Joint Learning of Multiplex Graphs & Nodal Attributes via Neural Network Gaussian Processes

Sharmistha Guha
Department of Statistics, Texas A&M University
Joint Work with:

Jose Rodriguez-Acosta (Texas A&M)

Lekha Patel (Sandia Labs)

Kurtis Shuler (Sandia Labs)

Supported by U.S. NSF Award No. 2413721 (DMS), LDRD (Sandia) and DGE-2139772 (NSF GRFP - Jose).

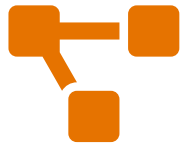
Motivation: The Challenge of Covert Terrorist Networks

Deliberately Obscured

Critical relationships hidden, threat identification exceptionally difficult.

Constantly Evolving

Ties shift and adapt under pressure.



Multiplex Relations

Networks are a mix of both alliances and rivalries, not a single type of tie.

Our Goal: A Unified Predictive Framework

- **Predict:** Evolving multiplex graphs (alliances & rivalries) and attributes.
- **Method:** Jointly model the time-varying network and its attributes.
- **To Enable:** Identification of threats, and prediction of behaviors for counter-terrorism efforts.

Data Description

Data Source:

- Big Allied and Dangerous (**BAAD**) dataset (Asal and Rethemeyer 2015).
- Observations of 15 terrorist organizations from 1998-2012.

Data Structure:

- **Multilayer Network:** One layer for alliances, one for rivalries.
- **Organizational (nodal) Attributes:** Organization size, lethality, ideology, capacity, leadership.
- Attibutes change, leaving observable traces - even as relationships remain hidden.

START
NATIONAL CONSORTIUM FOR THE
STUDY OF TERRORISM AND RESPONSES TO TERRORISM

PVC | PROJECT ON VIOLENT CONFLICT
Rockefeller College of Public Affairs & Policy
University at Albany, State University of New York

Core Premise & Goal:

- Jointly model the network and attributes to leverage signals that organizations cannot fully hide, allowing us to estimate concealed relationships.

Data Analysis

Three Critical Security Questions

Dynamic evolution of terrorist networks poses 3 key challenges for security intelligence:

Q1: The Problem of Incompleteness

Challenge: Data is incomplete and contains dual ties (alliance/rivalry) .
Need: A method to infer the complete multiplex network.

Q2: The Problem of Early Detection

Challenge: Detect pre-operational escalation signals - critical for preventive intelligence.
Need: A joint analysis of network nodes & attributes to detect escalation early.

Q3: The Problem of Cascading Effects

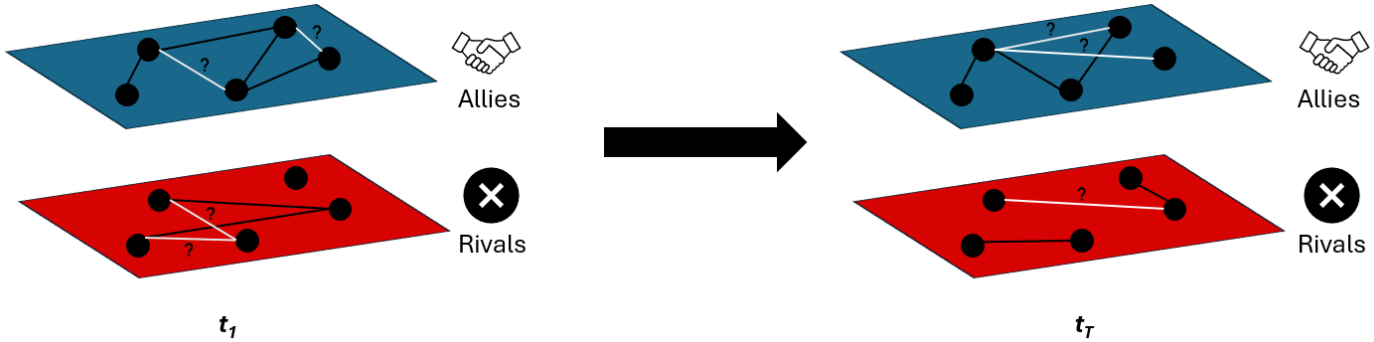
Challenge: To anticipate cascades across interconnected terrorist ecosystems (e.g., peace talks, leadership losses, territorial setbacks).
Need: Model the joint evolution of networks and organizational attributes.

