Network as a computer: Ranking paths to find flows

Dusko Pavlovic

Oxford University and Kestrel Institute

CSR, Moscow
June 2008

Outline

Outline

Introduction

Definition

A network is an annotated graph

$$A \quad = \quad \left(R \xleftarrow{\varphi} E \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$$

where

- N is a finite set of *nodes*,
- E is a finite set of *edges* (or links),
- R is an ordered rig of *rates* (e.g. $\mathbb{R}_+$).

Introduction

Definition

A network is an annotated graph

$$A \quad = \quad \left(R \xleftarrow{\varphi} E \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$$

where

- N is a finite set of *nodes*,
- E is a finite set of *edges* (or links),
- R is an ordered rig of *rates* (e.g. $\mathbb{R}_+$).

Notation

- $i \overset{e}{\underset{v}{\rightarrow}} j$ denotes $e \in E$ such that

$$\delta(e) = i, \;\; \varrho(e) = j \;\text{ and }\; \varphi(e) = v$$

Introduction

Networks are used to model

- social groups
- traffic, distribution systems,
- Web, Internet
- protein interactions, gene regulation, metabolism,
- food webs, populations,
- neural nets,
- probabilistic grammars (generative, phonological)

Problem and objective

Problem

Networks get large and complex.

Problem and objective

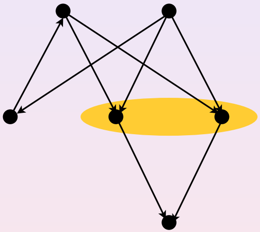
Problem

Networks get large and complex.

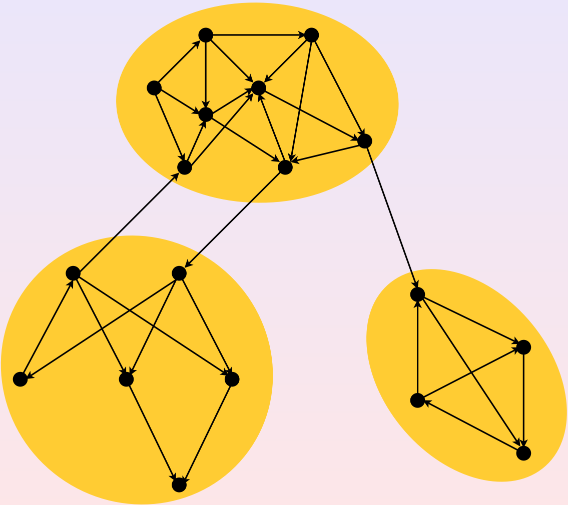
Objective

Simplify them. . .

. . . by clustering similar nodes



. . . by extracting functional modules



Problem and objective

Problem of the Web

Data structures and semantics vary from node to node.

Problem and objective

Problem of the Web

Data structures and semantics vary from node to node.

Solutions

- the Semantic Web
- search, latent semantics
  - extract structure from network
  - **concepts = communities = modules**

Approach

Notation

Given a network $A = \left(R \xleftarrow{\varphi} E \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$, define

- *total flow* $A_{ij} = \sum_{i \xrightarrow{e} j} \varphi(e)$

Approach

Notation

Given a network $A = \left(R \stackrel{\varphi}{\leftarrow} E \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$, define

- *total flow* $A_{ij} = \sum_{i \stackrel{e}{\to} j} \varphi(e)$
- *flow distribution* $\Phi_{ij} = \frac{A_{ij}}{A_{\bullet\bullet}}$,
  where $A_{\bullet\bullet} = \sum_{ij} A_{ij}$

Approach

Notation

Given a network $A = \left(R \stackrel{\varphi}{\leftarrow} E \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$, define

- *total flow* $A_{ij} = \sum_{i \xrightarrow{e} j} \varphi(e)$
- *flow distribution* $\Phi_{ij} = \frac{A_{ij}}{A_{\bullet\bullet}}$,
  where $A_{\bullet\bullet} = \sum_{ij} A_{ij}$
- *flow bias* $\Upsilon_{ij} = \Phi_{ij} - \Phi_{i\bullet}\Phi_{\bullet j}$
  where $\Phi_{i\bullet} = \sum_k \Phi_{ik}$ and $\Phi_{\bullet j} = \sum_k \Phi_{kj}$.

