Kernels for Link Prediction
with Latent Feature Models

Canh Hao Nguyen & Hiroshi Mamitsuka

Bioinformatics Center, Institute of Chemical Research
Kyoto University
{canhhao,mami}@kuicr.kyoto-u.ac.jp

Outline

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Link Prediction on Networks
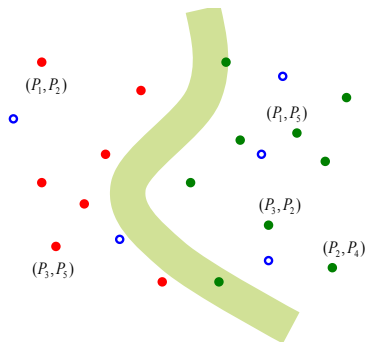Network Structures

Link Prediction on Networks

- A formal ways of representing relations between entities.
- Found in social networks, Collaborative Filtering, Systems Biology.
- A link is a relation between two entities.
- Problem: predicting new links given a networks.
- Information used: nodes' information or network structures.
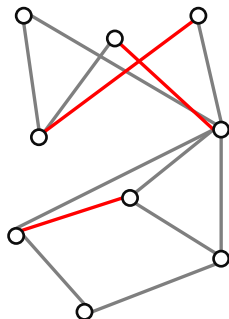
Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Link Prediction on Networks
Network Structures

Assumption of Link Prediction

Fundamental assumption:
independence and identical
distribution of links.

BUT...

Background
Latent Feature Models
Link Kernels
Experiments
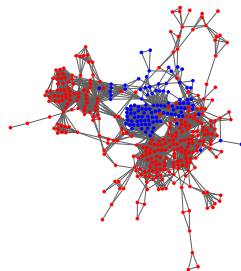Conclusion

Link Prediction on Networks
Network Structures

Assumption of Link Prediction
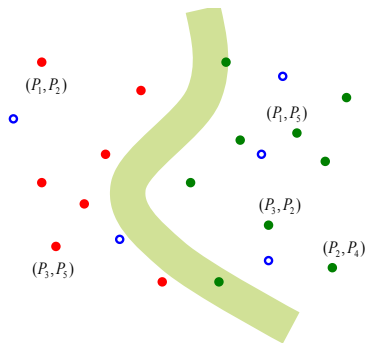
Fundamental assumption:
independence and identical
distribution of links.

BUT... networks have structures

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Link Prediction on Networks
Network Structures

Network Structures

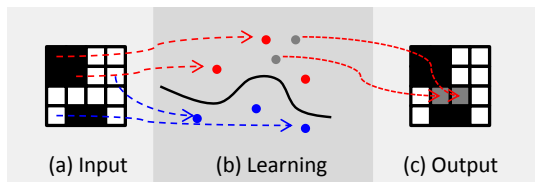Networks proved to have structures: scale-free, similarity, bipartite, clique, motif ...

- Similarity networks.

- Bipartite networks.

- Latent feature model based networks.

- Social networks, metabolic networks...

- Collaborative Filtering, protein-ligand bindings...

- Social networks, PPI, GR networks...

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Link Prediction on Networks
Network Structures

Objective

To model network structures with latent features

- To derive latent feature models for non-similarity networks.
- To approximate the model using kernel frameworks.
- To learn the model optimally and efficiently.

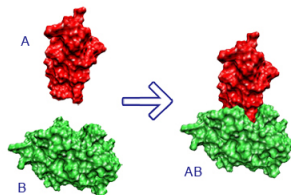and apply to the link prediction problem in biological networks.

(a) Input   (b) Learning   (c) Output

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Motivation
Latent Feature Models

Outline

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Motivation
Latent Feature Models

Biological Motivation

- Protein protein interactions are based on domain domain interactions.
- Protein docking is based on shape complementarity.
- These features (domain, shape) determine PPI ability.