BAAD: 15 Terrorist Organizations

1	Abu Sayyaf Group (ASG)
2	Hamas (Ha)
3	Hizballah (Hi)
4	Islamic Movement of Uzbekistan (IMU)
5	Kurdistan Workers' Party (PKK)
6	Lord's Resistance Army (LRA)
7	Moro Islamic Liberation Front (MILF)
8	Mujahedin-e Khalq (MEK)
9	National Liberation Army of Colombia (ELN)
10	Palestinian Islamic Jihad (PIJ)
11	Revolutionary Armed Forces Of Colombia (FARC)
12	Taliban (Ta)
13	United Liberation Front of Assam (ULFA)
14	Allied Democratic Forces (ADF)
15	Baloch Liberation Front (BLF)

List of the 15 terrorist organizations in BAAD dataset.

Notation



- Partially observed L layered time varying multiplex graph.
- $\mathbf{A}_l^{(o)}(t) = \{\, a_{jj',l}(t) \mid (j,j') \in \mathbf{U}_l^{(o)}(t) \,\}$: observed edges in layer l at time t.
- $\mathbf{U}_l^{(o)}(t) \subseteq \{1,\ldots,J\} \times \{1,\ldots,J\}$ denotes the index pairs of **observed** edges.
- Unobserved set, $\mathbf{A}_l^{(u)}(t)$ defined analogously.
- $\mathbf{x}_j(t) \in \mathbb{R}^m$ attribute vector for node $\mathcal{N}_j$ at time t.

Joint Latent Factor Model for Co-evolution

Modeling Network and Attribute Dynamics

Model Specification:

$$\text{logit}\Big(P\big(a_{jj',l}(t)=1\big)\Big) = \underbrace{\mu(t)}_{\substack{\text{Time-varying}\\ \text{baseline}}} + \underbrace{\zeta_j(t)^T\zeta_{j'}(t)}_{\substack{\text{Shared Factor}\\ \text{(Cross-layer effects)}}} + \underbrace{\xi_{j,l}(t)^T\xi_{j',l}(t)}_{\substack{\text{Layer-specific Factor}\\ \text{(Within-layer effects)}}}$$

- Model log-odds of a link forming between nodes j and j' in layer l at time t to capture the **co-evolution of networks and attributes**.
- **Shared Factors ($\zeta_j(t) \in \mathbb{R}^{R_\zeta}$):** R_ζ-dimensional vector capturing common, cross-layer dynamics for node j.
- **Layer-Specific Factors ($\xi_{j,l}(t) \in \mathbb{R}^{R}$):** R-dimensional vector capturing unique dynamics for node j within a single layer l.

Modeling Complex "Coopetition" Dynamics

Why a Dual Latent Factor Model?

- Dual latent factor structure is designed to capture nuanced **dual** reality of terrorist networks, beyond simple "yes-or-no" classifications.

- **Shared Factors - The Ideological Bedrock (ζ):**
  - Model fundamental, long-term drivers that create basis for alliance.
  - "Common ground" such as a shared ideology, common strategic goals, or a designated common enemy.

- **Layer-Specific Factors - Operational & Contextual Friction (ξ):**
  - Model practical deviations from the ideological baseline.
  - Capture the "on-the-ground" realities, like competition for resources, territory, or influence.

Key Advantage: This structure allows us to model "coopetition" - how groups can be **allies in one domain** but **rivals in another**.