Cohesion and adhesion

Definition

Cohesion of $U \subseteq N$ is the total flow bias between its members

$$\mathrm{Coh}(U) \;=\; \sum_{i,j \in U} \Upsilon_{ij}$$

Cohesion and adhesion

Definition

Cohesion of $U \subseteq N$ is the total flow bias between its members

$$\mathrm{Coh}(U) = \sum_{i,j \in U} \Upsilon_{ij}$$

Adhesion of $U \subseteq N$ is the total flow bias of its members and nonmembers

$$\mathrm{Adh}(U) = \sum_{i \in U, j \notin U} \Upsilon_{ij} + \Upsilon_{ji}$$

Modularity

Definition

Modularity of $U \subseteq N$ is the difference of its cohesion and its adhesion

$$\mathrm{Mdu}(U) = \mathrm{Coh}(U) - \mathrm{Adh}(U)$$

Modularity

Definition

Modularity of $U \subseteq N$ is the difference of its cohesion and its adhesion

$$\mathrm{Mdu}(U) \quad = \quad \mathrm{Coh}(U) - \mathrm{Adh}(U)$$

Idea

Find modules by maximizing modularity.

Task

Finding network modules boils down to evaluating

- modularity $\mathrm{Mdu}(U) = \mathrm{Coh}(U) - \mathrm{Adh}(U)$

Task

Finding network modules boils down to evaluating

- modularity $\text{Mdu}(U) = \text{Coh}(U) - \text{Adh}(U)$

  which is induced by
- flow bias $\Upsilon : N \times N \longrightarrow [-1, 1]$

Task

Finding network modules boils down to evaluating

- modularity $\mathrm{Mdu}(U) = \mathrm{Coh}(U) - \mathrm{Adh}(U)$

  which is induced by
- flow bias $\Upsilon : N \times N \longrightarrow [-1, 1]$

  which is induced by
- flow distribution $\Phi : N \times N \longrightarrow [0, 1]$,

Task

Finding network modules boils down to evaluating

- modularity $\text{Mdu}(U) = \text{Coh}(U) - \text{Adh}(U)$

  which is induced by
- flow bias $\Upsilon : N \times N \longrightarrow [-1, 1]$

  which is induced by
- flow distribution $\Phi : N \times N \longrightarrow [0, 1]$,

  which is induced by
- flow $\varphi : E \longrightarrow R$

Problems

- flows unknown
- network dynamics unknown
- modules disjoint and hard to compute

Problems

- flows unknown

  Step 1: use cost and paths to estimate flows

- network dynamics unknown

  Step 2: use Markovian and ranking methods

- modules disjoint and hard to compute

  Step 3: parametrize modularity

Outline

Cost instead of flow

Modified definition

A network is a labelled graph

$$A \quad = \quad \left(R \xleftarrow{\gamma} E \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$$

where the **cost** γ determines the likely flow φ

$$\varphi(e) \quad = \quad 2^{-\gamma(e)}$$

Cost instead of flow

Modified definition

A network is a labelled graph

$$A = \left(R \xleftarrow{\gamma} E \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$$

where the **cost** γ determines the likely flow φ

$$\varphi(e) = 2^{-\gamma(e)}$$

The estimated total flow $i \to j$ in A is now thus

$$A_{ij} = \sum_{i \xrightarrow{e} j} 2^{-\gamma(e)}$$

Path completion

Definition

Given

- network $A = \left(R \xleftarrow{\gamma} E \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$,
- cutoff value $v \in R$, and
- length penalty $d \in R$,

Path completion

Definition

Given

- network $A = \left(R \xleftarrow{\gamma} E \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$,
- cutoff value $v \in R$, and
- length penalty $d \in R$,

we define the *path completion* of A as

- network $A^{*vd} = \left(R \xleftarrow{\gamma} E^{*vd} \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$

Path completion

Definition

Given

- network $A = \left(R \overset{\gamma}{\leftarrow} E \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$,
- cutoff value $v \in R$, and
- length penalty $d \in R$,

we define the *path completion* of A as

- network $A^{*vd} = \left(R \overset{\gamma}{\leftarrow} E^{*vd} \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$ with
- links $E^{*vd} = \{a \in E^{+} \mid \gamma(a) \leq v\}$

