FASCINATE: Fast Cross-Layer Dependency Inference on Multi-Layered Networks

Presented by Chen Chen

Chen Chen  Hanghang Tong  Lei Xie  Lei Ying  Qing He
# Multi-Layered Networks are Everywhere!

## Infrastructure Networks
- Power Grid
- AS Network
- Transportation Network

## Cross-Layer
- **Power Supply, Control**
  - (Power Station→Routers)
- **Power Supply, Fuel Supply**
  - (Power Station→Transportation)
- **Control**
  - (Routers→Transportation)

## Collaboration Platforms
- Team Network
- Social Network
- Information Network

## Bio Systems
- Chemical Network
- Drug Network
- Disease Network
- PPI Network

## Intra-Layer
- Power Supply, Control
- Membership
  - (Team→Employees)
- Specialization
  - Employee→Information
- Composition
  - (Chemical→Drug)
- Treatment
  - (Drug→Disease)
- Association
  - (Disease→PPI)
Cross-Layer Dependency

- **Role:** Unique topology characteristic of multi-layered networks

- **Importance:** Key to multi-layered network mining tasks (e.g. connectivity control, robustness analysis)

- **Challenge:** Incomplete cross-layer dependencies
Q1: How to infer the hidden cross-layer dependencies?
Dependencies of Zero-Start nodes

- **Obs.** New nodes are emerging over time

- **Q2:** How to efficiently infer the dependencies of zero-start nodes?
Roadmap

- Motivation
- Q1: Cross-Layer Dependency Inference
  - Q2: Dependencies for Zero-Start Nodes
- Evaluations
- Conclusions
Background: A (Simplified) Multi-layered Network Model

- **A tuple** $\Gamma = \langle G, A, D \rangle$
  - $G$: layer-layer dependency network
  - $A$: intra-layer connectivity
  - $D$: cross-layer dependence

Q1: Dependency Inference

- **Key Idea 1**: Collaborative Filtering

Dependency Only

Users ≈ Routers | Movies ≈ Transportation | Known Ratings ≈ Observed Cross-Layer Dependency

\[ R \approx F_1 \times F_2' \]
Q1: Dependency Inference

- **Key Idea 2:** Collaborative Filtering with Side Information

Two-layered Network

![Diagram](image-url)

Movie-Movie Similarity \( \approx \) Transportation Network \| Social Network \( \approx \) AS Network

Known Ratings \( \approx \) Support from Routers to Transportation Network
Node Homophily

- **Assumption**: closely connected entities within each layer tend to have similar latent profiles.

\[
F(u_1, :) \approx F(u_2, :)
\]

\[
(\min \text{tr}(F'(D_U - U)F))
\]

Power Grid

AS Network

Transportation Network

Celebrities \approx Power Plants | Users \approx Routers | Movies \approx Transportation

Known Ratings, Movie Cast, Fans \approx Observed Cross-Layer Dependencies
Q1: Dependency Inference

- **Key Idea 3:** Collective Collaborative Filtering

Multi-layered Network

- Actor Similarity
  \[ A_1 \Rightarrow F_1 \]

- Actor-Movie Cast
  \[ D_{1,3} \approx F_1 \times F_3' \]

- Movie Similarity
  \[ A_3 \Rightarrow F_3 \]

- Actor-User Fans
  \[ D_{1,2} \approx F_1 \times F_2' \]

- User Similarity
  \[ A_2 \Rightarrow F_2 \]

- User-Movie Ratings
  \[ D_{2,3} \approx F_2 \times F_3' \]

- Known Ratings, Movie Cast, Fans \( \approx \) Observed Cross-Layer Dependencies
Optimization Problem

- **Objective Function:**

\[
\min_{F_i \geq 0 (i=1,\ldots,g)} J = \sum_{i,j:G(i,j)=1} \| W_{i,j} \odot (D_{i,j} - F_i F_j') \|_F^2 + J = \sum_{i,j:G(i,j)=1} \| W_{i,j} \odot (D_{i,j} - F_i F_j') \|_F^2 + 
\]

- **Challenge:** Not jointly convex w.r.t. \( F_i (i=1,\ldots,g) \)!

- **Q:** How to find a **local** optimal?