Modeling Nodal Attributes & Latent Factors

- A high edge probability between two nodes suggests they are also likely to have similar attribute vectors.
- The latent factors that drive graph connectivity also inform the evolution of nodal attributes.

$$\underbrace{x_{j,k}(t)}_{\substack{\text{Observable}\\ \text{Attribute}}} = \underbrace{\eta_k(t)}_{\substack{\text{Baseline}\\ \text{Value}}} + \underbrace{\sum_{l=1}^{L} \xi_{j,l}(t)^T \alpha_{k,l}(t)}_{\substack{\text{Effect of Node's}\\ \text{Latent Network Position}}} + \underbrace{\epsilon_{j,k}(t)}_{\substack{\text{Random}\\ \text{Noise}}}$$

- Approach leverages signals terrorists cannot fully hide (e.g., changes in operations or capacity) to reconstruct concealed relationships.
- **Core Idea:** This formula allows us to mathematically connect the invisible graph structure to visible information.

Model Characteristics

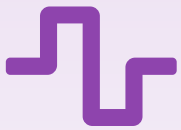
Inferring the Unseen from the Seen

- Model connects **observable signals** (e.g., propaganda, public statements, attack patterns) to the **hidden network** of relationships.
- This leverages information that groups *cannot fully conceal*, using their own actions to infer on their associations.

Capturing Emergent Alliances

- Model encodes that alliances are transitive (*a friend of a friend is often an ally*).
- This allows it to identify **emergent clusters** and indirect ties.

Neural Network Gaussian Process Prior



To model time-varying functions, we use **Gaussian Process (GP) priors**.

- ✔ **GP Priors are placed on:** $\{\zeta_{j,r}(t)\}$, $\{\xi_{j,l,r}(t)\}$, $\{\mu(t)\}$, $\{\eta_k(t)\}$, $\{\alpha_{k,l,r}(t)\}$.
- ✔ **Innovation:** We use **deep neural network kernels** within the GPs.
- → **Benefit:** Gives the **flexibility of deep learning** and **principled uncertainty** of a GP. Closed form kernel with ReLU activation.

Prior Specifications for Time-Varying Functions

Neural Network Gaussian Processes (NN-GPs)

All unknown time-varying functions are assigned **mean-zero NN-GP priors** with F-layer deep network covariance kernel κ_F.

Function	**Prior Distribution**
$\mu(\cdot)$	$\sim$ NN-GP$(0, \kappa_F(t_i, t_j; \boldsymbol{\beta}_\mu))$
$\eta_k(\cdot)$	$\sim$ NN-GP$(0, \kappa_F(t_i, t_j; \boldsymbol{\beta}_\eta))$
$\zeta_{j,r}(\cdot)$	$\sim$ NN-GP$(0, \kappa_F(t_i, t_j; \boldsymbol{\beta}_\zeta))$
$\xi_{j,l,r}(\cdot)$	$\sim$ NN-GP$(0, \kappa_F(t_i, t_j; \boldsymbol{\beta}_\xi))$
$\alpha_{k,l,r}(\cdot)$	$\sim$ NN-GP$(0, \kappa_F(t_i, t_j; \boldsymbol{\beta}_\alpha))$

Bayesian inference: MCMC enables full uncertainty quantification for graph and attribute estimates.

Why We Use the Neural Network GP (NN-GP) Kernel

Problem with Matérn

- **Assumes Stationarity:** Fundamentally unsuited for data with abrupt changes; expects smooth evolution.
- **Parameter Unidentifiability:** Length-scale and variance are difficult to distinguish. Complex tuning.

Solution: NN-GP

- **Captures Non-Stationarity:** Naturally designed to model the sudden structural shifts we expect.
- **Simple & Identifiable:** Requires two core parameters, avoiding tuning issues.

Critical Advantage for Terrorism Analysis

NeuralNet-GP is crucial as it models the abrupt, **non-stationary patterns** in terrorism graphs that stationary kernels (like Matérn) miss.