Background
Latent Feature Models
Link Kernels
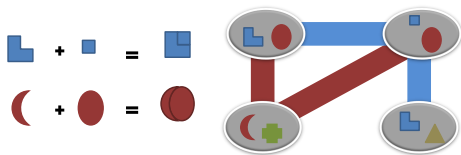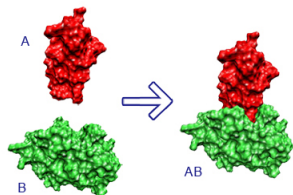Experiments
Conclusion

Motivation
Latent Feature Models
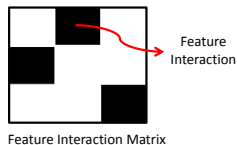
Biological Motivation

- Protein protein interactions are based on domain domain interactions.
- Protein docking is based on shape complementarity.
- These features (domain, shape) determine PPI ability.

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Motivation
Latent Feature Models

Latent Feature Models of Links

- A node in the graph contains some (latent) features (F).
- Some pairs of features interact (W).
- The nodes contain that pairs link to each other.

$$A = F \times W \times F^T$$

Latent Feature Vector

Latent Feature Matrix

Feature Interaction

Feature Interaction Matrix

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Motivation
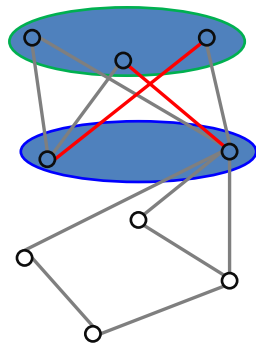**Latent Feature Models**
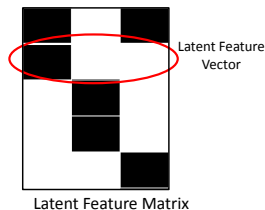
Inferring the Models

given the adjacency matrix A.

- Not globally optimal.
- Using Indian Buffet Processes (IBP)
- ... and Gibbs sampling.
- Sampling F from some distribution.
- Inferring W and repeat the process again.
- Very time consuming.
- Scalable to small graphs (< 500 nodes).

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

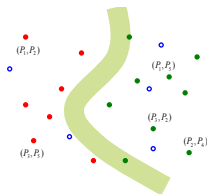Kernels from Latent Feature Models
Node Kernels
Link Kernels

Outline

1. Background

2. Latent Feature Models

3. Link Kernels
   - Kernels from Latent Feature Models
   - Node Kernels
   - Link Kernels

4. Experiments

5. Conclusion

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Kernels from Latent Feature Models
Node Kernels
Link Kernels

Latent Feature Models in Kernels

Kernels to encode similarity

- High similarity on the same class.
- Low similarity otherwise.

Link Similarity with latent features.

- Links are similar if nodes are similar.
- Nodes are similar if they share many latent features.

Ideal Kernels

for any diagonal, nonnegative $D \in R^{d \times d}$,

$$K^*(D) = F \times D \times F^T$$

Any node kernels should be close to Ideal Kernels

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Kernels from Latent Feature Models
Node Kernels
Link Kernels

Latent Feature Models in Kernels

Kernels to encode similarity

- High similarity on the same class.
- Low similarity otherwise.

Link Similarity with latent features.

- Links are similar if nodes are similar.
- Nodes are similar if they share many latent features.

Ideal Kernels
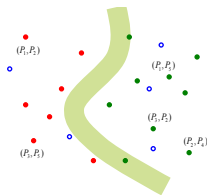
for any diagonal, nonnegative $D \in R^{d \times d}$,

$$K^*(D) = F \times D \times F^T$$

Any node kernels should be close to Ideal Kernels

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Kernels from Latent Feature Models
Node Kernels
Link Kernels

Latent Feature Models in Kernels

Kernels to encode similarity

- High similarity on the same class.
- Low similarity otherwise.

Link Similarity with latent features.

- Links are similar if nodes are similar.
- Nodes are similar if they share many latent features.

Ideal Kernels
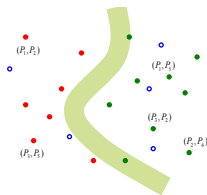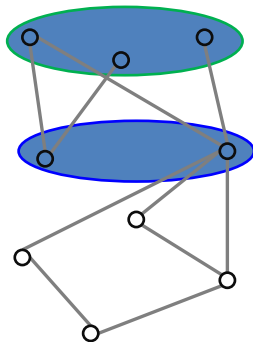
for any diagonal, nonnegative $D \in R^{d \times d}$,

$$K^*(D) = F \times D \times F^T$$

Any node kernels should be close to Ideal Kernels

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Kernels from Latent Feature Models
Node Kernels
Link Kernels

Node Kernels

- Idea: nodes with more common latent features are more similar.

