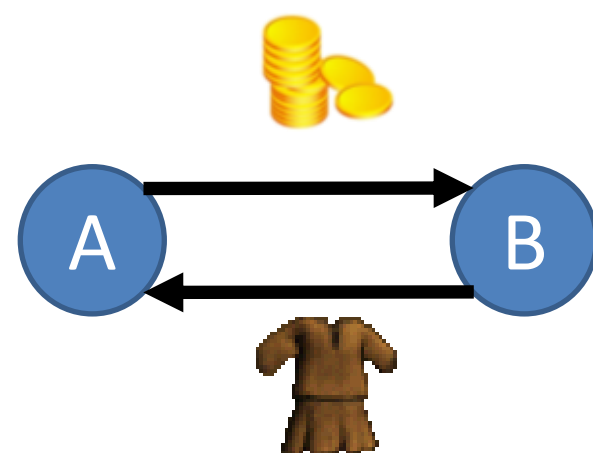
Multiplex trade relationships
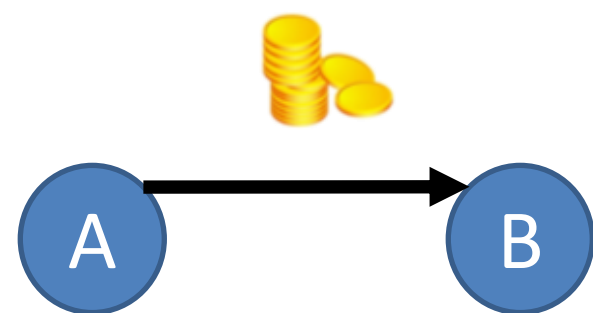
Donations
- Unreciprocated distribution
- Unusual & risky signal
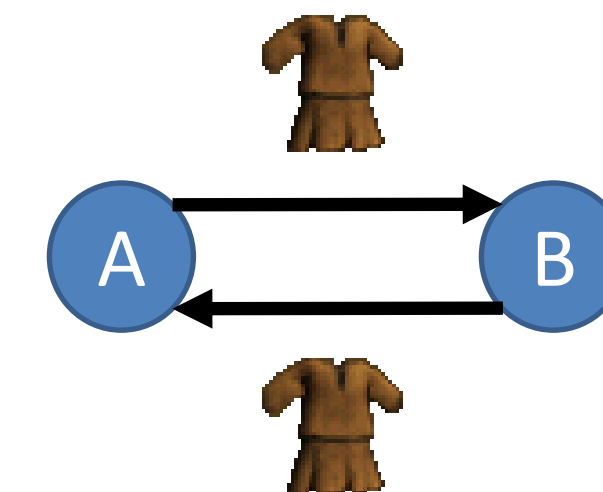- Nodes = 1519
- Edges = 1318

Gifting
- Unreciprocated distribution
- Unusual & risky signal
- Nodes= 5461
- Edges = 9239

Market exchange
- Goods for money
- Legitimate & common
- Nodes = 1022
- Edges = 768

Bartered exchange
- Goods for goods
- Legitimate & common
- Nodes = 1138
- Edges = 1323

Method - P*/ERGM modeling

- P*/exponential random graph models - "Regression" for networks

- Specify local endogenous tendencies & exogenous factors

- Networks of difference size and context structured by similar tendencies will have similar coefficients **[Faust & Skvoretz 2002]**

