Efficient Distributed Grounding in Markov Logic Networks

Kaustubh Beedkar and Rainer Gemulla

Background

Which citations refer to the same publication ?

	Citation 1	Citation 2	Citation 3
Author	Richardson Matt and Domingos Pedro	M. Richardon and P. Domingos	Domingos, Pedro and Richardson, Mathhew
Title	Markov Logic Networks	Markov Logic Networks	Markov Logic: A Unifying Framework for Statistical Relational Learning
Year	2006	2006	2007

✓ If authors for the citations are same
✓ If citation 1 and 2 are same and 2 and 3 are same, then 1 and 3 are same
✓ …

Markov Logic Networks

Combines first order logic and probability theory

Predicates
HasAuthor(title, citation) *e.g., HasAuthor('Domingos', C1)*
SameCitation(citation, citation) *Semantic equality of citations*

Formulas are uncertain

Weight	***Formulas***
2.4	HasAuthor(t, c1) ^ HasAuthor(t, c2) => SameCitation(c1, c2)
2.8	SameCitation(c1, c2) ^ SameCitation(c2, c3) => SameCitation(c1, c3)

First order logic formulas

Inference in MLN

MLN + Data → Grounding → Markov network → Inference

Motivation

MLN + Data → Grounding → Factor graph ⇐ Inference

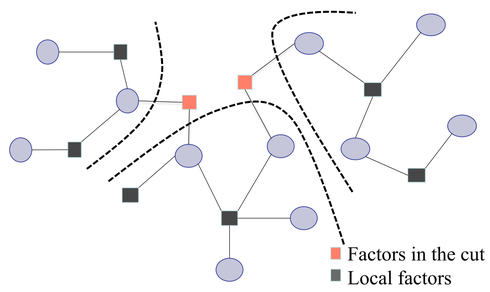
Challenges
- Grounding is expensive
- Models are large (Large factor graphs)
  - Link prediction
  - Entity resolution
  - Information extraction
- Inference is difficult

Explore parallel/distributed approaches!

Naïve Approach

MLN + Data → Grounding → Factor graph → Graph partitioning → Partitioned factor graph ⇐ Distributed Inference