Posterior Computation & Inference

Edge Prediction

- To draw inference on unobserved edge $a^{(u)}_{jj',l}(t_i), (j, j') \in \mathbf{U}^{(u)}_l(t_i)$, draw posterior predictive samples $a^{(u)}_{jj',l,q}(t_i)$ from $\text{Ber}(p^{(u)}_{jj',l,q}(t_i))$.
- An edge exists if its posterior probability exceeds a decision cutoff.

$$\underbrace{\frac{1}{Q}\sum_{q=1}^{Q} a^{(u)}_{jj',l,q}(t_i)}_{\text{Posterior Edge Probability}} > \underbrace{\Delta_c}_{\text{Decision Cutoff}}$$

UQ in a predicted link, allowing agencies to **prioritize investigations** based on evidence strength, not just a yes/no guess.

Posterior Computation & Inference: Inferring Attributes

Method

Sample from the posterior predictive distribution for each MCMC draw q:

$$x_{j,k} \sim \mathcal{N}\left(\eta_k^{(q)} + \sum_l \boldsymbol{\xi}_{j,l}^{(q)T} \boldsymbol{\alpha}_{k,l}^{(q)}, \sigma_k^{(q)2}\right)$$

- Point prediction of $x_{j,k}(t_i)$ obtained from mean/median of samples.
- 95% Predictive Intervals can be obtained.

Strategic Implication: Narrow prediction intervals signal high certainty for **direct action**. Wide intervals signal high uncertainty and the need for **more intelligence gathering**.

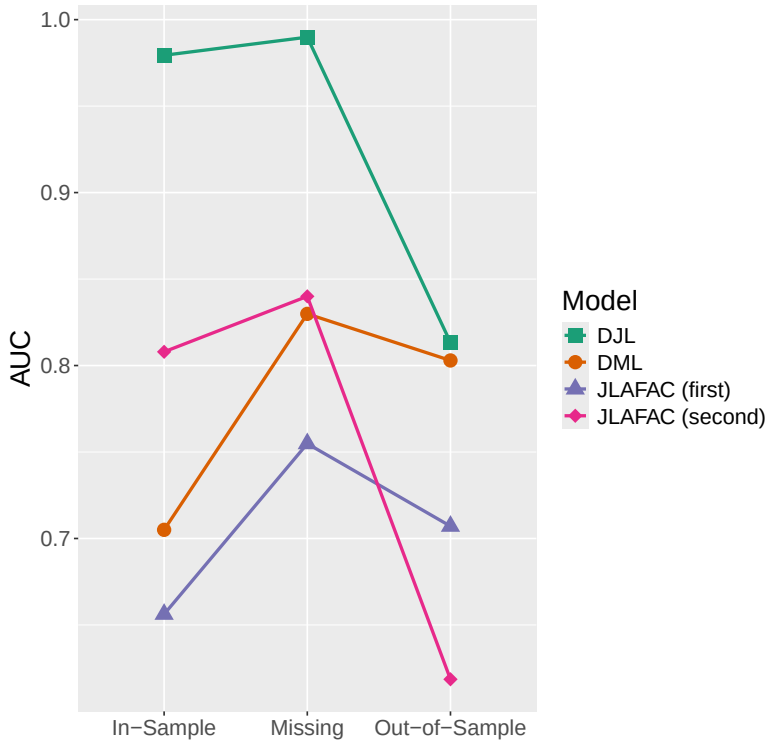
Analysis of Covert Terrorist Graph Data

- Terrorism graph data. Data Description
- $J = 15$ organizations, $L = 2$ relationship types, $T = 14$ time points, with 15th time point being held out.
- Number of nodal attributes $m = 14$.
- Analyses directly address 3 critical security questions.
  - reconstruct hidden relationships.
  - detect organizational transformations.
  - predict cascading effects within the terrorist ecosystem.