Matching observed cross-layer dependencies

Node homophily

Regularization

Hard to find **global** optimal solution!
**FACINATE: Proposed Solution**

- **Obs.:** $J$ becomes convex if we fix all but one (e.g. $F_i$) latent matrices

- **Method:** Block coordinate descent
  Fixing all other $F_j(j \neq i)$, the objective function w.r.t. $F_i$ is
  \[
  \min_{F_i \geq 0} J_i = \sum_{j:G(i,j)=1} \| W_{i,j} \odot (D_{i,j} - F_iF_j') \|_F^2 + \alpha tr(F_i'(T_i - A_i)F_i) + \beta \| F_i \|_F^2
  \]
  - Cross-layer dependencies that involve layer $i$
  - Homophily in layer $i$
  - Layer regularization

- **Multiplicative Update Rules:**
  \[
  F_i(u, v) \leftarrow F_i(u, v) \frac{X(u, v)}{Y(u, v)}
  \]
  \[
  X = \sum_{j:G(i,j)=1} (W_{i,j} \odot W_{i,j} \odot D_{i,j})F_j + \alpha A_iF_i
  \]
  \[
  Y = \sum_{j:G(i,j)=1} (W_{i,j} \odot W_{i,j} \odot (F_iF_j'))F_j + \alpha T_iF_i + \beta
  \]
Roadmap

- Motivation
- Q1: Cross-Layer Dependency Inference
- Q2: Dependencies for Zero-Start Nodes
- Evaluations
- Conclusions
Q2: Dependencies for zero-start nodes

\[ A_1 \]

\[ s \]

\[ \hat{A}_1 = \begin{bmatrix} A_1 & s' \\ s & 0 \end{bmatrix} \]

\[ F_i \rightarrow \hat{F}_i \ (i \neq 1) \]

\[ F_1 \rightarrow \hat{F}_1 = [\hat{F}_1(1_{n1} \times r), f] \]

Decompose Objective Function

\[ \hat{j} = J + J^1 \]

\[ \begin{cases} J: \text{objective function without zero-start node} \\ J^1: \alpha \sum_{v=1}^{n_1} s(v) \parallel f - \hat{F}_1(v,:) \parallel^2_2 + \parallel f \parallel^2_2 \end{cases} \]

Local Neighbors

Existing Nodes

New Node
Q2: Dependencies for zero-start nodes

- **Objective Function with Zero-Start Node:**

  \[
  \min_{\hat{F}_i \geq 0} \hat{j} = J + J^1 \\
  J: \text{objective function without zero-start node} \\
  J^1: \alpha \sum_{v=1}^{n_1} s(v) \| f - \hat{F}_1(v,:) \|_2^2 + \| f \|_2^2
  \]

- **Local Search Assumption:**

  \[
  \hat{F}_{1(n_1 \times r)} \approx F_1 \\
  \hat{F}_i \approx F_i \ (i \neq 1)
  \]

- **Solution:**

  \[
  \min_{\hat{F}_i \geq 0} \hat{j} = J + J^1 \\
  \rightarrow \min_{f \geq 0} J^1 \ \text{sub. to } \hat{F}_{1(n_1 \times r)} = F_1^* \\
  f = \frac{\alpha s F_1^*}{\beta + \alpha \sum_{v=1}^{n_1} s(v)}
  \]

  Only related to zero-start node’s local neighbors!
Roadmap

- Motivation
- Q1: Cross-Layer Dependency Inference
- Q2: Dependencies for Zero-Start Nodes
- Evaluations
- Conclusions
Experimental Set-up

- **Datasets:**

  - **SOCIAL**
    - Paper
    - Author
    - Venue

  - **BIO**
    - Chemical
    - Gene
    - Disease

  - **INFRA-5**
    - Internet
    - R1
    - R2
    - R3
    - R4

  - **INFRA-3**
    - Power
    - AS

<table>
<thead>
<tr>
<th>Datasets</th>
<th>#Layers</th>
<th>#Nodes</th>
<th>#Links</th>
<th>#CrossLinks</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOCIAL</td>
<td>3</td>
<td>125,344</td>
<td>214,181</td>
<td>188,844</td>
</tr>
<tr>
<td>BIO</td>
<td>3</td>
<td>35,631</td>
<td>253,827</td>
<td>75,456</td>
</tr>
<tr>
<td>INFRA-5</td>
<td>5</td>
<td>349</td>
<td>379</td>
<td>565</td>
</tr>
<tr>
<td>INFRA-3</td>
<td>3</td>
<td>15,126</td>
<td>29,861</td>
<td>28,023,500</td>
</tr>
</tbody>
</table>

