Graph Analytical Re-ranking for Entity Search

Takahiro Komamizu

Nagoya University
Japan

Recall@k (k=10, 100, 1000)

Model	Total		
	@10	@100	@1000
BM25	.1823	.5175	.7703
PRMS	.2522	.5919	.8009
MLM-all	.2571	.6136	.8009
LM	.2607	.6413	.8009
SDM	.2659	.6674	.8633
LM-ELR	.2646	.6483	.8006
SDM-ELR	.2739	.6782	.8633
MLM-CA	.2639	.6370	.8329
BM25-CA	.2782	.6727	**.8708**
FSDM	.2812	.6667	.8455
BM25F-CA	.2811	**.6912**	.8653
FSDM-ELR	**.2872**	.6765	.8450
max	.2872	.6912	.8708
gap	.5836	.1796	—

- DBpedia-Entity v2
- Recall@k
  $$= \frac{\#Relevant\ items@k}{\#Relevant\ items}$$
- Gaps
  - @10, 58% drops from @1000
  - @100, 18% drops from @1000
- Missing 13% in top-1000

Research Overview

- Obj.1: Improvement of ranking
  - To fill the gaps
  - To try graph analytical approach
    - ➢ Q1 "Do graph analytical approaches improve ranking?"
- Obj. 2: Investigation for non-perfect recall
  - @1000 miss more than 15% in harder tasks
  - To improve in the future researches
    - ➢ Q2 "How far query terms from relevant entities?"

Task	SemSearch ES	INEX-LD	ListSearch	QALD-2	Total
max	.9865	.8603	.8431	.8164	.8708

Current State of Entity Search

- Fielded document model
  - Entity has 1000 fields and 3 special fields
    - 1000 of most frequent predicates in DBpedia
    - specials: name, types, contents
      - contents: contents of fields in neighbor entities
- Many approaches
  - BM25, BM25-CA, LM, SDM, PRMS, MLM-all
  - Fielded extensions: MLM-CA, FSDM, BM25F-CA
  - With entity linking: LM-ELR, SDM-ELR, FSDM-ELR

1. Text-based matching and ranking models
2. Literals of entities at most 1-hop away are taken into account

Improving Ranking by Re-ranking

- Objective:
  improve top-10 and top-100 rankings
- Approach: re-ranking top-1000
  - Top-1000 includes more than 80% results
  - Graph analytical approaches (e.g., PageRank)
- Naive approaches are quite bad
  - PageRank and personalized PageRank

Re-ranking method	SemSearch ES		INEX-LD		ListSearch		QALD-2		Total	
	@10	@100	@10	@100	@10	@100	@10	@100	@10	@100
PageRank	.1545	.4664	.1171	.3639	.1059	.4438	.1561	.4519	.1344	.4198
Personalized PageRank	.1632	.4779	.1228	.3822	.1146	.4524	.1613	.4587	.1397	.4355

PPRSD: score distribution

(**P**ersonalized **P**age**R**ank based **S**core **D**istribution)

- Observation: existing approaches already have good results (scores for ranking).
- Approach: distribute the scores in a personalized PageRank manner

$$\mathbf{pprsd}_q = (1-d)\cdot \mathbf{pprsd}_q \underline{A} + d \cdot \underline{\mathbf{t}}$$

