Revisiting Query Variants:

The Advantage of Retrieval Over Generation of Query Variants for Effective QPP

Fangzheng Tian, Debasis Ganguly, Craig Macdonald

University of Glasgow

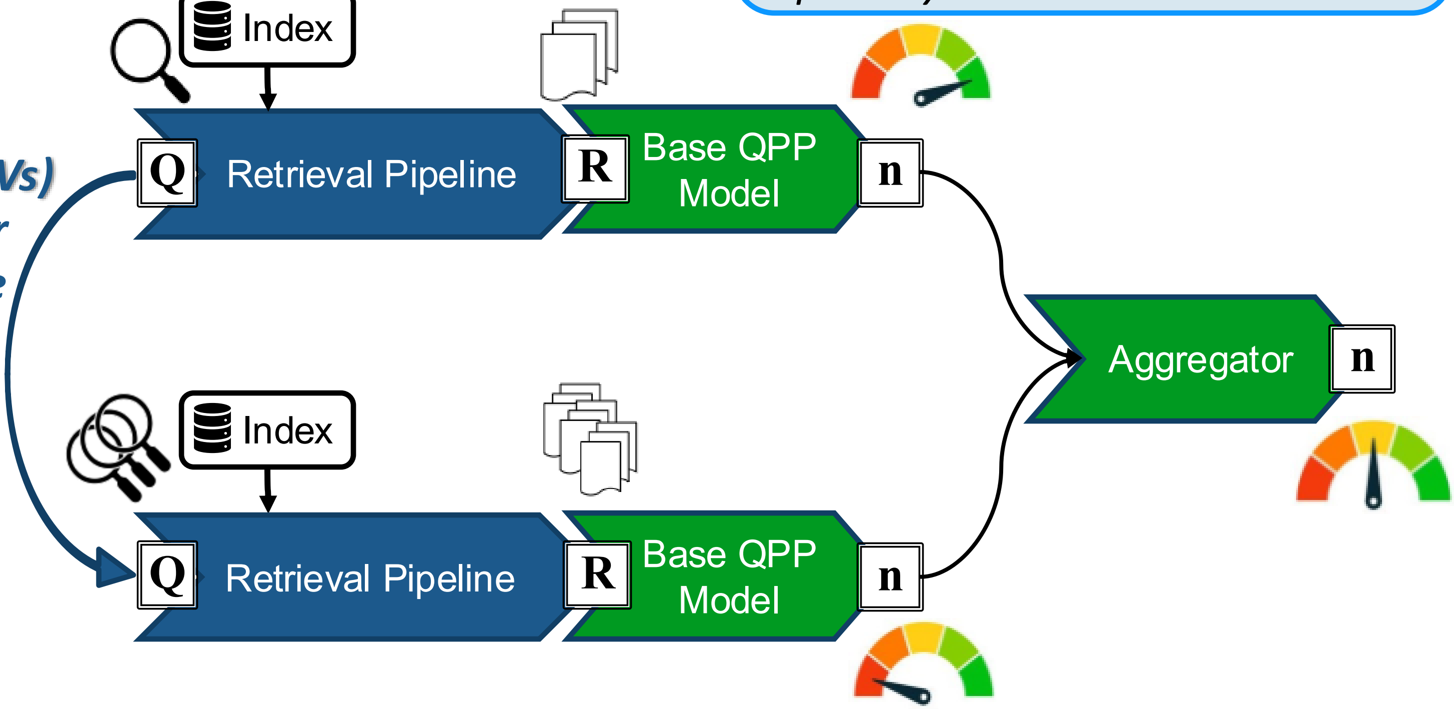
Introduction

Terrier Team — University of Glasgow

Post-retrieval QPP predicts target query's specificity based on the statistics of the retrieval scores.

QPP effectiveness this kind of methods is limited when the score distribution is unsuitable …… especially some neural rankers.

Query Variants (QVs)* *contain the similar information as the target query.



QV-based methods provide multiple observations for the prediction about the target query.

Motivation: To enhance QPP effectiveness for neural rankers with QVs.

Importance of Good QVs

QVs that are leveraged in QPP estimation should be valuable reference regarding to a target query.

Target query:
"how often to button quail lay eggs?"

Generated QVs by semantic expansion
"quail eggs pets breed"
*"lay **birds** large clutch"*
*"lay year **domestic eggs**"*

How to construct QVs for a target query?

If a query log or training set is provided, retrieve from it!

Incoherent?

Drift of topic?