- **Evaluation Objectives:**
  - Effectiveness: How accurate is FACSINATE?
  - Efficiency: How fast is FACSINATE?
**Effectiveness of FASCINATE (Q1)**

Cross-layer dependency inference on BIO dataset

<table>
<thead>
<tr>
<th>Methods</th>
<th>MAP</th>
<th>R-MPR</th>
<th>HLU</th>
<th>AUC</th>
<th>Prec@10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FASCINATE</strong></td>
<td>0.3979</td>
<td>0.4066</td>
<td>45.1001</td>
<td>0.9369</td>
<td>0.1039</td>
</tr>
<tr>
<td><strong>FASCINATE-CLUST</strong></td>
<td>0.3189</td>
<td>0.3898</td>
<td>37.4089</td>
<td>0.9176</td>
<td>0.0857</td>
</tr>
<tr>
<td>MulCol</td>
<td>0.3676</td>
<td>0.3954</td>
<td>42.8687</td>
<td>0.9286</td>
<td>0.0986</td>
</tr>
<tr>
<td>PairSid</td>
<td>0.3623</td>
<td>0.3403</td>
<td>40.4048</td>
<td>0.8682</td>
<td>0.0941</td>
</tr>
<tr>
<td>PairCol</td>
<td>0.3493</td>
<td>0.3153</td>
<td>38.4364</td>
<td>0.8462</td>
<td>0.0889</td>
</tr>
<tr>
<td>PairNMF</td>
<td>0.1154</td>
<td>0.1963</td>
<td>15.8486</td>
<td>0.6865</td>
<td>0.0393</td>
</tr>
<tr>
<td>PairRec</td>
<td>0.0290</td>
<td>0.2330</td>
<td>3.6179</td>
<td>0.7105</td>
<td>0.0118</td>
</tr>
<tr>
<td>FlatNMF</td>
<td>0.2245</td>
<td>0.2900</td>
<td>26.1010</td>
<td>0.8475</td>
<td>0.0615</td>
</tr>
<tr>
<td>FlatRec</td>
<td>0.0613</td>
<td>0.3112</td>
<td>8.4858</td>
<td>0.8759</td>
<td>0.0254</td>
</tr>
</tbody>
</table>

**FASCINATE** performs best!
Parameter Studies

- Parameters: \( \alpha, \beta, r \)

\[
\min_{F_i \geq 0} J = \sum_{i,j: G(i,j) = 1} \| W_{i,j} \odot (D_{i,j} - F_i F_j') \|^2_F + \alpha \sum_i \text{tr}(F_i'(T_i - A_i) F_i) + \beta \sum_i \| F_i \|^2_F
\]

\( (r: \text{rank of } F_i (i = 1, \ldots, g)) \)

Impact of \( \alpha \) and \( r \)  
Impact of \( \beta \) and \( r \)  
Impact of \( \alpha \) and \( \beta \)

FASCINATE is stable in wide range of parameter settings!
Effectiveness of FASCINATE-ZERO (Q2)

- FASCINATE-ZERO vs. FASCINATE

FASCINATE-ZERO: similar performance, faster speed!
Scalability

FASCINATE (Q1)
Linear

FASCINATE-ZERO (Q2)
Sub-linear
Roadmap

✓ ▪ Motivation
✓ ▪ Q1: Cross-layer Dependency Inference
✓ ▪ Q2: Dependencies for Zero-Start Nodes
✓ ▪ Evaluations
✓ ▪ Conclusions
Conclusions

- **Cross-Layer Dependency Inference**
  - **Key Ideas:**
    - Collective Collaborative Filtering + Node Homophily
    - Local Search (for zero-start nodes)
  - **Methods:** FASCINATE & FASCINATE-ZERO

- **Results**
  - **Effectiveness:** 8.2%-41.9% over best competitors
  - **Efficiency:** linear (FASCINATE), sublinear (FASCINATE-ZERO)

- **More in paper**
  - Variants
  - Convergence analysis & Effectiveness results

- **Code:** [http://www.public.asu.edu/~cchen211]