Path completion

Definition

Given

- network $A = \left(R \stackrel{\gamma}{\leftarrow} E \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$,
- cutoff value $v \in R$, and
- length penalty $d \in R$,

we define the *path completion* of A as

- network $A^{*vd} = \left(R \stackrel{\gamma}{\leftarrow} E^{*vd} \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$ with
- links $E^{*vd} = \{a \in E^{+} \mid \gamma(a) \le v\}$
  - E^{+} is the set of nonempty paths

Path completion

Definition

Given

- network $A = \left(R \xleftarrow{\gamma} E \overset{\delta}{\underset{\varrho}{\rightrightarrows}} N\right)$,
- cutoff value $v \in R$, and
- length penalty $d \in R$,

we define the *path completion* of A as

- network $A^{*vd} = \left(R \xleftarrow{\gamma} E^{*vd} \overset{\delta}{\underset{\varrho}{\rightrightarrows}} N\right)$ with
- links $E^{*vd} = \{a \in E^+ \mid \gamma(a) \leq v\}$
  - E^+ is the set of nonempty paths
- cost $\gamma\left(i_0 \xrightarrow{a_1} i_1 \xrightarrow{a_2} \cdots \xrightarrow{a_n} i_n\right) = (n-1)d + \gamma(a_1) + \cdots + \gamma(a_n)$.

Outline

Simple network dynamics

forward: probability that traffic at i flows to j

$$A^{\triangleright}_{ij} = \frac{A_{ij}}{A_{i\bullet}} \text{ where}$$

$$A_{i\bullet} = \sum_{k=1}^{N} A_{ik}$$

backward: probability that traffic at j flows from i

$$A^{\triangleleft}_{ij} = \frac{A_{ij}}{A_{\bullet j}} \text{ where}$$

$$A_{\bullet j} = \sum_{k=1}^{N} A_{kj}$$

Simple node ranking

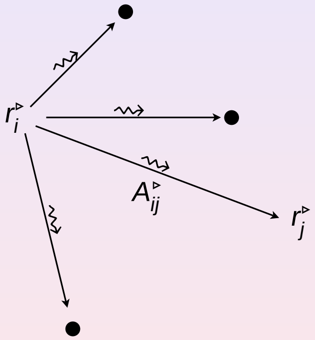
pull rank (reputation): probability that the traffic arrives at j

$$r^{\triangleright}_j = \sum_{k=1}^{N} r^{\triangleright}_k A^{\triangleright}_{kj}$$

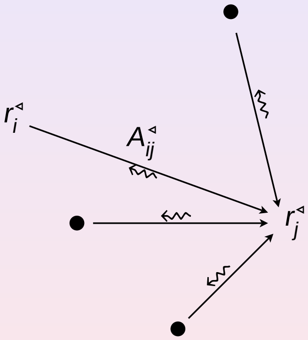
push rank (promotion): probability that the traffic departs from i

$$r^{\triangleleft}_i = \sum_{k=1}^{N} A^{\triangleleft}_{ik} r^{\triangleleft}_k$$

Simple pull rank: Reputation



Simple push rank: Promotion



One-hop dynamics

forward out: probability that traffic at i flows to a hub j

$$A^{\blacktriangleright}_{ij} = A^{\triangleright}_{ij} \cdot \Phi_{j\bullet} \text{ where}$$

$$\Phi_{j\bullet} = \frac{\sum_k A_{jk}}{\sum_{\ell k} A_{\ell k}}$$

backward in: probability that traffic at j flows from an authority i

$$A^{\blacktriangleleft}_{ij} = \Phi_{\bullet i} \cdot A^{\triangleleft}_{ij} \text{ where}$$

$$\Phi_{\bullet i} = \frac{\sum_k A_{ki}}{\sum_{k\ell} A_{k\ell}}$$

One-hop ranking

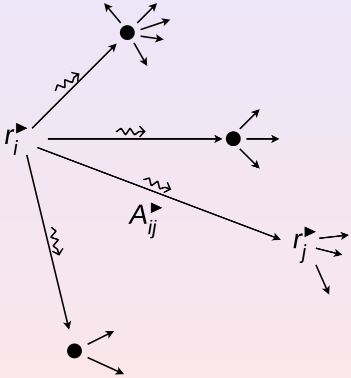
pull-out rank: probability that traffic arrives to a hub j

$$r^{\blacktriangleright}_j = \sum_{k=1}^{N} r^{\blacktriangleright}_k A^{\blacktriangleright}_{kj}$$

