Applying Tracking Theory to the Network Domain

Dr. James P. Ferry
Metron, Inc.

Graph Exploitation Symposium

Lincoln Labs

April 13-14, 2010

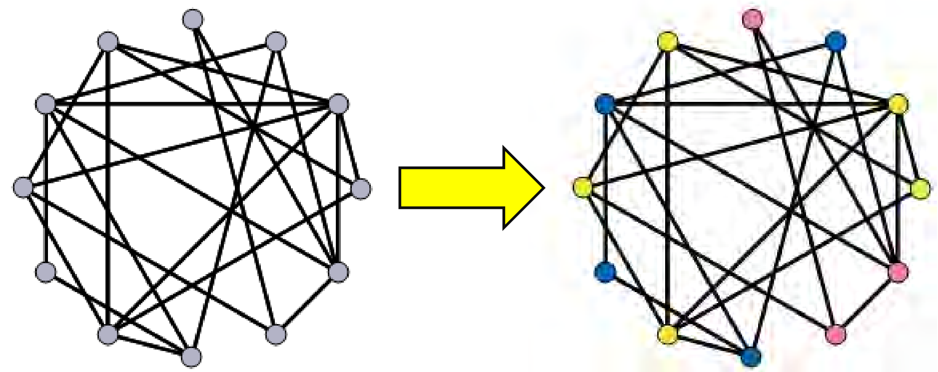
Funding Source: Phase II SBIR, N08-081
Contract Number: N00014-09-C-0563
Sponsoring Agency: ONR, Code 311
Program Manager: Dr. Donald Wagner

Overview

- Definitions
  - Group finding (on networks)
  - Group tracking
- Toward a “Kalman filter for networks”
  - The static homophily model
  - The dynamic homophily model
  - The inference equations
  - Example
- Marginalization of vast state space
  - Pairwise probability equations
  - Closures
  - Group partitions from pairwise probabilities
- Summary

Group Finding Algorithms

- Problem: given a network, find the "best" partition of its nodes into groups
  - Nodes in same group tend to be more linked
- Substantial literature [1]
- Seminal work (2002): Newman's "modularity" metric [2]
- State of the art: three excellent algorithms developed in 2008-2009 (Infomap, etc.) [3]
- What does "best" mean?



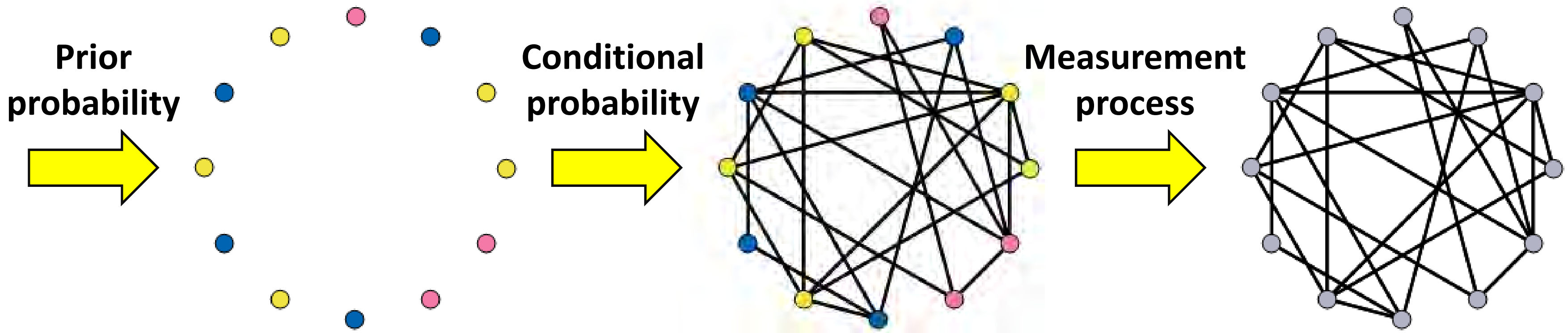
[1] S. Fortunato, "Community detection in graphs," Phys. Rep., **486**(3-5), pp. 75-174 (2010).

[2] M. Girvan and M. Newman, "Community structure in social and biological networks," Proc. Natl. Acad. Sci. USA **99**(12), pp. 7821-6 (2002).

[3] A. Lancichinetti and S. Fortunato, "Community detection algorithms: A comparative analysis," Phys. Rev. E, **80**(5) , 056117 (2009).

