FEMA: FLEXIBLE EVOLUTIONARY MULTI-FACETED ANALYSIS FOR DYNAMIC BEHAVIOR PATTERN DISCOVERY

Meng Jiang, Tsinghua University, Beijing, China

Joint work with Peng Cui, Fei Wang, Xinran Xu, Wenwu Zhu and Shiqiang Yang
August 25, 2014 – NYC, USA
Behavior Analysis

Modeling

How to formulate human behavior?

KDD’13
KDD’14

Pattern discovery

How to understand human behavior?

Prediction

What is the missing human behavior?

Data Mining
Our Goals

• Given: Behavioral data sequence
• Find: A general framework that fast and best fit the behavioral data

• Goals:
  • G1. Model the human behavior
  • G2. Understand the hidden patterns
  • G3. Predict the missing behavior
Human Behavior

• Write a paper/book

• Post a photo on Facebook
Human Behavior: Multi-faceted

- Write a paper/book

- Post a photo on Facebook
Human Behavior: Dynamic

- Write a paper/book
Human Behavior: Dynamic

- Post Facebook messages

- Tsinghua
- WWW’14
- Tsinghua
- KDD’14
Human Behavior

• Multi-faceted
• Dynamic

• How to model human behavior?
Model Human Behavior

Human behavior

Multi-faceted

Dynamic

Problem

Behavior modeling

Pattern discovery

Behavior prediction

Tensor sequence

Decomposition Completion

≈

x

x

x
Challenges

• High sparsity
  • High-order tensors

• High complexity
  • Long sequence of tensors
  • Too slow if decomposing at each time
Idea

• High sparsity

• Auxiliary knowledge as regularizations
Idea

• High complexity

• Update projection matrices with new coming piece of data
OUTLINE

1. Background
2. Model Formulation
3. The Framework
4. Experiments
5. Visualization
FEMA: Flexible Evolutionary Multi-faceted Analysis
Tensor Perturbation Theory
**FEMA Algorithm**

**Approximation**

Require: $\mathcal{X}_t, \Delta \mathcal{X}_t, A_t^{(m)}|_{m=1}^M, \lambda_t^{(m)}|_{m=1}^M$

for $m = 1, \ldots, M$ do
  for $i = 1, \ldots, r(m)$ do
    Compute $\Delta \lambda_t^{(m)}$ using
    \[
    \Delta \lambda_t^{(m)} = a_i^{(m)\top}(X^{(m)}\Delta X^{(m)\top} + \Delta X^{(m)}X^{(m)\top})a_i^{(m)}
    \]
    and compute
    \[
    \lambda_{t+1, i} = \lambda_{t, i} + \Delta \lambda_{t, i};
    \]
    Compute $\Delta a_t^{(m)}$ using
    \[
    \Delta a_t^{(m)} = \sum_{j \neq i} a_j^{(m)\top}(X^{(m)}\Delta X^{(m)\top} + \Delta X^{(m)}X^{(m)\top})a_i^{(m)} / \lambda_t^{(m)} - \lambda_j^{(m)}a_j^{(m)}
    \]
    and compute
    \[
    a_{t+1, i} = a_{t, i} + \Delta a_{t, i} \text{ and } A_{t+1}^{(m)} = \{a_{t+1, i}\};
    \]
  end for
end for
\[
\mathcal{Y}_{t+1} = (\mathcal{X}_t + \Delta \mathcal{X}_t) \prod_{m=1}^M \times (m) A_t^{(m)\top};
\]
return $A_{t+1}^{(m)}|_{m=1}^M, \lambda_{t+1}^{(m)}|_{m=1}^M, \mathcal{Y}_{t+1}$

**Bound Guarantee**

Core tensor

\[
|\Delta \lambda_t^{(m)}| \leq 2(\lambda_{X^{(m)\top}X^{(m)}}^{\max})^{\frac{1}{2}} \|\Delta X^{(m)}\|_2
\]

