Multi-relational Community Detection in Social Platforms with GNNs

Nouamane Arhachoui[1], Vincent Gauthier[2],
Anastasios Giovanidis[1], Lionel Tabourier[1]

[1] *Sorbonne Université, CNRS – LIP6, Paris, France*
[2] *Télécom SudParis, Institut Polytechnique de Paris, Palaiseau, France*

May 22nd, 2025

Motivation

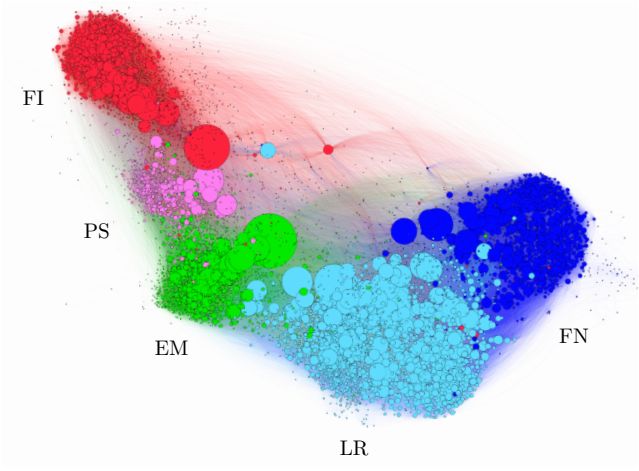
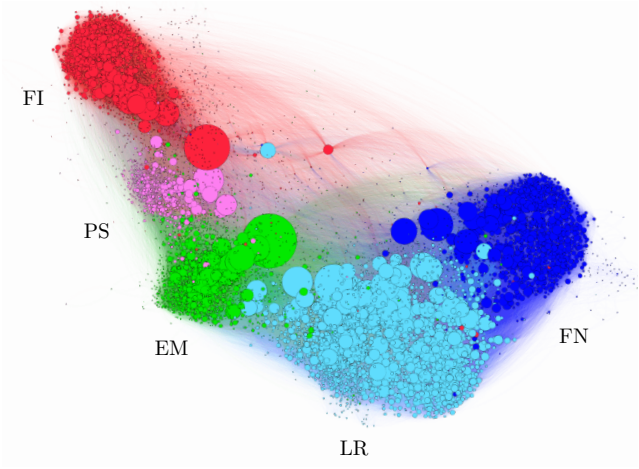


*Image by Effrosyni Papanastasiou

Motivation



Problem: Most existing methods focus on single-relational networks.

*Image by Effrosyni Papanastasiou

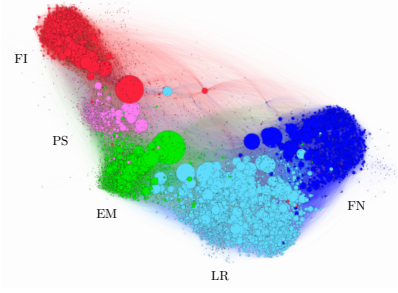
Motivation

SORBONNE UNIVERSITÉ

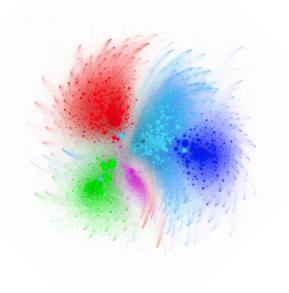
However... Most real-world networks are **multi-relational**.

Motivation

However... Most real-world networks are **multi-relational**.

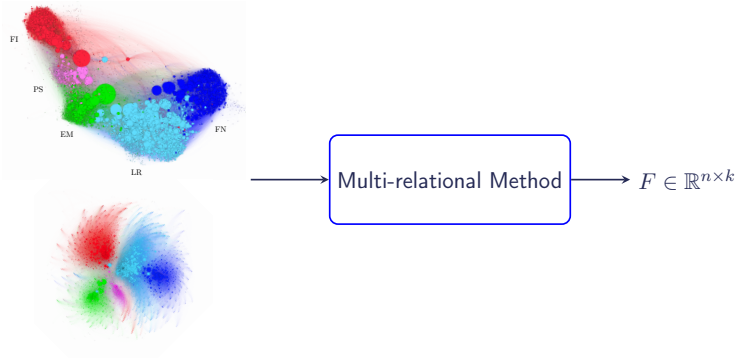


Follower Graph



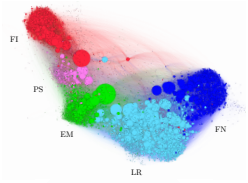
Retweet Graph

Problem Statement



F = community structure of the multi-relational network.