Group Finding as an Inference Problem

- Posit causal model for how groups and networks arise [4]
- Bayesian inference finds group distribution given network data

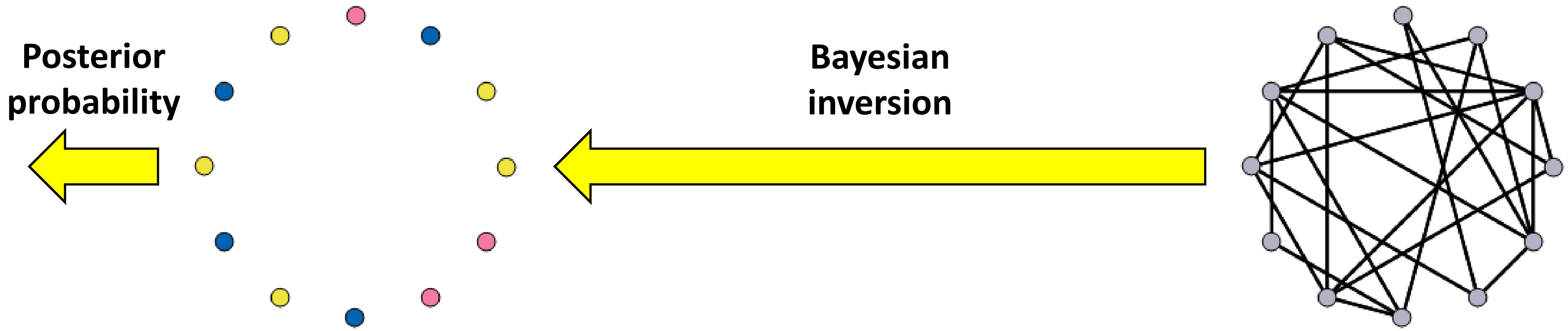


- Trade-off of modeling: meaningfulness vs. oversimplicity
- Much mileage from simple models (later extended)
  - Erdős–Rényi model [5] → "6 degrees of separation" & "tipping point"
  - Kalman filter [6] based on variants of Brownian motion model

[4] M. Hastings, "Community detection as an inference problem," Phys. Rev. E, **74**(3) , 035102(R) (2006).
[5] P. Erdős and A. Rényi, "On the Evolution of Random Graphs," Pub. Math. Inst. Hung. Acad. Sci. **5** (1960).
[6] R. Kalman, "A new approach to linear filtering and prediction problems," J. Basic Eng. **82**(1), 35-45 (1960).

Group Finding as an Inference Problem

- Posit causal model for how groups and networks arise [4]
- Bayesian inference finds group distribution given network data