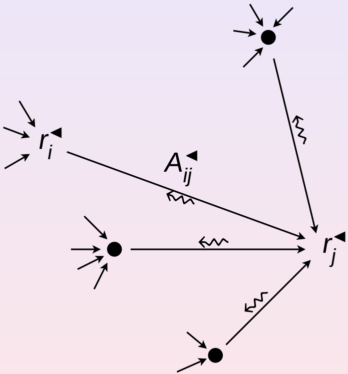
push-in rank: probability that traffic departs from an authority i

$$r^{\blacktriangleleft}_i = \sum_{k=1}^{N} A^{\blacktriangleleft}_{ik} r^{\blacktriangleleft}_k$$

Pull-out rank



Push-in rank



$r_i^\blacktriangleleft$
$A_{ij}^\blacktriangleleft$
$r_j^\blacktriangleleft$

Unibased flow

If the traffic from j to k is only driven by

- j's push $r_j^{\blacktriangleright}$, and by
- k's pull $r_k^{\blacktriangleleft}$,

which are assumed to be mutually independent

Unibased flow

If the traffic from j to k is only driven by

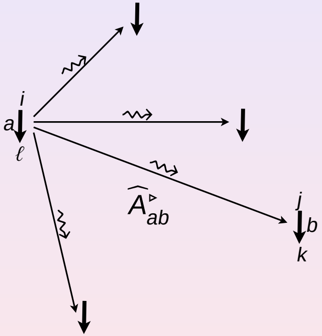
- j's push $r_j^{\blacktriangleright}$, and by
- k's pull $r_k^{\blacktriangleleft}$,

which are assumed to be mutually independent , then

expected unbiased flow from j to k is

$$
\begin{aligned}
r_{jk}^{\bowtie} &= r_j^{\blacktriangleright} r_k^{\blacktriangleleft} \\
&= \sum_{i\ell} A_{ij}^{\blacktriangleright} r_{i\ell}^{\bowtie} A_{k\ell}^{\blacktriangleleft} \\
&= \sum_{i\ell} \frac{A_{ij} A_{j\bullet} A_{\bullet k} A_{k\ell}}{A_{i\bullet} A_{\bullet\bullet}^2 A_{\bullet\ell}} r_{i\ell}^{\bowtie}
\end{aligned}
$$

Idea: capture path dynamics



Outline

Path network

Definition

Given

- path complete network $A = \left(R \stackrel{\gamma}{\leftarrow} E \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$

we define

- *path network* $\widehat{A} = \left(R \stackrel{\widehat{\gamma}}{\leftarrow} \widehat{E} \underset{\widehat{\varrho}}{\overset{\widehat{\delta}}{\rightrightarrows}} \widehat{N}\right)$, with

Path network

Definition

Given

- path complete network $A = \left(R \stackrel{\gamma}{\leftarrow} E \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$

we define

- *path network* $\widehat{A} = \left(R \stackrel{\widehat{\gamma}}{\leftarrow} \widehat{E} \underset{\widehat{\varrho}}{\overset{\widehat{\delta}}{\rightrightarrows}} \widehat{N}\right)$, with
- nodes $\widehat{N} = E$

Path network

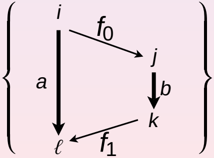
Definition

Given

- path complete network $A = \left(R \xleftarrow{\gamma} E \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$

we define

- *path network* $\widehat{A} = \left(R \xleftarrow{\widehat{\gamma}} \widehat{E} \underset{\widehat{\varrho}}{\overset{\widehat{\delta}}{\rightrightarrows}} \widehat{N}\right)$, with
- nodes $\widehat{N} = E$
- links $\widehat{E} = \sum_{a,b\in E} \widehat{E}_{ab}$, where $\widehat{E}_{ab} =$