Adjacency matrix of an induced subgraph for top-1000 entities

Score vector of top-1000 entities in a method

Obj.1: Improvement of ranking

PPRSD (*-ed) improves ranking? – Yes, both recall and NDCG

Red: improved Blue: degraded

Recall

Model	Total @10	Total @100	Model	Total @10	Total @100
BM25	.1823	.5175	SDM-ELR	.2739	.6782
BM25*	.1983	.5466	SDM-ELR*	.2749	.6786
imp.	+7.57%	+4.70%	imp.	+0.37%	+0.06%
PRMS	.2522	.5919	MLM-CA	.2639	.6370
PRMS*	.2522	.5919	MLM-CA*	.2639	.6371
imp.	0.00%	0.00%	imp.	0.00%	+0.02%
MLM-all	.2571	.6136	BM25-CA	.2782	.6727
MLM-all*	.2571	.6136	BM25-CA*	.2826	.6795
imp.	0.00%	0.00%	imp.	+1.33%	+0.97%
LM	.2607	.6413	FSDM	.2812	.6667
LM*	.2607	.6410	FSDM*	.2813	.6671
imp.	0.00%	-0.03%	imp.	+0.04%	+0.06%
SDM	.2659	.6674	BM25F-CA	.2811	.6912
SDM*	.2671	.6684	BM25F-CA*	.2865	**.6963**
imp.	+0.41%	+0.13%	imp.	+1.99%	+0.72%
LM-ELR	.2646	.6483	FSDM-ELR	.2872	.6765
LM-ELR*	.2646	.6473	FSDM-ELR*	**.2873**	.6769
imp.	0.00%	-0.12%	imp.	+0.03%	+0.06%

NDCG

Model	Total @10	Total @100	Model	Total @10	Total @100
BM25	.2558	.3582	SDM-ELR	.4261	.5211
BM25*	.2812	.3847	SDM-ELR*	.4271	.5218
imp.	+9.93%	+7.40%	imp.	+0.23%	+0.13%
PRMS	.3905	.4688	MLM-CA	.4365	.5143
PRMS*	.3913	.4698	MLM-CA*	.4361	.5150
imp.	+0.20%	+0.21%	imp.	-0.09%	+0.14%
MLM-all	.4021	.4852	BM25-CA	.4399	.5329
MLM-all*	.4030	.4863	BM25-CA*	.4475	.5404
imp.	+0.22%	+0.23%	imp.	+1.73%	+1.41%
LM	.4182	.5036	FSDM	.4524	.5342
LM*	.4191	.5046	FSDM*	.4527	.5350
imp.	+0.22%	+0.20%	imp.	+0.07%	+0.15%
SDM	.4185	.5143	BM25F-CA	.4605	.5505
SDM*	.4191	.5152	BM25F-CA*	**.4673**	**.5581**
imp.	+0.14%	+0.17%	imp.	+1.48%	+1.38%
LM-ELR	.4230	.5093	FSDM-ELR	.4590	.5408
LM-ELR*	.4240	.5103	FSDM-ELR*	.4587	.5416
imp.	+0.24%	+0.20%	imp.	-0.07%	+0.15%

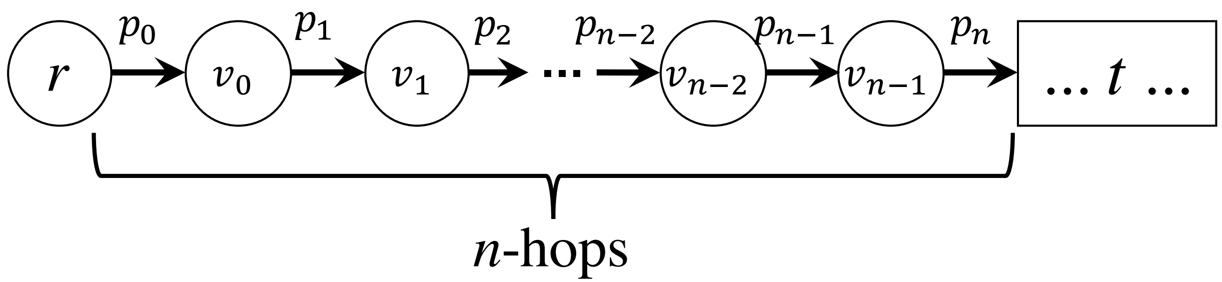
Research Overview

- Obj.1: Improvement of ranking
  - To fill the gaps
  - To try graph analytical approach
    - ➢ Q1 “Do graph analytical approaches improve ranking?”
    - ➢ A1 “Yes, but still limited improvement.”
- Obj. 2: Investigation for non-perfect recall
  - @1000 miss more than 15% in harder tasks
  - To improve in the future researches
    - ➢ Q2 “How far query terms from relevant entities?”

Task	SemSearch ES	INEX-LD	ListSearch	QALD-2	Total
max	.9865	.8603	.8431	.8164	.8708

Where are relevant terms?

