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)

Cocktail Party Problem

- Given the linear mixture of some unknown variables

$$\mathbf{x} = G \times \mathbf{y}$$

(data) (linear mixing matrix) (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**)

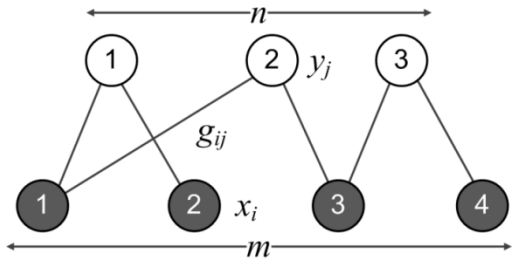
$$x_i = \bigvee_{j=1}^{n} (g_{ij} \wedge y_j), \quad i = 1, \ldots, m$$

- 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: $\mathbf{y} = [y_1, y_2, \ldots, y_n]$
- m monitors: $\mathbf{x} = [x_1, x_2, \ldots, x_m]$
- Binary mixing matrix:

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



$$\boldsymbol{G} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

- Binary ICA model: $\mathbf{x} = \underset{\text{(unknown)}}{\boldsymbol{G}} \otimes \underset{\text{(unknown)}}{\mathbf{y}}$

Binary ICA Inference Algorithm

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

```
FindBICA ()
if m = 1 then
    p_0 = F(x_1 = 0)
    p_1 = F(x_1 = 1)
else
    p^0_{1:2^{m-1}-1} = FindBICA (X^0_{(m-1)×T})
    p^*_{1:2^{m-1}-1} = FindBICA (X_{(m-1)×T})
    for l = 1, ... 2^{m-1} - 1 do
        p_{l+2^{m-1}} = 1 - (1-p^*_l)/(1-p^0_l)
```

Huy Nguyen and Rong Zheng, "Binary Independent Component Analysis with OR Mixtures", http://arxiv.org/abs/1007.0528

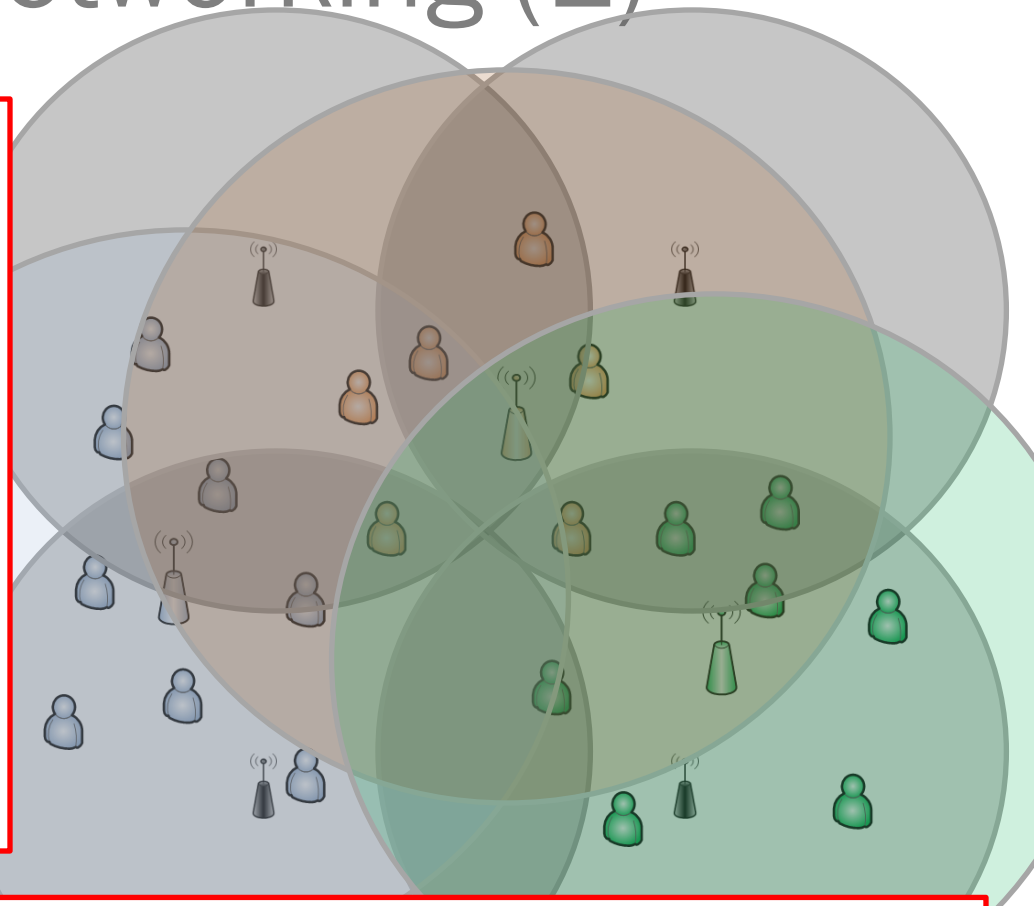
Application in Networking (1)

Problem: Maximize number of monitored users

x: binary observations from sniffers

y: user activities

G: relationship between sniffers and users



A. Chhetri, Huy Nguyen, G. Scalosub, and R. Zheng, "On Quality of Monitoring for Multi-channel Wireless Infrastructure Networks", In Proc. of Mobihoc'10