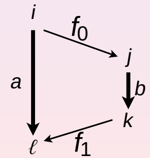
Path network

Definition

Given

- path complete network $A = \left(R \overset{\gamma}{\leftarrow} E \underset{\varrho}{\overset{\delta}{\rightrightarrows}} N\right)$

we define

- *path network* $\widehat{A} = \left(R \overset{\widehat{\gamma}}{\leftarrow} \widehat{E} \underset{\widehat{\varrho}}{\overset{\widehat{\delta}}{\rightrightarrows}} \widehat{N}\right)$, with
- nodes $\widehat{N} = E$
- links $\widehat{E} = \sum_{a,b\in E} \widehat{E}_{ab}$, where $\widehat{E}_{ab} = \left\{ \right.$  $\left. \right\}$

- cost $\widehat{\gamma}(f) = \gamma(f_0) + \gamma(b) + \gamma(f_1) - \gamma(a) + 2d \leq v$

Dynamics of path selection

attraction: probability that traffic through a will traverse b

$$\widehat{A}^{\triangleright}_{ab} = \frac{\widehat{A}_{ab}}{\widehat{A}_{a\bullet}} \text{ where}$$

$$\widehat{A}_{a\bullet} = \sum_x A_{ax}$$

repulsion: probability that traffic through b is diverted away from a

$$\widehat{A}^{\triangleleft}_{ab} = \frac{\widehat{A}_{ab}}{\widehat{A}_{\bullet b}} \text{ where}$$

$$\widehat{A}_{\bullet b} = \sum_x A_{xb}$$

Path ranking

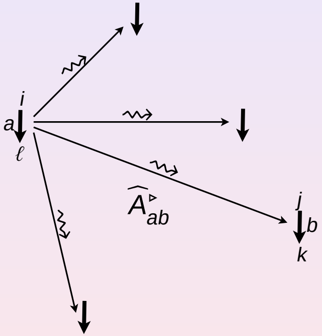
path pull: probability that traffic traverses b

$$\widehat{r}^{\triangleright}_b = \sum_a \widehat{r}^{\triangleright}_a \widehat{A}^{\triangleright}_{ab}$$

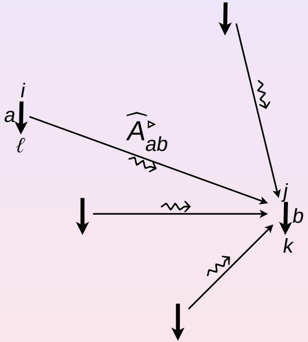
path push: probability that traffic is diverted from a

$$\widehat{r}^{\triangleleft}_a = \sum_{b \in \widehat{N}} \widehat{A}^{\triangleleft}_{ab} \widehat{r}^{\triangleleft}_b$$

Path pull rank



Path push rank



Attraction dynamics

Definition

The *node attraction* between j and k is the total attraction of all paths between them:

$$\widehat{r}_{jk} = \sum_{j \underset{b}{\to} k} \widehat{r}_b$$

Attraction dynamics

Definition

The *node attraction* between j and k is the total attraction of all paths between them:

$$\widehat{r}_{jk} = \sum_{j \xrightarrow[b]{} k} \widehat{r}_b$$

Idea

Estimate the traffic bias as the difference between the node attraction and the unbiased flow

$$\Upsilon_{jk} = \widehat{r}_{jk} - r^{\bowtie}_{jk}$$

Attraction dynamics

Definition

The *attraction dynamics* of a network A is the Markov chain $\widehat{A} = \left(\widehat{A}_{(ij)(k\ell)}\right)_{N^2\times N^2}$, with the entries

$$\widehat{A}_{(ij)(k\ell)} = \frac{A_{ij}A_{jk}A_{k\ell}}{A_{i\bullet}A_{\bullet\bullet}A_{\bullet\ell}}$$

where $A_{i\bullet}A_{\bullet\bullet}A_{\bullet\ell} = \sum_{m,n\in N} A_{im}A_{mn}A_{n\ell}$.