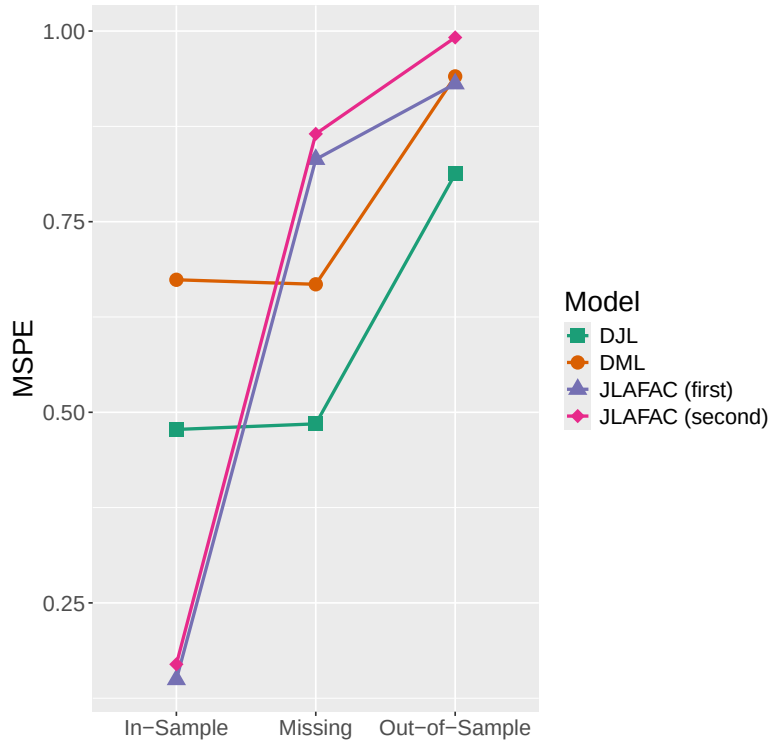
Security Question 1: Graph Reconstruction & Prediction of Organizational Attributes