- Observation: nodes with more common latent features have similar connectivity patterns.

- We use connectivity pattern to define node similarity.

- $K_n = norm(A^2)$ (diagonally normalized).

- In sparse networks, this kernel K_n approximates ideal kernels.

- PPI networks are sparse ($> 80\%$ nodes in yeast, all in fruit-fly are of degree ≤ 10).

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Kernels from Latent Feature Models
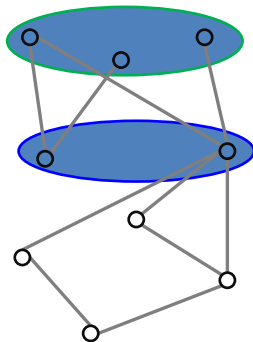**Node Kernels**
Link Kernels

Node Kernels

- Idea: nodes with more common latent features are more similar.

- Observation: nodes with more common latent features have similar connectivity patterns.

- We use connectivity pattern to define node similarity.

- $K_n = norm(A^2)$ (diagonally normalized).

- In sparse networks, this kernel K_n approximates ideal kernels.

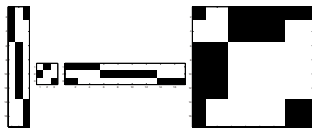- PPI networks are sparse ($> 80\%$ nodes in yeast, all in fruit-fly are of degree ≤ 10).

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Kernels from Latent Feature Models
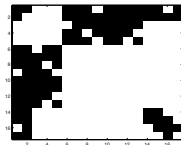Node Kernels
Link Kernels

Kernels on Links

Given node similarity are encoded in node kernel K_n.

- We use pairwise kernels to encode link similarity.
- If corresponding nodes are similar, links are similar.
- $K(\{a, b\}, \{c, d\}) = K_n(a, c) * K_n(b, d) + K_n(a, d) * K_n(b, c)$
- Feature-wise: if a_i, b_j are feature of a and b
  $a_i b_j + a_j b_i$ is a feature of $\{a, b\}$
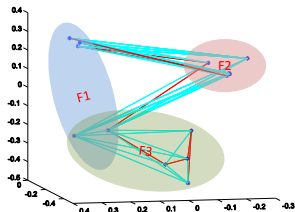
Background
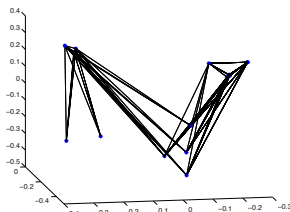Latent Feature Models
Link Kernels
Experiments
Conclusion

Kernels from Latent Feature Models
Node Kernels
Link Kernels

Demonstration

Network model

Incomplete observation

Nodes and Links

Nonlinks class

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Kernels from Latent Feature Models
Node Kernels
Link Kernels

Properties (K_n similar to ideal kernels)

- Positivity: $(K_n)_{ij}$ for a pair of nodes (i, j) is positive if they share a common feature.
  if there is a shared feature according to the generative model, K_n should recognize it

- Monotonicity: the more a pair of nodes share features, the higher the kernel value is.
  K_n respects the monotonicity as an ideal kernels: more features, higher kernel values

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Kernels from Latent Feature Models
Node Kernels
Link Kernels

Properties (K_n similar to ideal kernels) cont.

Recall that $K_n = FDF^T = F(WEW^T)F^T$, ($E = F^T F$), ideal kernels $K^*(\hat{D}) = F\hat{D}F^T$.