Attraction dynamics

Definition

The *attraction dynamics* of a network A is the Markov chain $\widehat{A} = \left(\widehat{A}_{(ij)(k\ell)}\right)_{N^2 \times N^2}$, with the entries

$$\widehat{A}_{(ij)(k\ell)} = \frac{A_{ij}A_{jk}A_{k\ell}}{A_{i\bullet}A_{\bullet\bullet}A_{\bullet\ell}}$$

where $A_{i\bullet}A_{\bullet\bullet}A_{\bullet\ell} = \sum_{m,n\in N} A_{im}A_{mn}A_{n\ell}$.

Recall

$$A^{\bowtie}_{(ij)(k\ell)} = \frac{A_{ij}A_{j\bullet}A_{\bullet k}A_{k\ell}}{A_{i\bullet}A_{\bullet\bullet}A_{\bullet\bullet}A_{\bullet\ell}}$$

and

$$r^{\bowtie}_{jk} = \sum_{i,\ell\in N} A^{\bowtie}_{(ij)(k\ell)} r^{\bowtie}_{i\ell}$$

Proposition

Let a network A be path complete for a sufficiently large cutoff value v.

Then the node attraction $\widehat{r}$ is the stationary distribution of the attraction dynamics:

$$\widehat{r}_{jk} \quad = \quad \sum_{i,\ell \in N} \widehat{A}_{(ij)(k\ell)} \widehat{r}_{i\ell}$$

Corollary

The directed reputation and promotion ranks are the marginals of the node attraction

$$\sum_{k \in N} \widehat{r}_{jk} = r_j^{\blacktriangleright}$$

$$\sum_{j \in N} \widehat{r}_{jk} = r_k^{\blacktriangleleft}$$

Interpretation

$$r_j^{\blacktriangleright} = \mathrm{Prob}\left(\bullet \xrightarrow{\xi} j \mid \xi \in \widehat{A}\right)$$

$$r_k^{\blacktriangleleft} = \mathrm{Prob}\left(k \xrightarrow{\xi} \bullet \mid \xi \in \widehat{A}\right)$$

$$\widehat{r_{jk}} = \mathrm{Prob}\left(j \xrightarrow{\xi} k \mid \xi \in \widehat{A}\right)$$

Interpretation

The mutual information

$$I(r^{\blacktriangleright} \; ; \; r^{\blacktriangleleft}) \;=\; D(\widehat{r} \,\|\, r^{\bowtie}) \;=\; \sum_{j=1}^{N} \sum_{k=1}^{N} \widehat{r}_{jk} \log \frac{\widehat{r}_{jk}}{r^{\blacktriangleright}_{j} r^{\blacktriangleleft}_{k}}$$

quantifies the *non-local* information processing in A.

Interpretation

The mutual information

$$I(r^{\blacktriangleright} \; ; \; r^{\blacktriangleleft}) \;=\; D(\widehat{r} \,\|\, r^{\bowtie}) \;=\; \sum_{j=1}^{N}\sum_{k=1}^{N} \widehat{r}_{jk} \log \frac{\widehat{r}_{jk}}{r_j^{\blacktriangleright} r_k^{\blacktriangleleft}}$$

quantifies the *non-local* information processing in A.

E.g, in the extremal cases,

- if $I(r^{\blacktriangleright} \; ; \; r^{\blacktriangleleft}) = 0$, i.e. $r^{\blacktriangleright}$ and $r^{\blacktriangleleft}$ are independent, all information is generated by the nodes,
- if $I(r^{\blacktriangleright} \; ; \; r^{\blacktriangleleft}) = H(r)$ for $r = r^{\blacktriangleright} = r^{\blacktriangleleft}$, all information is generated by the network.