Prediction Performance for Edges & Nodal Attributes

Edge Prediction



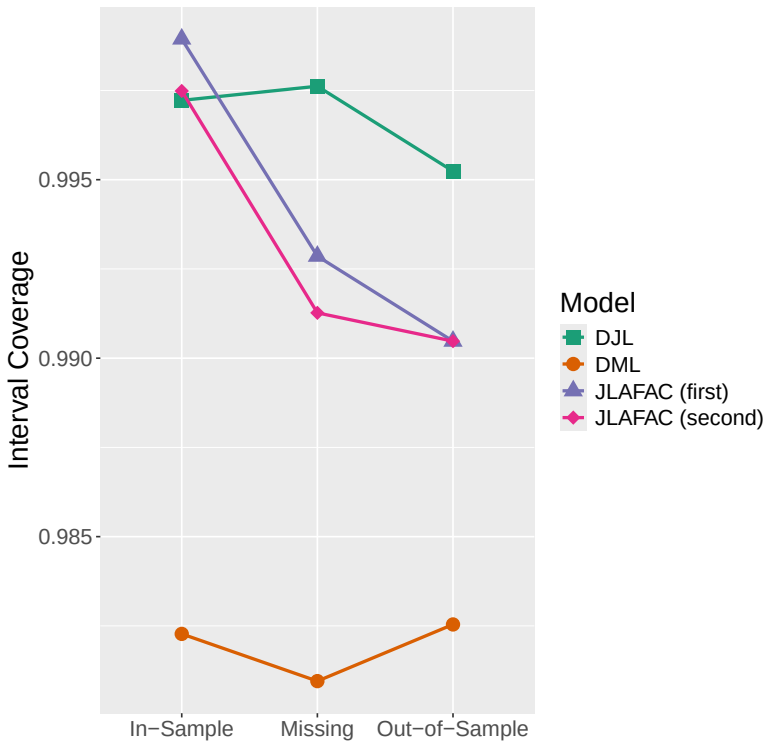
Nodal Attribute Prediction



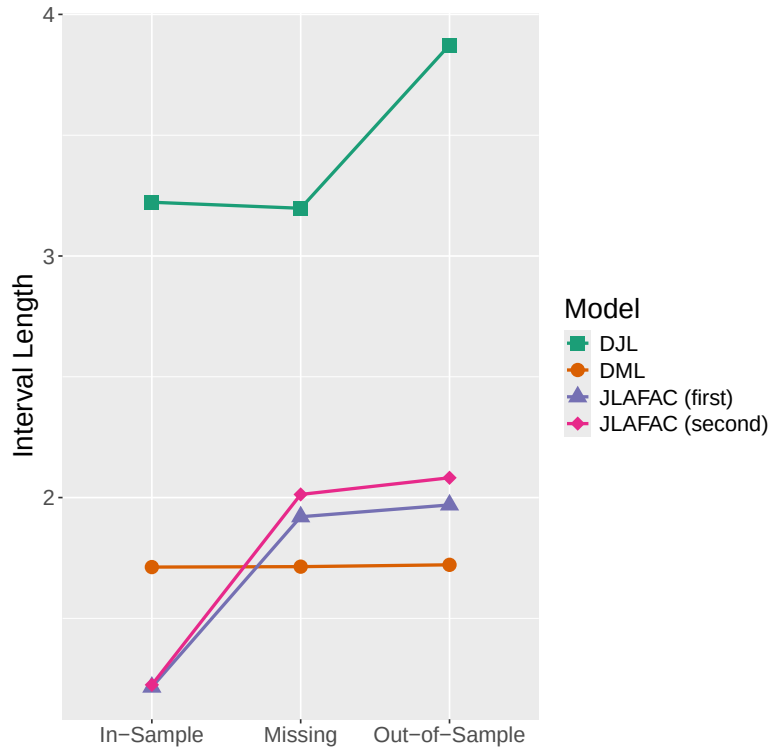
Security Question 1: Graph Reconstruction & Prediction of Organizational Attributes

Coverage and Length of 95% PIs for Nodal Attributes

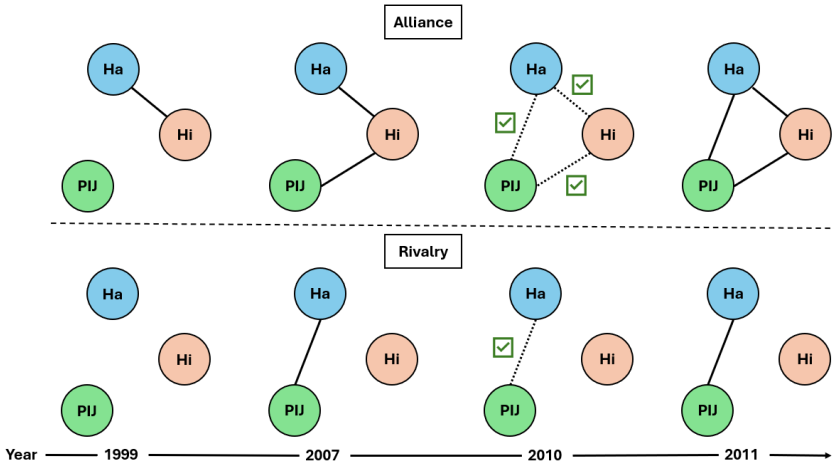
95% PI Coverage



95% PI Length



Security Question 1: Predicting Terrorist Group Dynamics



Evolving relations: Hizballah, Hamas, & PIJ.

Historical Validation

Predictions align with some major events like the Second Intifada (2000-2005) and Gaza Civil War (2007).

Model's Predictive Successes

- **Future Alliances:** Forecasted the Hizballah-PIJ alliance in future years, despite no initial ties (1998-99).
- **Persistent Alliances:** Correctly forecast ongoing Hizballah-Hamas alliance in 2010.
- **Nuanced Duality:** Identified simultaneous **alliance-rivalry** between Hamas and PIJ, reflecting their complex history.

