Binary Independent Component Analysis: Theory and Applications in Networking

Huy Nguyen
WiSeR – Wireless System Research Group
Department of Computer Science
University of Houston, TX, USA

Ph.D. Showcase Presentation
Independent Component Analysis (ICA)

- Given the linear mixture of some unknown variables
  \[ x = G \times y \]
  (data) \hspace{1cm} \hspace{1cm} (linear mixing matrix) \hspace{1cm} (sources)
  
- \( y \): mutually independent variables

- Revealing underlying sources

- Application: Image processing, document databases, financial analysis …
Binary ICA with OR Mixtures

• Consider the model where observations ($x$) are disjunctive mixtures of binary independent sources ($y$)

$$x_i = \bigvee_{j=1}^{n} (g_{ij} \land y_j), \quad i = 1, \ldots, m$$

• ICA assumes continuous variables → not directly applicable

• Binary Independent Component Analysis: From observation matrix $x$, infer the mixing matrix $G$ and activity matrix $y$
Binary ICA Model

• \( n \) independent binary sources: \( y = [y_1, y_2, \ldots, y_n] \)
• \( m \) monitors: \( x = [x_1, x_2, \ldots, x_m] \)
• Binary mixing matrix:

\[
G = g_{ij} \in \{0, 1\}, i = [1, \ldots, m], j = [1, \ldots, n]
\]

\[
G = \begin{bmatrix}
1 & 1 & 0 \\
1 & 0 & 0 \\
0 & 1 & 1 \\
0 & 0 & 1 
\end{bmatrix}
\]

• Binary ICA model: \( x = G \odot y \) (unknown)
Binary ICA Inference Algorithm

- **Input:** Observation matrix $X$
- **Output:** Mixing matrix $G$, active probability $p$

```plaintext
FindBICA ()
if $m = 1$ then
    \[ p_0 = \mathcal{F}(x_1 = 0) \]
    \[ p_1 = \mathcal{F}(x_1 = 1) \]
else
    \[ p_{1:2^{m-1}-1} = \text{FindBICA}(X^{0}_{(m-1)\times T}) \]
    \[ p_{1:2^{m-1}-1} = \text{FindBICA}(X^{1}_{(m-1)\times T}) \]
for $l = 1, \ldots 2^{m-1} - 1$ do
    \[ p_l = 1 - \frac{1-p_l^*}{1-p_l^0} \]
```

Problem: Maximize number of monitored users

**x**: binary observations from sniffers

**y**: user activities

**G**: relationship between sniffers and users

Problem: PU Separation

\( x \): SU observations

\( y \): PU activities

\( G \): relationship between PUs and SUs

Huy Nguyen, Rong Zheng, and Zhu Han, “Binary is Good: A Binary Inference Framework for Primary User Separation in Cognitive Radio Networks”, In Proc. of CrownCom’10
Problem: Multicast topology inference

\( x \): binary packet loss observations from monitors

\( y \): link loss events

\( G \): relationship between links and monitors

Huy Nguyen and Rong Zheng, “Revisiting Tree Topology Inference: A Binary Independent Component Analysis Approach”, In submission to INFOCOM’11
THANK YOU FOR YOUR ATTENTION

nahuy@cs.uh.edu
http://www2.cs.uh.edu/~nahuy/