Problem Statement

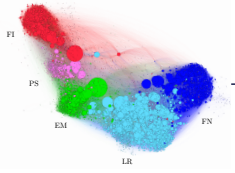


Multi-relational Method → $F \in \mathbb{R}^{n \times k}$

F = community structure of the multi-relational network.

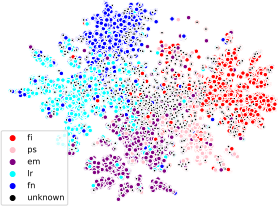
Multi-relational community: A group of nodes that are densely connected across multiple relations.

From Graphs to Embeddings



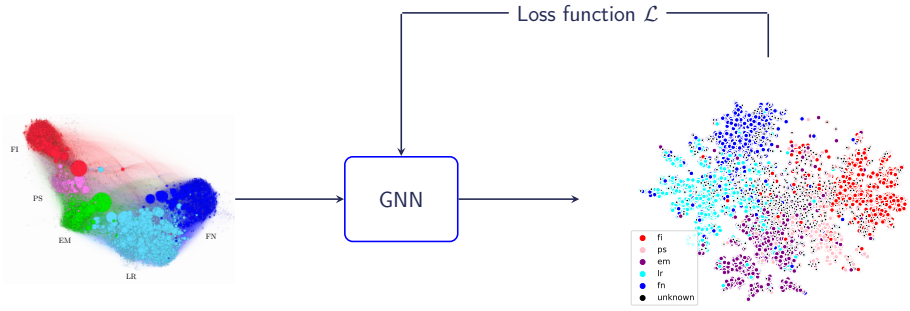
Input graph

GNN

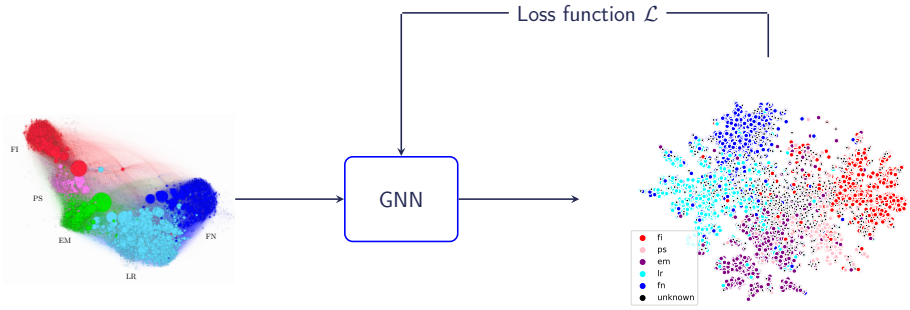


Embedding Space

From Graphs to Embeddings

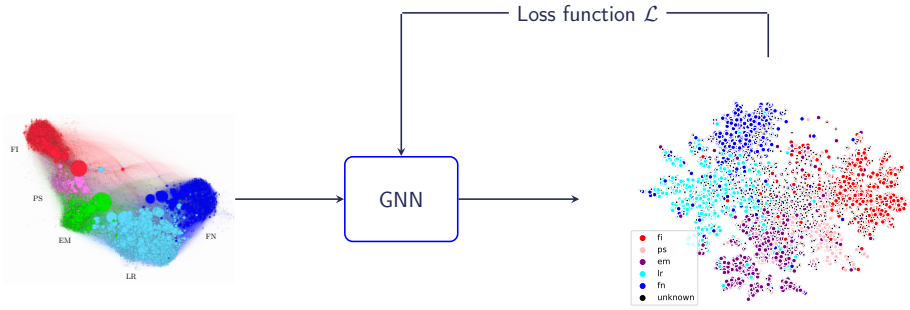


From Graphs to Embeddings



$$\mathcal{L} = -\log(P(G|F))$$

From Graphs to Embeddings



$$\mathcal{L} = -\log(P(G|F))$$

Goal: to fit a graph model that captures the community structure in the data.

BigClam model

In BigClam*:

$$P((u,v) \in E|F) = 1 - e^{-F_u^T \cdot F_v} \quad (1)$$

F_u: community affiliation vector of node u.