Security Question 2: Early Warning Through Organizational Transformation

Evolution of the FARC (Revolutionary Armed Forces of Colombia) - ELN (National Liberation Army) Relationship in Colombia

1998: Alliance $\longrightarrow$ **1999: Alliance + Rivalry**

(Joint enemy) *(Triggered by exclusive peace talks with FARC)*

Model Prediction for Missing Year 2005

Our model (DJL) correctly predicted this complex "coopetition" state for the unobserved year 2005, by which the dual nature of their relationship had solidified.

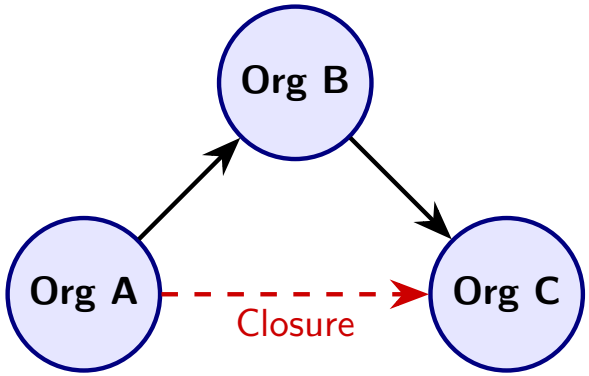
- This demonstrates the model's ability to identify and forecast intricate, non-obvious relationship dynamics.

Security Question 2: Strategic Implications

- Can we achieve early detection of organizational transformations?
- **Mechanism:** Joint analysis of graph evolution and organizational attributes. How external shocks (like peace talks) alter both group behavior and their relationship dynamics.
- **Evidence:** This predictive power is confirmed by Mean Squared Prediction Error (MSPE) values below 0.5 for attributes in missing data scenarios.

Intelligence Advantage - Proactive Lead Time: The framework detected transformation patterns **years** before they fully manifested, offering months or even years of lead time before organizational changes result in operational escalation.

Security Question 3: Mechanism of Cascading Effects



Question

Changes in one/few organizations cascade to others?

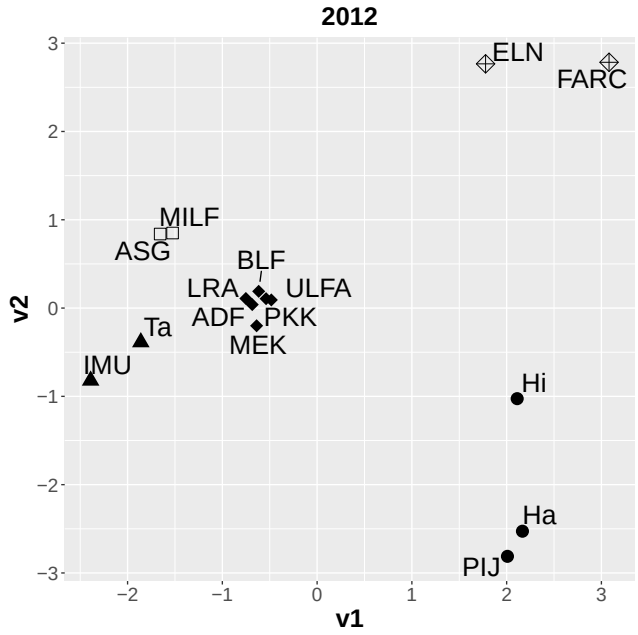
Finding

The latent position plots reveal **dynamic community structures** over time.

- **Transitivity** fosters clusters, enabling diffusion of tactics & ideology.
- **Implication:** Clusters drive **cascading effects**, directly addressing how influence spreads across organizations.
- **Approach:** Analyze temporal evolution of **latent positions**.
- **Visualization:** Resolve rotational ambiguity via **Procrustean transformations** on MCMC samples; visualize with the first two principal components.

Security Question 3: Mechanism of Cascading Effects

The model successfully identified distinct clusters of terrorist organizations that align with known geopolitical contexts.