Outline

Communities

- *Attraction bias* is the difference between total attraction and the expected flow

$$\Upsilon_{jk} = \widehat{r}_{jk} - r^{\bowtie}_{jk}$$

Communities

- *Attraction bias* is the difference between total attraction and the expected flow

$$\Upsilon_{jk} \quad = \quad \widehat{r}_{jk} - r^{\bowtie}_{jk}$$

- *Coherence* of $U \subseteq N$ is the minimal attraction bias of its members, in either direction

$$\Upsilon(U) \quad = \quad \bigwedge_{i,j \in U} (\Upsilon_{ij} \vee \Upsilon_{ji})$$

Communities

- *Attraction bias* is the difference between total attraction and the expected flow

$$\Upsilon_{jk} = \widehat{r}_{jk} - r^{\bowtie}_{jk}$$

- *Coherence* of $U \subseteq N$ is the minimal attraction bias of its members, in either direction

$$\Upsilon(U) = \bigwedge_{i,j \in U} (\Upsilon_{ij} \vee \Upsilon_{ji})$$

- *Communities* are coherent sets of nodes

$$\wp_\epsilon A = \{U \subseteq N \mid \Upsilon(U) \geq \epsilon\}$$

Communities

- Each $\wp_\epsilon A$, ordered by

$$U \sqsubseteq V \iff U \subseteq V \wedge \Upsilon(U) \leq \Upsilon(V)$$

is a *directed complete partial order (dcpo)*.

Communities

- Each $\wp_\epsilon A$, ordered by

$$U \sqsubseteq V \iff U \subseteq V \wedge \Upsilon(U) \leq \Upsilon(V)$$

is a *directed complete partial order (dcpo)*.

- An ϵ-*concept* is a maximal element of $\wp_\epsilon(A)$.

Communities

- Each $\wp_\epsilon A$, ordered by

$$U \sqsubseteq V \iff U \subseteq V \wedge \Upsilon(U) \leq \Upsilon(V)$$

  is a *directed complete partial order (dcpo)*.

- An ϵ-*concept* is a maximal element of $\wp_\epsilon(A)$.

- $\epsilon_1 \leq \epsilon_2$ implies $\wp_{\epsilon_1} A \supseteq \wp_{\epsilon_2} A$
  - $\wp_\epsilon$ is easy for large and small ϵ

Concept modules

Definition

$U \subseteq N$ is an ϵ-*concept module* if

- $\forall i, j \in U.\ \Upsilon(\{i, j\}) \geq \epsilon$, but
- $\forall k \in N \setminus U\ \exists j \in U.\ \Upsilon(\{k, j\}) < \epsilon$.

Concept modules

Definition

$U \subseteq N$ is an ϵ-*concept module* if

- $\forall i, j \in U.\ \Upsilon\left(\{i, j\}\right) \geq \epsilon$, but
- $\forall k \in N \setminus U\ \exists j \in U.\ \Upsilon\left(\{k, j\}\right) < \epsilon$.

Let $\mathcal{N}^{\epsilon}$ denote the set of ϵ-concept modules in a network A.

Association networks

Given a path complete network A, the induced concept network

$$\mathcal{A}^\epsilon = \left(R \xleftarrow{\gamma} \mathcal{E}^\epsilon \underset{\varrho}{\overset{\delta}{\rightrightarrows}} \mathcal{N}^\epsilon\right)$$

consists of

$$\begin{aligned}
\mathcal{N}^\epsilon &= \epsilon\text{-concept modules in } A \\
\mathcal{E}^\epsilon &= \sum_{U,V\in\mathcal{N}^\epsilon} \mathcal{E}^\epsilon_{UV} \text{ where} \\
\mathcal{E}^\epsilon_{UV} &= \sum_{U \underset{a}{\to} U\cap V} \; \sum_{U\cap V \underset{b}{\to} V} \tilde{E}_{ab} \\
\tilde{E}_{ab} &= \left\{ \begin{array}{ccc} i & \xrightarrow{f_0} & j \\ a \big\downarrow & & \big\downarrow b \\ \ell & \xrightarrow{f_1} & k \end{array} \right\}
\end{aligned}$$

Outline

Summary

- extended the ranking methods to paths
  - path rank is a measure of nonlocal information
  - allows estimating the flow bias to extract modules

Summary

- extended the ranking methods to paths
  - path rank is a measure of nonlocal information
  - allows estimating the flow bias to extract modules
- extracted modules (= communities = concepts)
  - parametric, richer structure for simpler algorithmics
  - concept networks for latent semantics

Current work

Is this method "real"?

- experimental validation
  - PL networks
  - IMDB
- relate with spectral methods
- algorithmics, convergence. . .