  - Senate co-sponsorship more similar to cow-licking than monkey grooming

Out-edge + In-edge + Distribution + Popularity + Brokering + Reciprocity + Cyclicality + Transitivity
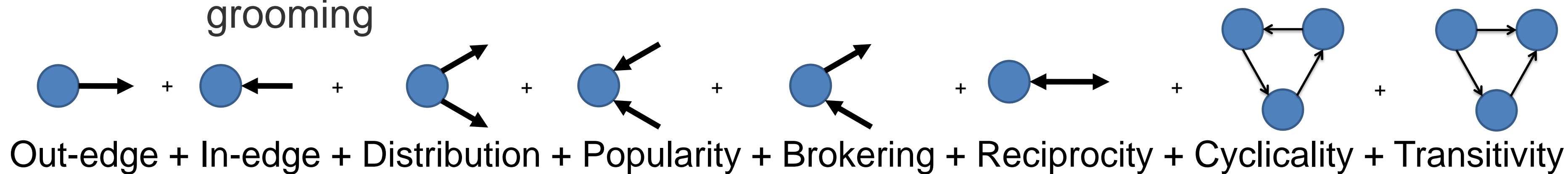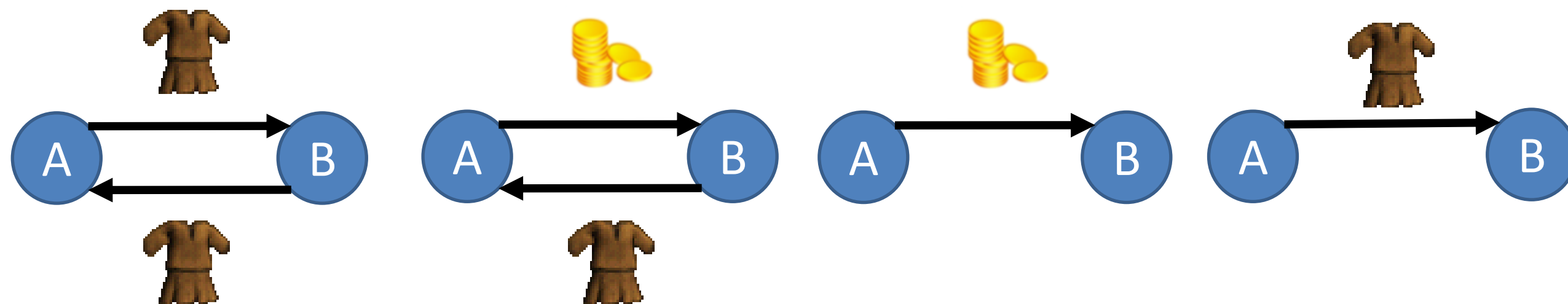
|	Barter			Market			Donations			Gifting			CAVIAR		
	Est.	SE	P	Est.	SE	P	Est.	SE	P	Est.	SE	P	Est.	SE	P
In-edge	-7.357	2.52E-01	***	-7.075	4.89E-02	***	-6.786	4.57E-02	***	-8.811	1.83E-02	***	-5.559	1.35E-01	***
Popularity	0.913	1.41E-01	***	0.126	4.58E-02	**	0.090	4.57E-02	*	0.135	3.77E-04	***	0.044	2.86E-02	
Out-edge	0.853	0.00E+00	***	-0.039	4.89E-02		-0.122	4.57E-02	**	-0.040	1.83E-02	*	-0.002	0.00E+00	***
Distribution	-0.324	1.72E-01	.	0.181	4.01E-02	***	-0.461	6.91E-02	***	0.074	1.30E-02	***	0.099	2.07E-05	***
Brokering	-2.159	2.88E-02	***	-0.314	4.76E-02	***	-0.480	4.46E-02	***	-0.035	2.12E-03	***	-0.012	4.40E-05	***
Reciprocity	12.681	1.09E-01	***	3.524	3.93E-01	***	6.408	6.32E-01	***	7.195	6.06E-02	***	4.163	1.31E-03	***
Cyclicality	2.699	1.07E+01		2.843	1.12E+00	*	0.147	8.56E-01		-0.642	2.07E-02	***	-0.533	6.98E-05	***
Transitivity	-0.094	9.43E-02		1.228	6.84E-01	.	3.004	2.71E-01	***	1.359	1.25E-02	***	0.434	4.77E-05	***