Projection matrix

\[
|\Delta a_t^{(m)}| \leq 2\|\Delta X^{(m)}\|_2 \sum_{j \neq i} \frac{(\lambda_{X^{(m)\top}X^{(m)}}^{\max})^{\frac{1}{2}}}{|\lambda_t^{(m)} - \lambda_j^{(m)}|}
\]
Experiments: Test Behavior Prediction

- Data sets
- Leveraging multi-faceted information
- Leveraging flexible regularizations
- Efficiency, loss and parameters
Data Sets

• Microsoft Academic Search
  • Subset of top 100 experts from query “data mining”
  • Paper: <author, affiliation and keyword>
  • Regularization: co-authorship <author, author>
  • 7,777 x 651 x 4,566 x 32 years: 171,519 tuples

• Tencent Weibo
  • 43 days: Nov. 9, 2011 to Dec. 20, 2011
  • Tweet: <user-who-@, @-ed-user, word>
  • Regularization: social relation <user, user>
  • 6,200 x 1,813 x 6,435 x 43 days: 519,624 tuples
Leveraging Multi-faceted Information

Predict “Who”-“What keyword”
FEMA uses “Where” (affiliation).

Predict “Who”-“@Whom”
FEMA use “What” (tweet word).

<table>
<thead>
<tr>
<th></th>
<th>Microsoft Academic Search</th>
<th>Tencent Weibo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAE</td>
<td>RMSE</td>
</tr>
<tr>
<td>FEMA [X]</td>
<td>0.735</td>
<td>0.944</td>
</tr>
<tr>
<td>EMA [X]</td>
<td>0.794</td>
<td>1.130</td>
</tr>
<tr>
<td>EA [X]</td>
<td>0.979</td>
<td>1.364</td>
</tr>
</tbody>
</table>

Precision vs Recall

Graphs showing precision vs recall for FEMA, EMA, and EA.
Leveraging Flexible Regularizations


<table>
<thead>
<tr>
<th></th>
<th>Microsoft Academic Search</th>
<th>Tencent Weibo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAE</td>
<td>RMSE</td>
</tr>
<tr>
<td>FEMA</td>
<td>0.893</td>
<td>1.215</td>
</tr>
<tr>
<td>EMA</td>
<td>0.909</td>
<td>1.466</td>
</tr>
<tr>
<td>DTA [Sun et al.]</td>
<td>0.950</td>
<td>1.556</td>
</tr>
</tbody>
</table>

Precision vs Recall

![Graphs showing precision vs recall for different methods](image)
Efficiency, Loss and Parameters

Evolutionary analysis: update $\lambda$ and $a$ with $\Delta X$

Re-decompose updated matrices

Insensitive to regularization weight

The loss is small.
OUTLINE

1. Background
2. Model Formulation
3. The Framework
4. Experiments
5. Visualization
Visualization: Test Pattern Discovery

• Microsoft Academic Search
• Tencent Weibo (see our paper 😊)

• Behavior Patterns
  • Multi-faceted
  • Dynamic
Microsoft Academic Search
Microsoft Academic Search

Year 2005
Microsoft Academic Search

Year 2012

- Jiawei Han
  - Yongjian Fu
  - Philip Yu
  - Osmar Zaïane
  - Nick Cercone
  - Jian Pei
  - Deng Cai
  - Xiaofei He
  - Xifeng Yan
  - Yizhou Sun
  - Gao Jing

University of Illinois at Urbana Champaign

- Simon Fraser University
- State University of New York at Buffalo
- LBNL J. J. Wilson Research Center

recursive query
 frequent pattern
 association rules
 sequential pattern
 relational database

data clustering
 factorization model
 heterogeneous information network
 community detection
 social networks

University of Illinois at Chicago

University of California at Santa Barbara
 Zhejiang University
Conclusion

• **Human behavior**: multi-faceted and dynamic
• **Challenges**: high sparsity and high complexity
• **Solutions**: flexible regularizations & evolutionary analysis
• **FEMA**: approximation algorithm and bounds
• **Experiment**: behavior prediction
• **Visualization**: pattern discovery
Questions?

Meng Jiang
mjiang89@gmail.com
http://www.meng-jiang.com