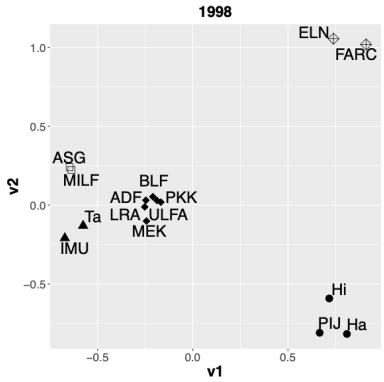


Interpretable Geopolitical Clusters Emerged

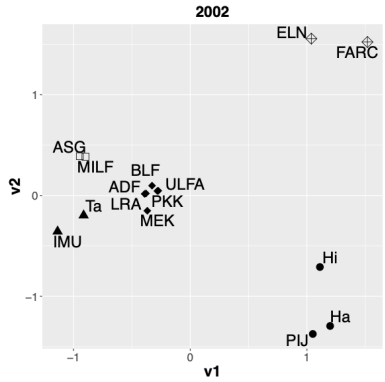
- **Palestinian Conflict:** Hizballah, Hamas, PIJ
- **Colombian Conflict:** ELN, FARC
- **Central Asian Front:** Taliban, IMU
- **Filipino Separatists:** MILF, ASG
- **Inter-Cluster Proximity:** Model shows stronger latent ties between geographically closer clusters (e.g., Central Asian & Filipino).

Thus transitivity reveals pathways for cascading effects across the network.

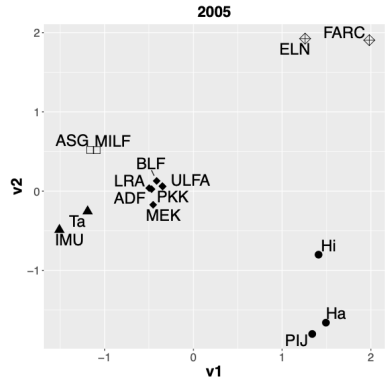
Security Question 3: Mechanism of Cascading Effects



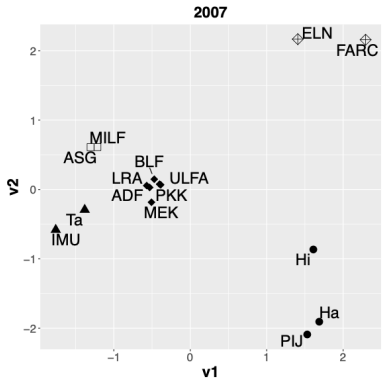
(a) 1998



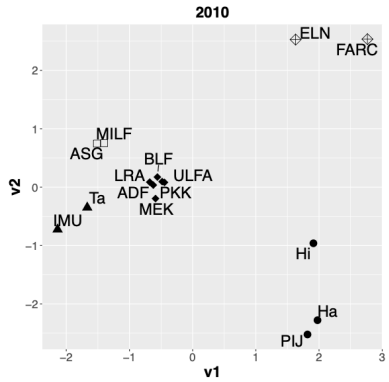
(b) 2002



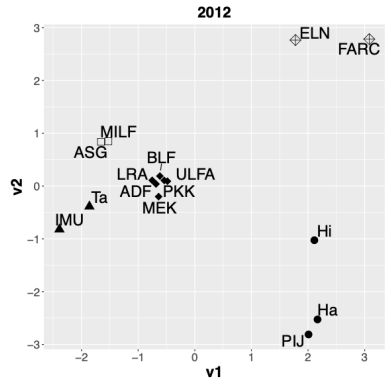
(c) 2005



(d) 2007



(e) 2010



(f) 2012

Emerged Geopolitical Clusters

To Sum Up...

- Joint modeling addresses critical intelligence questions.
- Accurate attribute and edge prediction (graph reconstruction).
- Principled UQ for reliable intelligence decision-making. Combines deep-learning predictive power with calibrated uncertainty.
- Identification of clusters reflecting actual geopolitical reality.
- Future Work : Scale and extend to larger networks using distributed inference.

Reference (Please scan for Preprint)

Thank you!