*Overlapping Community Detection at Scale: A Nonnegative Matrix Factorization Approach, Yang and Leskovec, WSDM 2013

BigClam model

In BigClam*:

$$P((u,v) \in E|F) = 1 - e^{-F_u^T \cdot F_v} \tag{1}$$

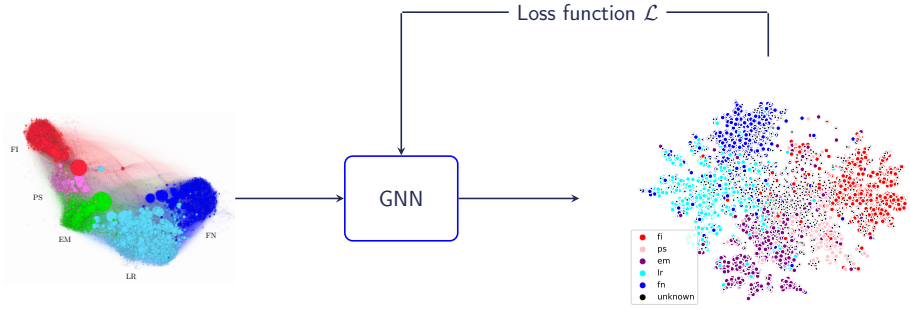
F_u: community affiliation vector of node u.

Loss:

$$\mathcal{L}(F) = -\sum_{(u,v)\in E} \log\left(1 - e^{-F_u^T \cdot F_v}\right) + \sum_{(u,v)\notin E} F_u^T \cdot F_v \tag{2}$$

*Overlapping Community Detection at Scale: A Nonnegative Matrix Factorization Approach, Yang and Leskovec, WSDM 2013

Minimization problem

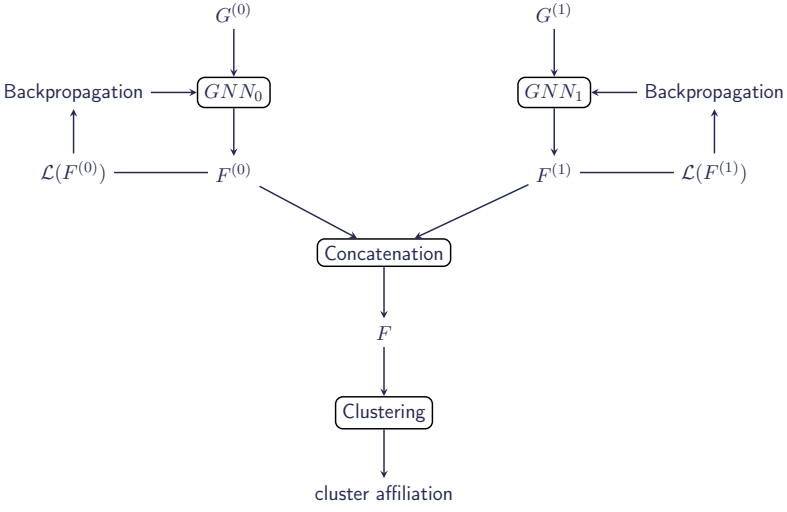


Repeat until convergence:

1. $F = \text{GNN}(G)$
2. Compute $\mathcal{L}(F)$
3. Update the GNN

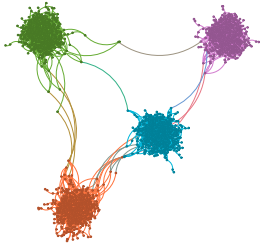
Learning pipeline

Case with 2 graphs/relations:



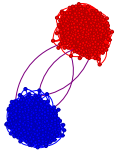
Experiment 1: same *a priori* communities

- 2 Planted Partition Model (PPM) graphs
  - n nodes, k communities

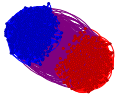


Experiment 1: same *a priori* communities

- 2 Planted Partition Model (PPM) graphs
  - n nodes, k communities
- Mixing parameter μ varies



Low μ

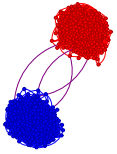


High μ

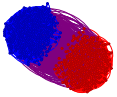
Experiment 1: same *a priori* communities

- 2 Planted Partition Model (PPM) graphs
  - n nodes, k communities
- Mixing parameter μ varies