- Hypothesis:
  too short hops for documents of entities
  - At most 1 hop in existing works
  - Missing relevant terms in distant entities
- Investigation
  - Distance n form entity r to term t in a query



Investigation Methodology

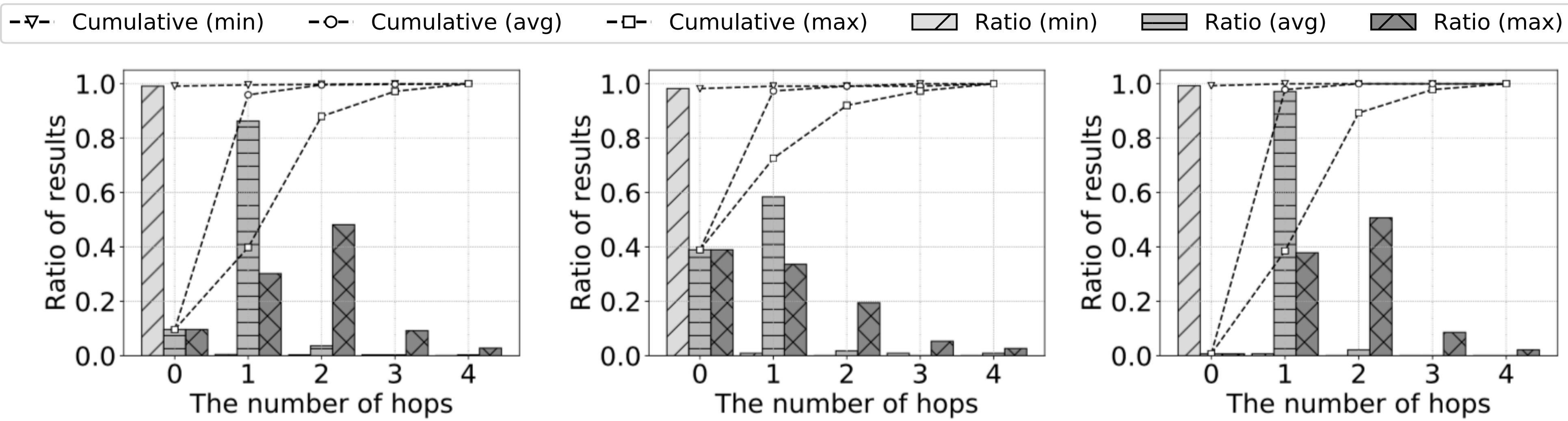
Given a query and relevant entities

1. Parse the query into terms
2. For each relevant entity r
    1. For each term t in the query
        1. Calculate minimum distance n from r to t by SPARQL queries like the following

           ```
           ASK{ ⟨r⟩ ?p0 ?v0. ?v0 ?p1 ?v1.
                ?v1 ?p2 ?l. ?l bif:contains 't'.
                FILTER isLiteral(?l).}
           ```

        2. Record n for (r, t)
3. Analyze representative distances
    - min, max, avg

Result: one hop is not enough



(a) Total (b) SemSearch ES (e) QALD-2

- Ratio of entities
  - min : at least one term is in a hop
  - avg : average number of terms are in a hop
  - max: all terms are in a hop

Discussion

Literals in longer hops should be taken into documents of entities

- Issue: noisy terms will be in the documents
  - Increasing distances for the documents explosively increases the number of reachable literals.
- Possible solutions
  - *"important"* path selection
    - Prioritization of predicates (e.g., ObjectRank)
  - Graph-based proximity
    - e.g., Random walk with restart

Conclusion

- Obj.1: Improvement of ranking
  - To fill the gaps
  - To try graph analytical approach
    - Q1 “Do graph analytical approaches improve ranking?”
    - A1 “Yes, but still limited improvement.”
- Obj. 2: Investigation for non-perfect recall
  - @1000 miss more than 15% in harder tasks
  - To improve in the future researches
    - Q2 “How far query terms from relevant entities?”
    - A2 ”More than one hop, but careful selection of paths is required.”