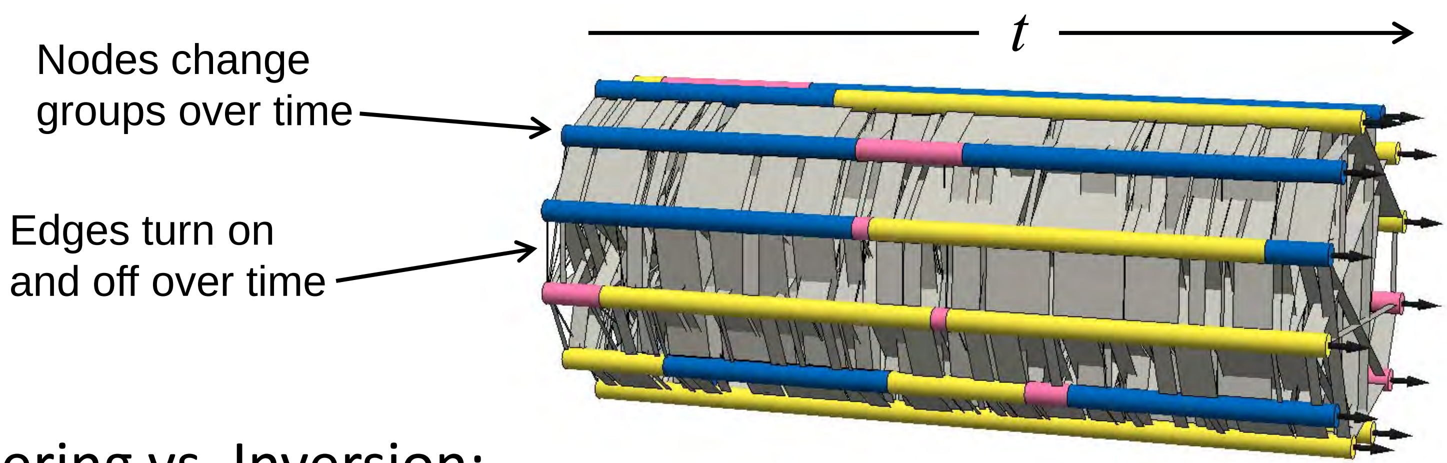


- Trade-off of modeling: meaningfulness vs. oversimplicity
- Much mileage from simple models (later extended)
  - Erdős–Rényi model [5] → "6 degrees of separation" & "tipping point"
  - Kalman filter [6] based on variants of Brownian motion model

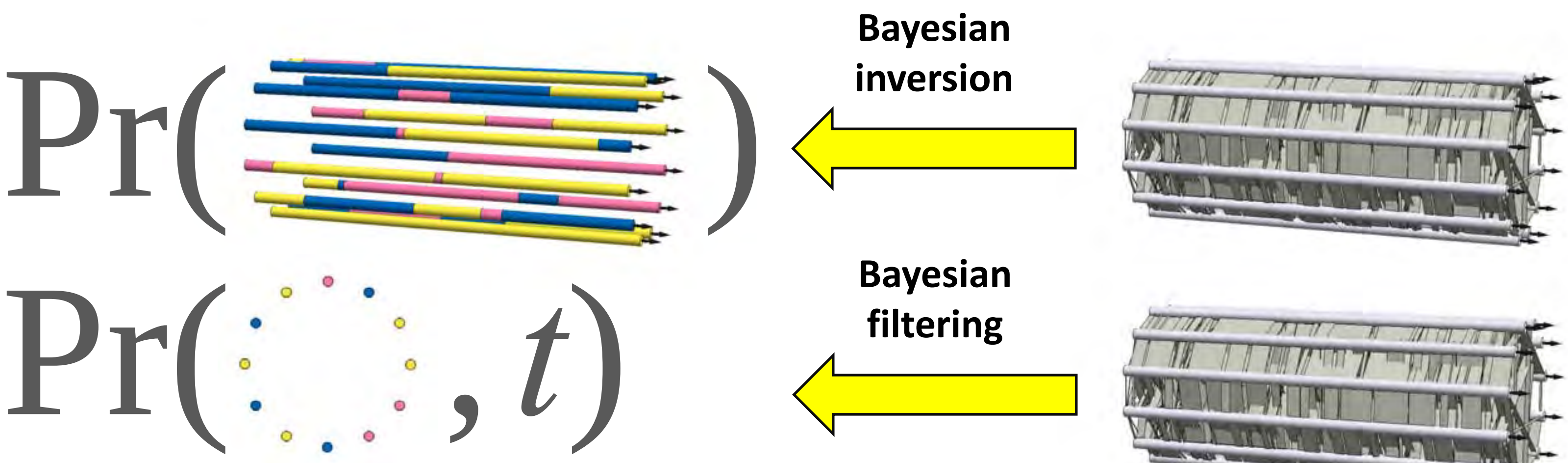
[4] M. Hastings, "Community detection as an inference problem," Phys. Rev. E, **74**(3) , 035102(R) (2006).
[5] P. Erdős and A. Rényi, "On the Evolution of Random Graphs," Pub. Math. Inst. Hung. Acad. Sci. **5** (1960).
[6] R. Kalman, "A new approach to linear filtering and prediction problems," J. Basic Eng. **82**(1), 35-45 (1960).

Group Tracking

- Extension of group finding to include time dimension:



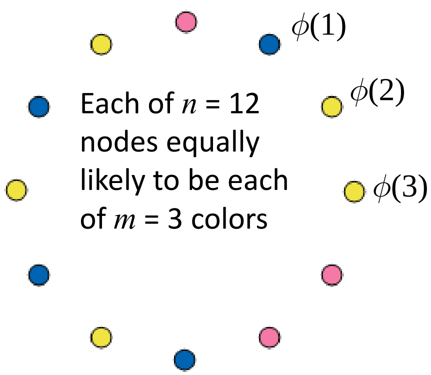
- Filtering vs. Inversion:



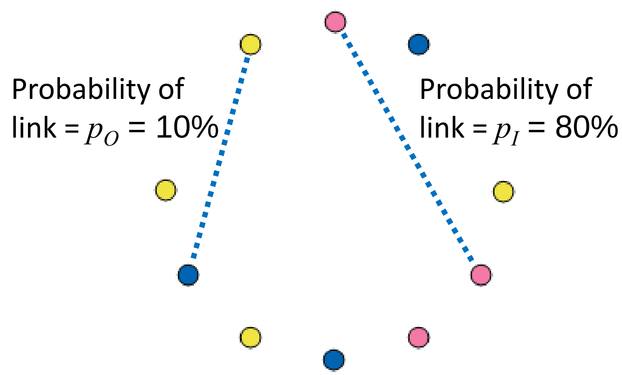
Static Homophily Model

- Static homophily model: $(\phi, G) \sim \mathrm{H}(n, m, p_I, p_O)$
  - n = number of nodes
  - m = number of groups
  - p_I = probability of intragroup link
  - p_O = probability of intergroup link

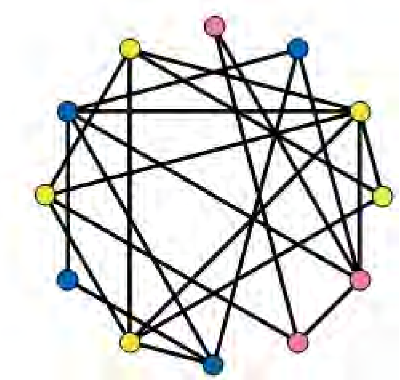
Model for generating group assignment ϕ



Given ϕ, model for generating graph G



An instance (ϕ, G) of $\mathrm{H}(12,3,0.8,0.1)$



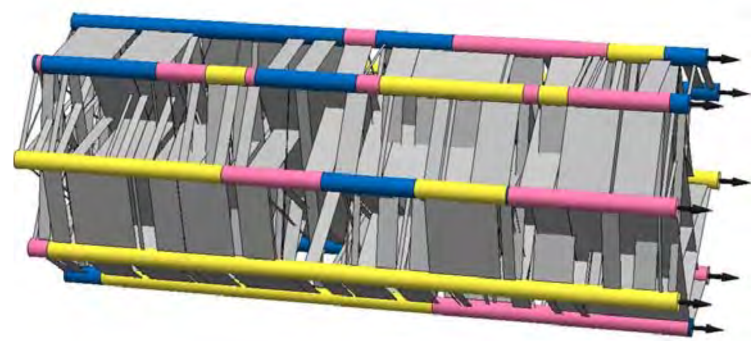
- What is probability of group assignment ϕ given a graph G?
  - Model enables us to ask question meaningfully
  - Answer straightforward

METRON SCIENTIFIC SOLUTIONS

Dynamic Homophily Model

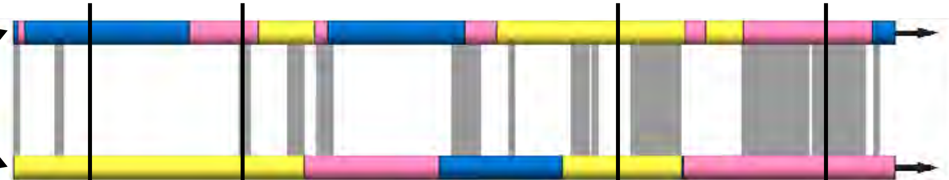
- Dynamic case: $(\phi_t, G_t) \sim \mathcal{H}(n, m, a, \lambda_I, \mu_I, \lambda_O, \mu_O)$
  - n = number of nodes
  - m = number of groups
  - a = rate that nodes change group
  - λ_I = turn-on rate of intragroup link
  - μ_I = turn-off rate of intragroup link
  - λ_O = turn-on rate of intergroup link
  - μ_O = turn-off rate of intergroup link

An instance (ϕ_t, G_t) of $\mathcal{H}(8,3,1.5,16,4,2,18)$



time: 0 to 5

nodes' rates of change: $a = 1.5$



links' rates of change:

	$\lambda_O = 2$ (*links turn on slowly*)	$\mu_O = 18$ (*links turn off quickly*)		$\lambda_I = 16$ (*links turn on quickly*)	$\mu_I = 4$ (*links turn off slowly*)	
between groups						*within groups*