$\Rightarrow$ Completely overlapping PPMs with different level of mixing



Low μ

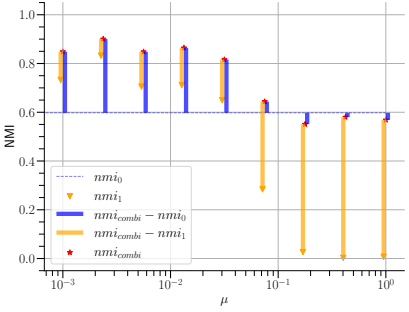


High μ

Experiment 1: same *a priori* communities

- 2 Planted Partition Model (PPM) graphs
  - n nodes, k communities
- Mixing parameter μ varies

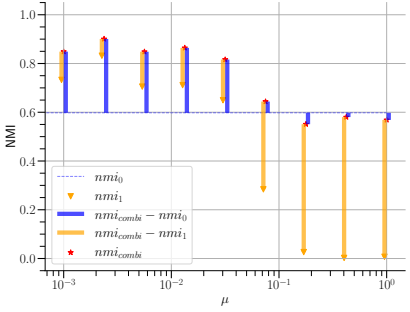
$\Rightarrow$ Completely overlapping PPMs with different level of mixing



Experiment 1: same *a priori* communities

- 2 Planted Partition Model (PPM) graphs
  - n nodes, k communities
- Mixing parameter μ varies

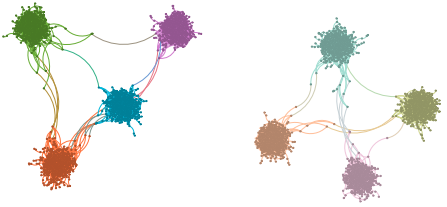
$\Rightarrow$ Completely overlapping PPMs with different level of mixing



Conclusion: Better performance with correlated communities and low mixing.

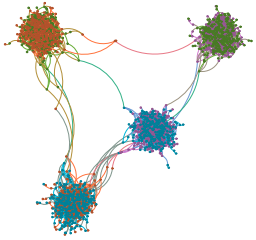
Experiment 2: partially overlapping PPMs

- 2 Planted Partition Model (PPM) graphs
  - n nodes, k communities



Experiment 2: partially overlapping PPMs

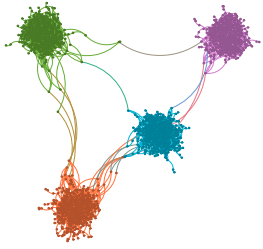
- 2 Planted Partition Model (PPM) graphs
  - n nodes, k communities
- Shifted communities parameterized by $s \in [0, 1]$



Experiment 2: partially overlapping PPMs

- 2 Planted Partition Model (PPM) graphs
  - n nodes, k communities
- Shifted communities parameterized by $s \in [0, 1]$

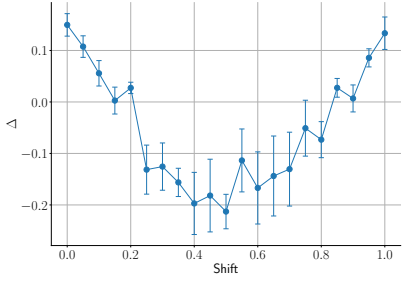
  Note: complete overlap for $s = 0$ or $s = 1$



Experiment 2: partially overlapping PPMs

- 2 Planted Partition Model (PPM) graphs
  - n nodes, k communities
- Shifted communities parameterized by $s \in [0, 1]$

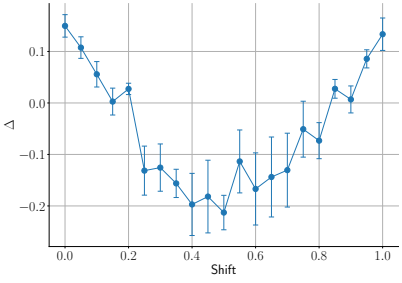
  Note: complete overlap for $s = 0$ or $s = 1$



Experiment 2: partially overlapping PPMs

- 2 Planted Partition Model (PPM) graphs
  - n nodes, k communities
- Shifted communities parameterized by $s \in [0, 1]$

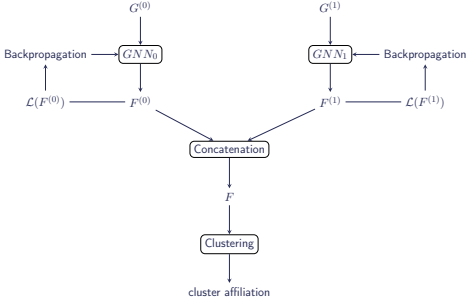
  Note: complete overlap for $s = 0$ or $s = 1$



Conclusion: $s \simeq 0.5 \Rightarrow$ hard to find the consensus k-community structure.