- Ideal Condition: K_n is an ideal kernel iff WF^T has all row vectors uncorrelated.
- On Sparse Networks: If the model is sparse in the sense that $(\sum_{i \neq j} |W_{il} W_{jk} E_{lk}|^p)^{\frac{1}{p}} \leq \delta$ then K_n is close to an ideal kernels in the sense that

$$\| D - \hat{D} \|_p \leq \delta$$

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Latent Feature Assumption
Time Complexity
Link Prediction Results

Outline

1. Background

2. Latent Feature Models

3. Link Kernels

4. Experiments
   - Latent Feature Assumption
   - Time Complexity
   - Link Prediction Results

5. Conclusion

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Latent Feature Assumption
Time Complexity
Link Prediction Results

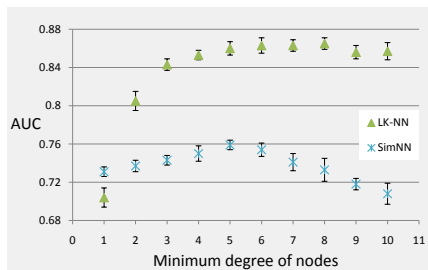# Experiments and Data

- Input: adjacency matrix A.
- 90/10 train/test splits.
- Learn SVMs on the kernels.
- Show: running times, assumption comparison, prediction AUCs on PPI networks.
- Compare to using nodes' attributes: sequence kernels.

- PPI networks of yeast and fruit fly: the largest networks available.
- Extracted from DIP databases (hand-curated, only direct interactions).
- We used only the connected part of the networks.
- We filtered out the nodes with degrees less than m.
- For $m = 1$, network sizes: 4762 proteins in yeast and 6644 proteins in fruit fly.

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Latent Feature Assumption
Time Complexity
Link Prediction Results

## On the Latent Feature Model Assumpsion

Reasonability of the latent feature assumption as opposed to the similarity assumption: nearest neighbor classifier.

Latent feature assumption is more suitable than than similarity.

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Latent Feature Assumption
Time Complexity
Link Prediction Results
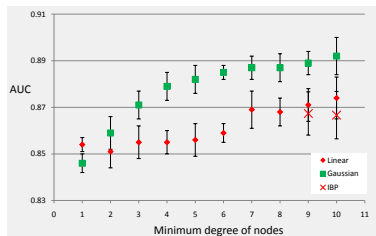
Execution Time of Our Method vs. IBP

Our method scales to this size of data as opposed to explicit feature generation methods.
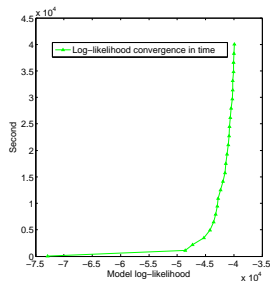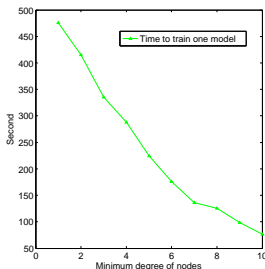
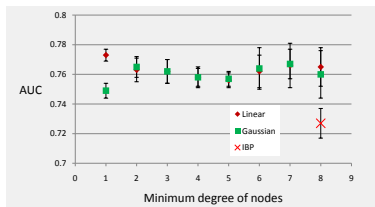Our method on all subnetworks   IBP on the smallest one.

Our method takes minutes on this sizes of data. IBP only scales to the smallest ones in many hours.

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Latent Feature Assumption
Time Complexity
Link Prediction Results

Link Prediction Performance on yeast PPI networks

- Our method gives very high performance compared to IBP.
- The results are significantly higher than random, meaning that networks have significant structures.
- Using sequence (spectrum kernels) to predict links in these networks: 0.71 ± 0.008.

Background
Latent Feature Models
Link Kernels
Experiments
Conclusion

Latent Feature Assumption
Time Complexity
Link Prediction Results

Link Prediction Performance on fruit fly PPI networks

- Our method gives higher result than IBP.
- The networks have significant structures.
- Using sequence (spectrum kernels) to predict links in these networks: 0.65 ± 0.016.

Conclusion

- We argued to use network structures to predict links.
- We proposed a latent feature model to describe the network structures.
- We used kernels to encode the model implicitly and to train it efficiently and optimally.
- Our method scaled to real data of PPI networks, unlike IBP.
- Our method gave a high performance of predicting PPI.

Our method is efficient and effective for latent feature models of networks

Conclusion

- We argued to use network structures to predict links.
- We proposed a latent feature model to describe the network structures.
- We used kernels to encode the model implicitly and to train it efficiently and optimally.
- Our method scaled to real data of PPI networks, unlike IBP.
- Our method gave a high performance of predicting PPI.

Our method is efficient and effective for latent feature models of networks