Mathematical Treatment

- Let $\mathcal{R}$ be the vector of probabilities of all possible (ϕ, G)
- $\mathcal{R}$ governed by a Markov process with transition rate matrix $\mathcal{C}$

$$\dot{\mathcal{R}} = \mathcal{C}\mathcal{R}$$

- $\mathcal{C}$ huge, $m^n 2^{n(n-1)/2}$ $m^n 2^{n(n-1)/2}$ entries, but sparse
- Explicit formula for $\mathcal{C}$:

$$\mathcal{C}_{(\phi,G^+)(\phi',G)} = \delta_{G^+G}\alpha_{G,\phi\phi'} + \delta_{\phi\phi'}\beta_{\phi,G^+G}$$

$$\alpha_{G,\phi\phi'} = \begin{cases} -na - \sum_{\{v,w\}\subseteq V} \gamma^{vw}(\phi,G), & \text{if } \phi = \phi', \\ a/(m-1), & \text{if } \phi = \phi' \text{ except at 1 node, and} \\ 0, & \text{otherwise.} \end{cases}$$

$$\beta_{\phi,G^+G} = \begin{cases} \gamma^{vw}(\phi,G), & \text{if } G^+ = G \text{ except at } \{v,w\}, \text{ and} \\ 0, & \text{otherwise.} \end{cases}$$

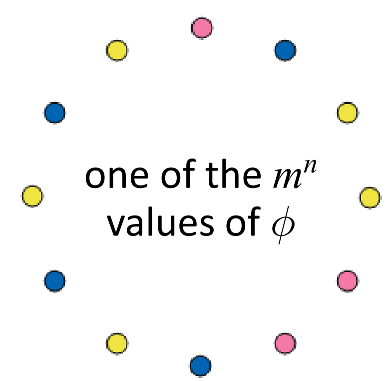
$\gamma^{vw}(\phi,G) =$ pertinent link rate $(\lambda_I, \mu_I, \lambda_O,$ or $\mu_O)$

METRON SCIENTIFIC SOLUTIONS

ONR Revolutionary Research . . . Relevant Results

Bayesian Filter

- Explicit evolution law for $\mathcal{R}$ (joint probability of ϕ_t and G_t)…
- …converted into law for $\Pr(\phi_t)$ given network data $\{G_s : s < t\}$
- Form of law:
  - Let $\tilde{p}^{\phi}$ denote the <u>unnormalized</u> probability of ϕ (at some time t)
  - During time intervals in which G remains constant: $\dot{\tilde{p}}^{\phi} = \sum_{\phi'} \alpha_{G,\phi\phi'} \tilde{p}^{\phi'}$
  - When graph jumps from G to G^{+}: $\tilde{p}^{\phi +} = \beta_{\phi,G^{+}G} \tilde{p}^{\phi}$
    - (Assuming jump is at a single edge. Special handling for multi-edge jumps)
- Sample visualization (next slide):
  - Dynamic network data
    - (Different color links: generalized version)
  - Full probability vector p^{ϕ} maintained (m^n values)
    - Too big to visualize fully
    - Visualization shows only "first-order statistics": probability of each node being in each group



Marginalize, Symmetrize, Close

- Inference equations require maintain a huge state space
  - E.g., n = 12 nodes and m = 3 groups needs 3^{12} = 530,000 states!
  - Completely impractical
- But exact equation is good start for an accurate approximation
- Step 1: Marginalize
  - E.g., sum equations over the groups of all nodes but 3 (v, w, and x)
  - Yields equations for the m^3 probabilities p_{ijk}^{vwx}: node v in group i, etc.
- Step 2: Symmetrize
  - Reduce assignments ϕ = (3,1,2,1,1,3) to partitions π = {{1,6},{2,4,5},{3}}
  - The m^3 values p_{ijk}^{vwx} reduce to 5: $p^{\{v,w,x\}}, p^{\{v,w\}\{x\}}, p^{\{v,x\}\{w\}}, p^{\{w,x\}\{v\}}, p^{\{v\}\{w\}\{x\}}$
- Step 3: Close
  - Equations for $p^{\{v,w,x\}}$ involve 4th- and 5th-order statistics
  - Need models for such higher-order statistics (results now approximate)