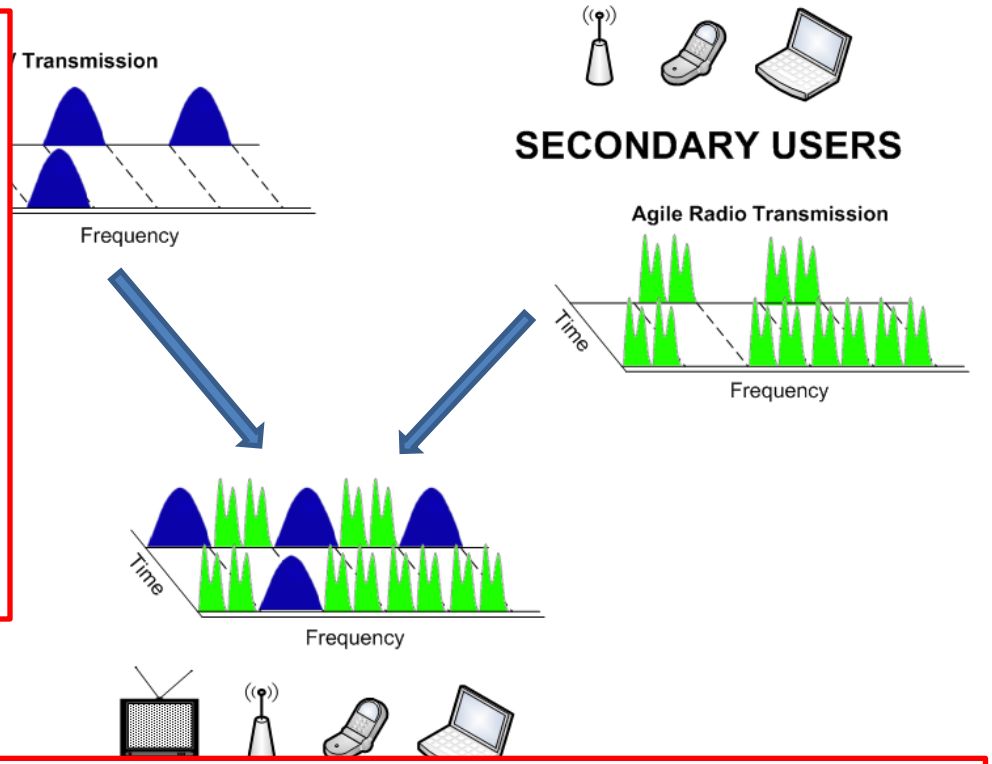
Application in Networking (2)

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

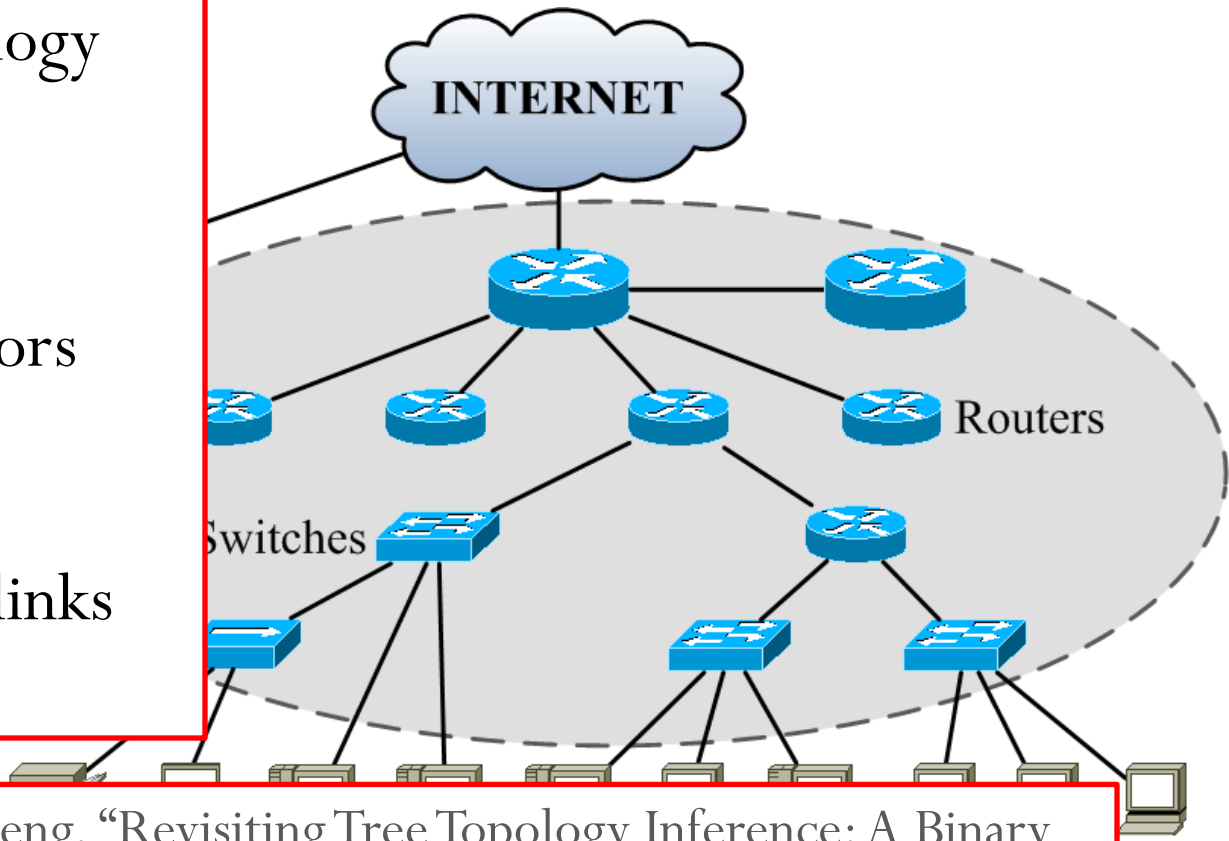
Application in Networking (3)

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/