Factor graph partitioning

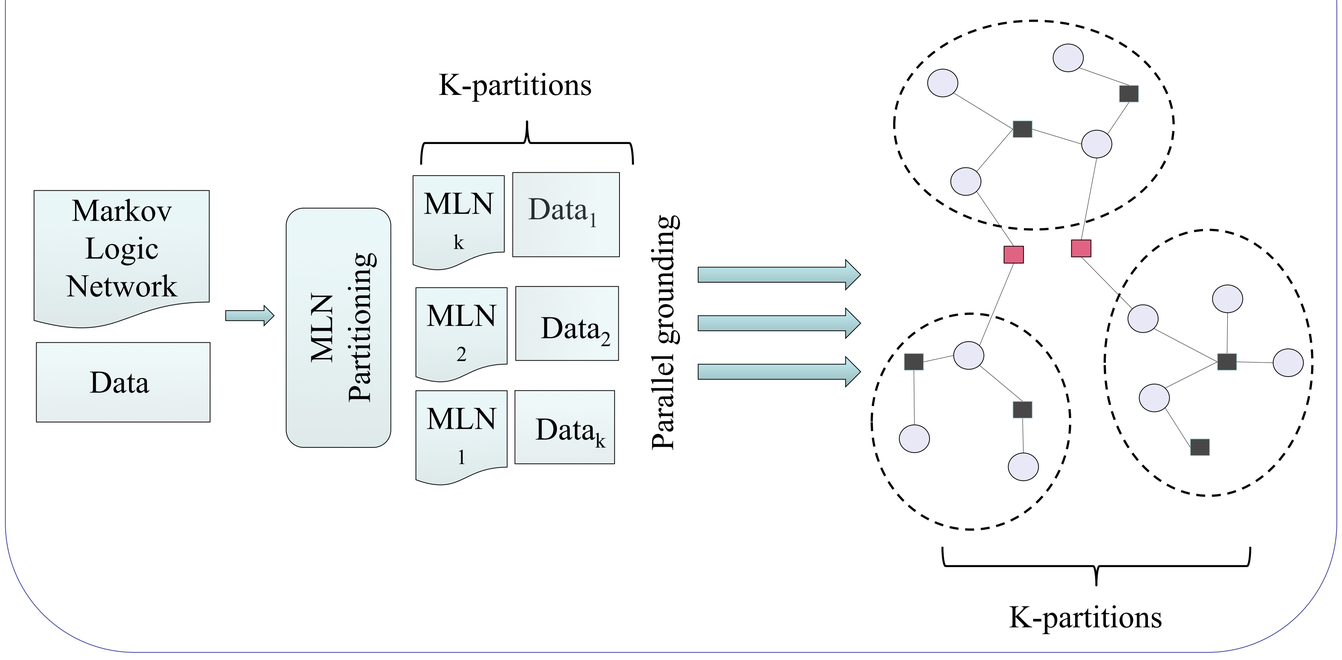


- Partitioning facilitates parallel/ distributed inference
- Computationally hard for large graphs

Proposed Approach

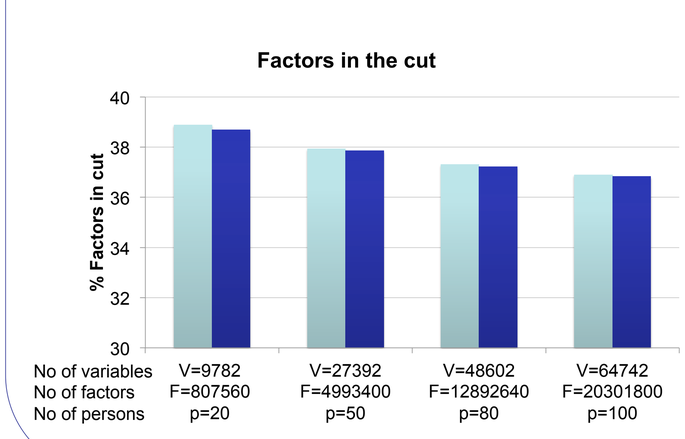
MLN + Data → MLN partitioning → Partitioned MLN → Parallel/Distributed Grounding → Partitioned factor graph ⇐ Distributed Inference

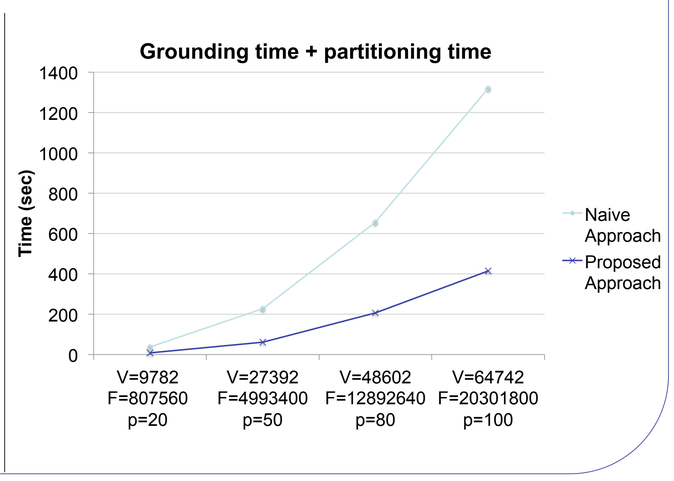
Partition markov logic and ground each partition to a partitioned factor graph.



Experimental Results

- ❖ Basic Idea – Partition at semantic level
  Ground fragments in parallel
- ❖ Partitioning at Rule level
  - Model MLN as a join graph
  - Estimate join sizes and encode join conditions
  - Co-partition relations such that it maximizes Σ(Join size) in co -partitions – optimization problem
  - **Encode as an ILP**





max planck institut informatik
UNIVERSITÄT DES SAARLANDES
International Max Planck Research School for Computer Science