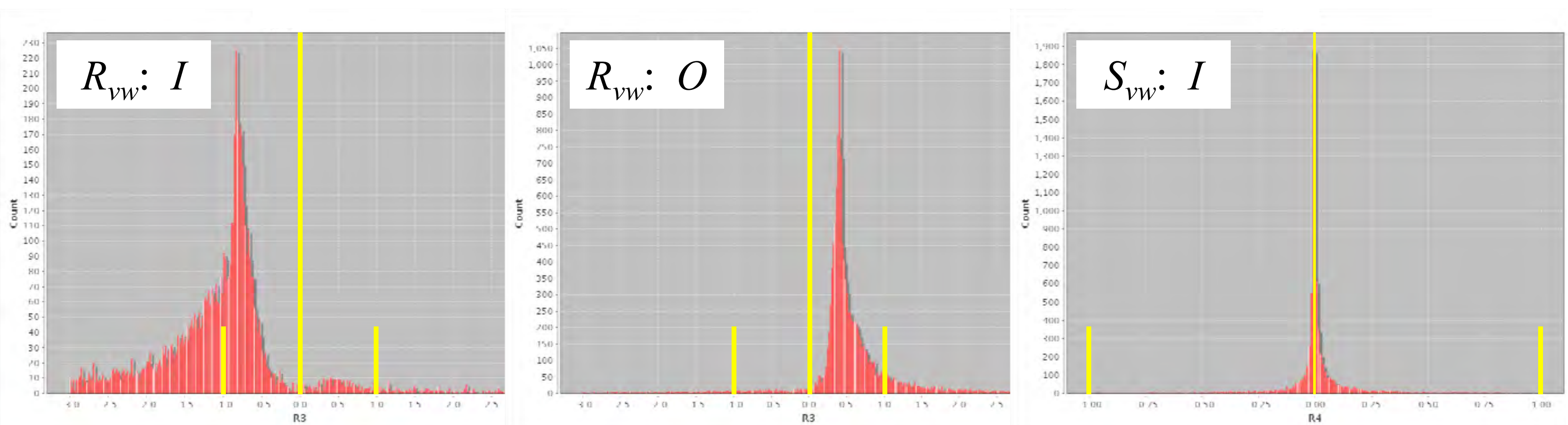
METRON SCIENTIFIC SOLUTIONS

Second-order Equations

- First-order results: $p^{\{v\}} = 1$ (not very helpful)
- Second-order results:

$$\dot{p}^{\{v,w\}} = \frac{2am}{m-1}\left(\frac{1}{m} - p^{\{v,w\}}\right) + \left(\gamma_O^{vw} - \gamma_I^{vw}\right) p^{\{v,w\}}\left(1 - p^{\{v,w\}}\right) - \overbrace{\sum_{x \neq v,w} r_{vwx}}^{R_{vw}} - \overbrace{\sum_{x,y \neq v,w} s_{vwxy}}^{S_{vw}}$$

  - r_{vwx} a function of the 5 third-order statistics ($p^{\{v,w,x\}}, p^{\{v,w\}\{x\}}$, …)
  - s_{vwxy} a function of the 15 fourth-order statistics
- Sample histograms:
  - Safe to truncate fourth-order terms?



Maximum Entropy Closure

- Need model for third-order statistics
- Use values of p_{ijk}^{vwx} which maximize entropy $H = -\sum_{i,j,k} p_{ijk}^{vwx} \log p_{ijk}^{vwx}$
- Subject to constraints
  - Five third-order statistics…
  - …satisfy the four equations →
  - leaving 1 free variable $p^{\{v,w,x\}}$

$p^{\{v,w,x\}} + p^{\{v,w\}\{x\}} = p^{\{v,w\}}$,

$p^{\{v,w,x\}} + p^{\{v,x\}\{w\}} = p^{\{v,x\}}$,

$p^{\{v,w,x\}} + p^{\{w,x\}\{v\}} = p^{\{w,x\}}$, and

$p^{\{v,w,x\}} + p^{\{v,w\}\{x\}} + p^{\{v,x\}\{w\}} + p^{\{w,x\}\{v\}} + p^{\{v\}\{w\}\{x\}} = 1.$