Results

- Gifting and drug trafficking structured by similar processes
  - Selective with partners
  - Avoid brokering
  - Strong reciprocity
  - Strong hierarchy (transitivity)
  - No generalized reciprocity (cyclicality)
- Cosine similarity: 0.997
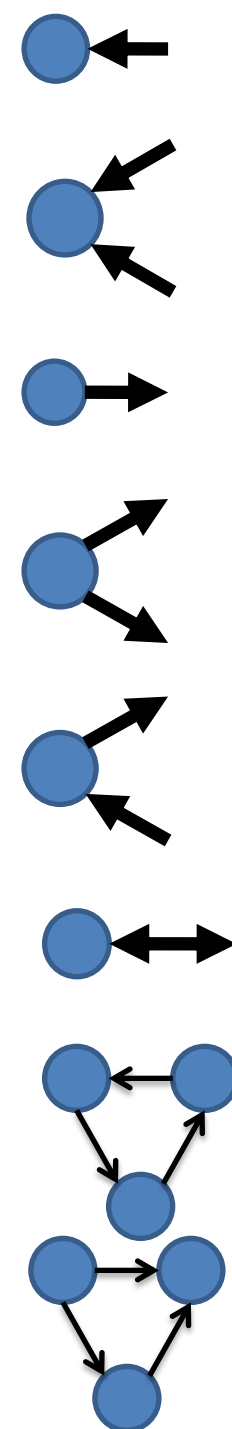
	Gifting			Caviar		
	Est.	SE	P	Est.	SE	P
In-edge	**-8.811**	1.83E-02	***	**-5.559**	1.35E-01	***
Popularity	**0.135**	3.77E-04	***	**0.044**	2.86E-02	
Out-edge	**-0.040**	1.83E-02	*	**-0.002**	0.00E+00	***
Distributor	**0.074**	1.30E-02	***	**0.099**	2.07E-05	***
Brokering	**-0.035**	2.12E-03	***	**-0.012**	4.40E-05	***
Reciprocity	**7.195**	6.06E-02	***	**4.163**	1.31E-03	***
Cyclicality	**-0.642**	2.07E-02	***	**-0.533**	6.98E-05	***
Transitivity	**1.359**	1.25E-02	***	**0.434**	4.77E-05	***

NORTHWESTERN UNIVERSITY

SONIC
advancing the science of networks in communities

Gold farming isn't drug trafficking

- Different constraints and affordance of online vs. offline world
  - Teleportation, instantaneous healing, fixed NPC vendor prices, can't kill the admins
- Online norms can be orthogonal to offline norms
  - Killing, stealing, and loitering tolerated or encouraged
- Consequences of detection have very different significance
  - Banned account online vs. violent retribution or incarceration offline
- In spite of differences, "structuring logic" of high risk transactions both online and offline are similar
- MMOG issues not provincial – foreshadow future debates

Dramaturgical stages and data science

- Are users fully rational about data being collected on their interactions?
    - Are you about your cell phone? Internet cookies?
- Or are users making assumptions about public, private, or anonymous interactions online based on offline analogues?
    - Do users (wrongly) assume private mediated behavior remains private or do users act deviantly in spite of surveillance?
- Real identity and cues substantially obscured online
    - Still possible to identify real world cues based on online behavior

Ethical quandaries

- Euphemisms abound: "removing" links and nodes
  - Should scholars be engaged in destructive science?

- Clandestine networks as "dual use" technologies
  - Used for good (Arab spring), ill (al-Qaeda), unclear (Wikileaks)

- Legal dimensions of information theory and methods
  - Different assumptions foreground different suspects [Ahmad, Keegan, et al. 2009]
  - Minimize false positives? Minimize false negatives? Maximize true positives? Maximize true negatives?
  - What standards for quality of data, stability of model outcomes, or complexity of models should govern "removal" decisions?

Legal questions

- Mapping legal and regulatory frameworks
  - Connecting behavioral trace data to individual action
  - Internal and external validity of data
  - Abstraction of behavior into complex models & data structures
- False positives and due process
  - Disclose proprietary methodological approaches during discovery?
  - Heightened burden of proof given superabundance of data?
  - Jurisdiction, peers, right to representation, due process…

NORTHWESTERN
UNIVERSITY

SONIC
advancing the science of networks in communities

- bkeegan@northwestern.edu
- Muhammad Ahmad (UMN), Noshir Contractor (NU), Jaideep Srivastava (UMN), Dmitri Williams (USC), Sophia Sullivan (NU)

http://www.vwobservatory.org/
http://sonic.northwestern.edu/

FA8650-10-C-7010   IIS-0729505   W91WAW-08-C-0106

NORTHWESTERN UNIVERSITY

SONIC

advancing the science of networks in communities