Can they really enhance QPP effectiveness?

Related Work

Terrier Team

University of Glasgow

Leveraging QVs can enhance QPP effectiveness.

Reference-list-based QPP (Shtok et al., 2016)

JM Smoothing and Weighted Relative Gain can be applied in QV-based QPP as aggregation method.

Information need, query and QPP (Zendel et al., 2019), WRIG (Datta et al., 2022)

These frameworks are potential to be applied in our work.

Applying retrieved QV in a supervised QPP model.

QPP with Contextualized Representations, (Ebrahim et al., 2024)

Limited QPP effectiveness for neural rankers.

We should use the retrieved QVs to improve QPP effectiveness in an efficient manner, e.g. unsupervised method.

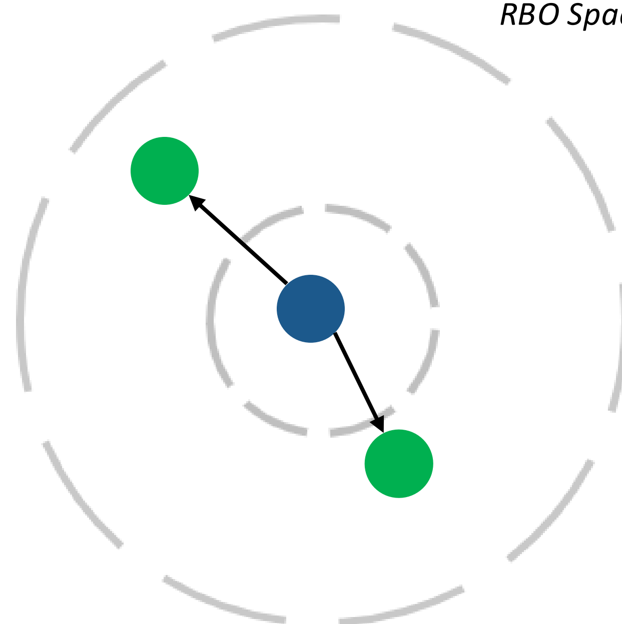
Query Retrieval Methodology

Retrieved QVs can be helpful in enhancing effectiveness of QV-based QPP method.

A single retrieval may not be able to include all the potential QVs from a query set.

1-hop Query Variants: The queries which are directly retrieved from the query.

RBO Space

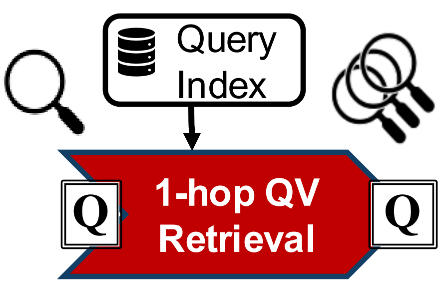


Target query:

- *"how often to button quail lay eggs?"*

1-hop QVs:

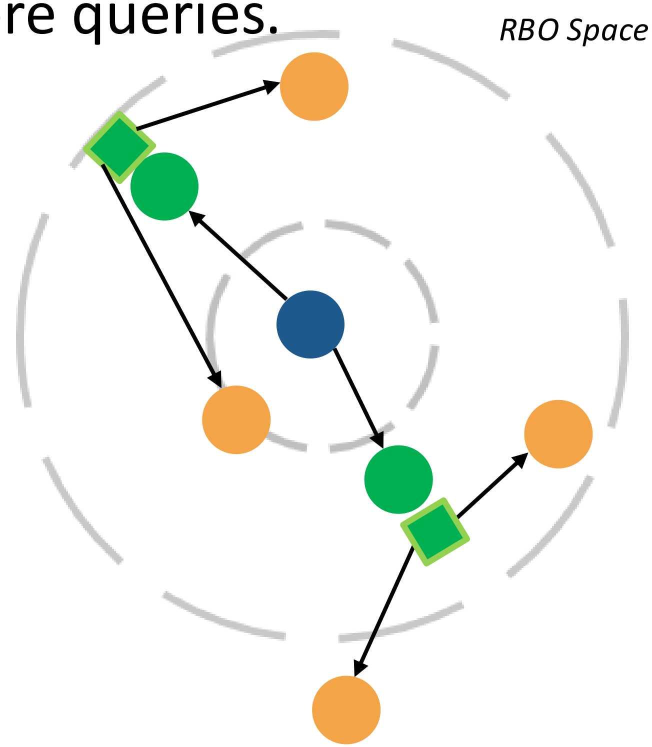
- *"how many eggs do quail lay a year?"*
- *"how long quail lay eggs?"*



Query Retrieval Methodology

Relevant information about a query can be used to represent the its information need.