- Solution: unique root of

$$\frac{dH}{dp^{\{v,w,x\}}} = \log\left(\frac{(m-2)^2}{4(m-1)} \frac{\left(p^{\{v,w\}} - p^{\{v,w,x\}}\right)\left(p^{\{v,x\}} - p^{\{v,w,x\}}\right)\left(p^{\{w,x\}} - p^{\{v,w,x\}}\right)}{p^{\{v,w,x\}}\left(p^{\{v,w,x\}} - p^{-}\right)^2}\right) = 0$$

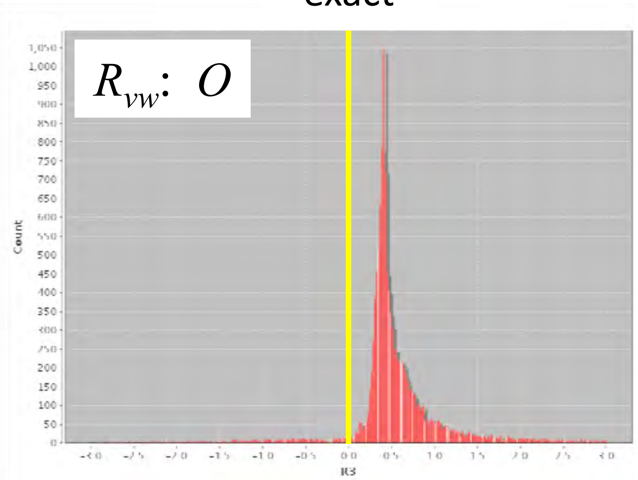
in the range $\left(p^{\{w,x\}} + p^{\{v,x\}} + p^{\{v,w\}} - 1\right)/2 = p^{-} \le p^{\{v,w,x\}} \le p^{+} = \min\left(p^{\{v,w\}}, p^{\{v,x\}}, p^{\{w,x\}}\right)$

- Requires solution of a cubic polynomial

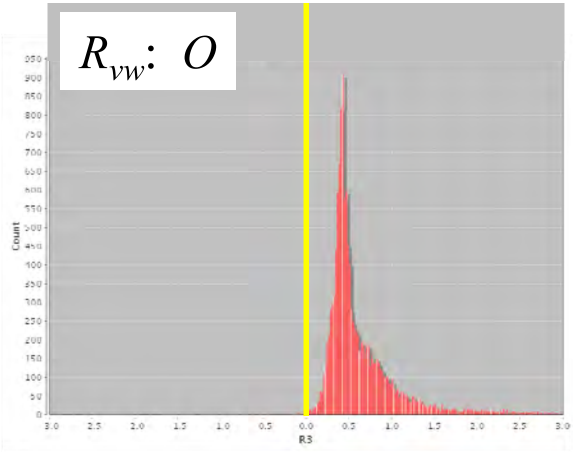
Maximum Entropy Closure Results

- Instantaneous agreement is quite good

exact



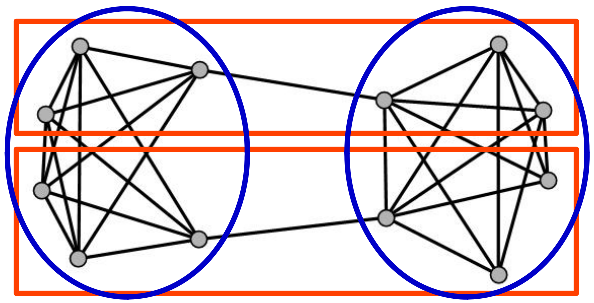
maximum entropy closure



- However, system evolved with this closure blows up
  - Investigation under way
  - Subtle consistency requirements for second-order statistics
  - Fourth-order terms play critical role in maintaining this consistency

What about Finding Groups?

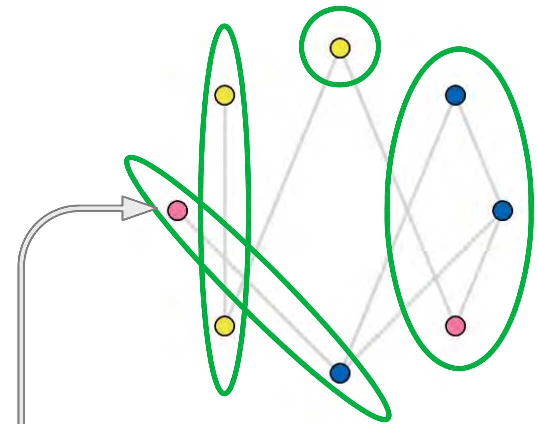
- What makes one group partition better than another?