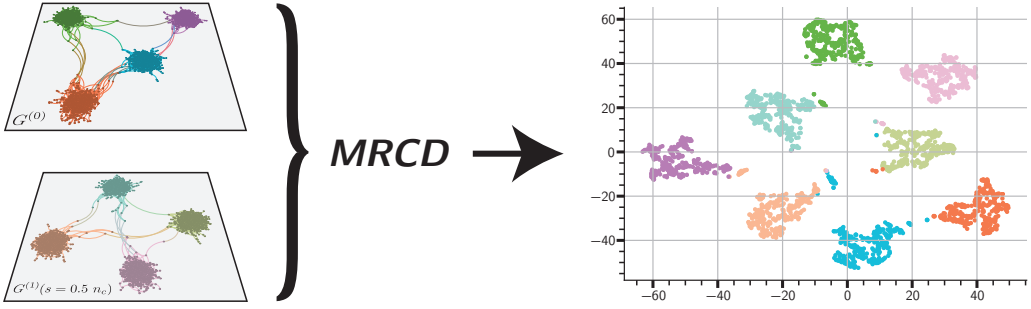
Experiment 2: partially overlapping PPMs

For $s = 0.5$, what happens to F when trying to find the $2k$ groups?



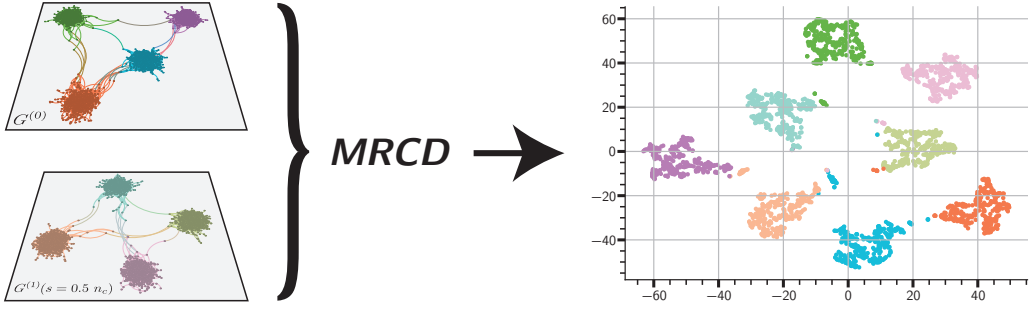
Experiment 2: partially overlapping PPMs

For $s = 0.5$, what happens to F when trying to find the $2k$ groups?



Experiment 2: partially overlapping PPMs

For $s = 0.5$, what happens to F when trying to find the $2k$ groups?

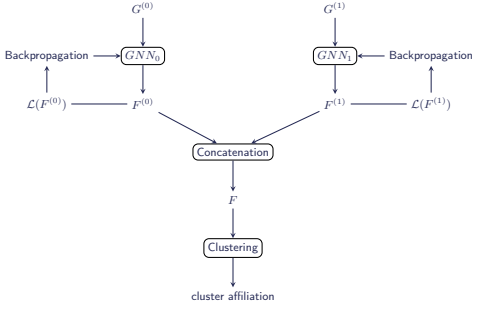


Conclusion: ability to find the overlaps between the predefined communities.

Conclusion

Conclusion:

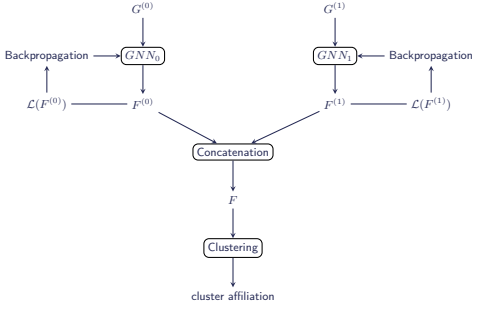
- Method for multi-relational community detection
- Better performance with correlated communities across relations



Conclusion

Conclusion:

- Method for multi-relational community detection
- Better performance with correlated communities across relations



Future work: explore real-world data with this method (*e.g.* polarization in social platforms)