- Usual answer: some network statistic
  - E.g., Newman's modularity
  - Sounds reasonable
  - Difficult to validate or generalize without extensive experimentation
- Bayesian answer: expected utility
  - Network statistic = f (computed partition, network)
  - (Bayesian) Utility = f (computed partition, *ground truth partition*)
  - Expected utility requires knowledge of $\Pr$(ground truth | network)
    - Which is what is being computed, in principle…
  - Do the second-order statistics suffice?

Utility Function

- Utility function reflects why one cares whether a group partition is accurate
- Proposed generic utility function
  - Based on a simple but representative scenario of how group partitions could be used
  - Tally effects of each vertex v being in “bad” group
    - “Kill” vertices in the same computed group as v
    - For each v tally bad, good, dead, and alive
- Example
  - Tally yields 12 cases of Bad & Dead, etc.
  - Let $U = 12\ u_{BD} + 10\ u_{BA} + 6\ u_{GD} + 36\ u_{GA}$
    - Positive utility: dead bad guys, alive good guys
    - Negative utility: alive bad guys, dead good guys
  - Different coefficients…
    - Reflect different priorities
    - Yield different group partitions



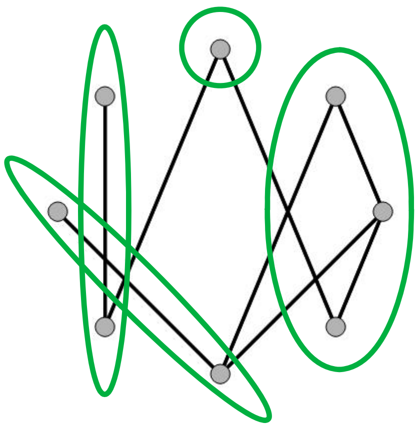
	● D	● D	● D	● A	● A	● A
● B	***1***	**1**	**0**	*1*	2	3
● B	***1***	**2**	**0**	*1*	1	3
● B	**1**	***2***	**0**	1	*1*	3
● B	**1**	***2***	**0**	1	*1*	3
● B	**1**	***1***	**0**	1	*2*	3
● B	**0**	**0**	***1***	2	3	*2*
● B	**0**	**0**	***2***	2	3	*1*
● B	**0**	**0**	***2***	2	3	*1*

Totals	
12	Bad & Dead
10	Bad & Alive
6	Good & Dead
36	Good & Alive

METRON SCIENTIFIC SOLUTIONS

Expected Utility

- Utility defines how good an answer is when ground truth is known
  - Flexible: allows one to specify what one wants out of a solution
  - Not immediately useful: ground truth unknown



- Thus compute *expected utility* given the evidence:

$$\mathbb{E}[U](\hat{\pi}|G) \doteq \sum_{\pi} U(\hat{\pi}|\pi)\Pr(\pi|G)$$

- Result:

$$\mathbb{E}[U](\hat{\pi}|G) = \sum_{\hat{C}\in\hat{\pi}} \sum_{\{v,w\}\subseteq\hat{C}} \left(p^{\{v,w\}} - \theta\right), \text{ where } \theta = \frac{u_{GA} - u_{GD}}{(u_{GA} - u_{GD}) + (u_{BD} - u_{BA})}$$

  - Expected utility = average amount by which pairwise probabilities within (computed) groups exceeds a threshold θ
  - Yes, pairwise probabilities suffice

Summary

- Introduction of dynamic homophily model
- Exact inference results: Bayesian filter equations
- Exact equations for pairwise probabilities
  - Involves third- and fourth-order statistics: needs closure
  - Of intrinsic interest + used to create maximal expected utility partitions
- Possible closure via maximum entropy
  - Reveals delicate consistency requirements for pairwise probabilities
- Status:
  - Investigating pairwise consistency requirements
    - Last hurdle in creating a Kalman filter for networks
  - Developing IGNITE (Inter-Group Network Inference and Tracking Engine)
- References:

[6] J. Ferry, "Group Tracking on Dynamic Networks," Proceedings of the 12th International Conference on Information Fusion, Seattle, WA, July 6-9, 2009.

[7] J.P. Ferry and J.O. Bumgarner, "Tracking Group Co-membership on Networks," *submitted to* the 13th International Conference on Information Fusion, Edinburgh, Scotland, July 26-29, 2010.