2-hop Query Variants: Taking the relevant documents of the 1-hop queries as query to retrieve more queries.



Target query:

- *"how often to button quail lay eggs?"*

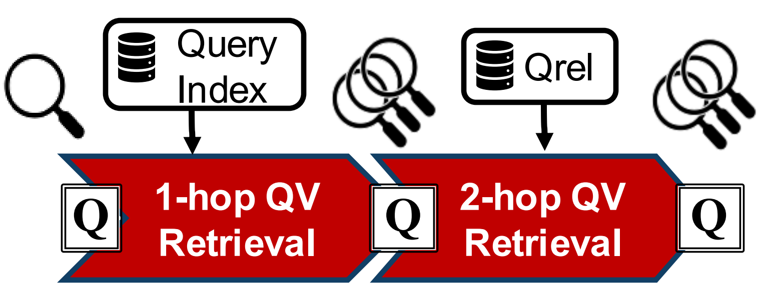
1-hop QVs:

- *"how many eggs do quail lay a year?"*
- *"how long quail lay eggs?"*

2-hop QVs:

- *"how old are quils before they lay eggs?"*
- *"when do bobwhite quails start laying?"*

……



Query Retrieval Methodology

Relevant information about a query can be used to represent the its information need.

2-hop Query Variants: Taking the relevant documents of the 1-hop queries as query to retrieve more queries.

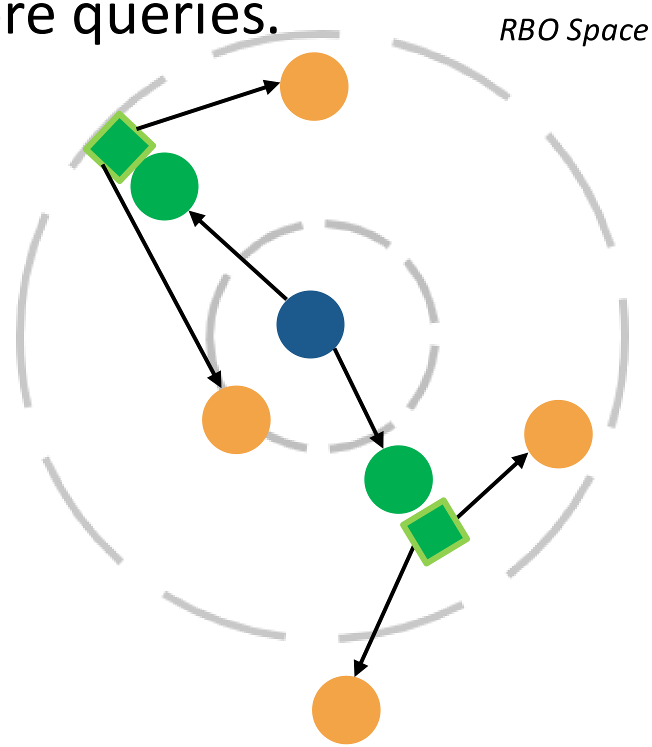
Target query:

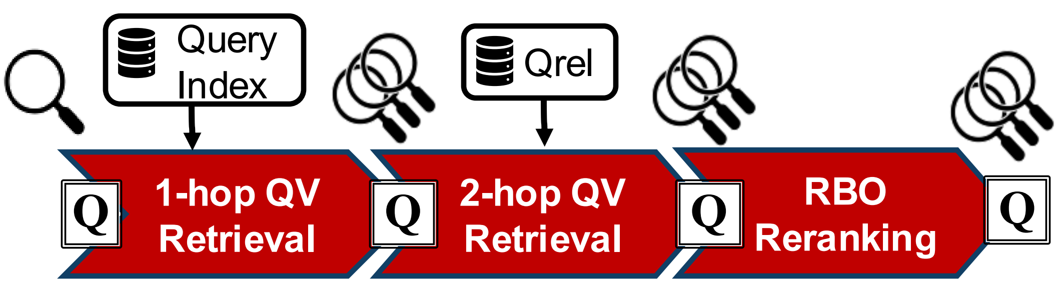
- *"how often to button quail lay eggs?"*

RBO Re-ranked QVs:

- *"how old are quils before they lay eggs?"*
- *"how many eggs do quail lay a year?"*
- *"how long quail lay eggs?"*
- *"when do bobwhite quails start laying?"*

……



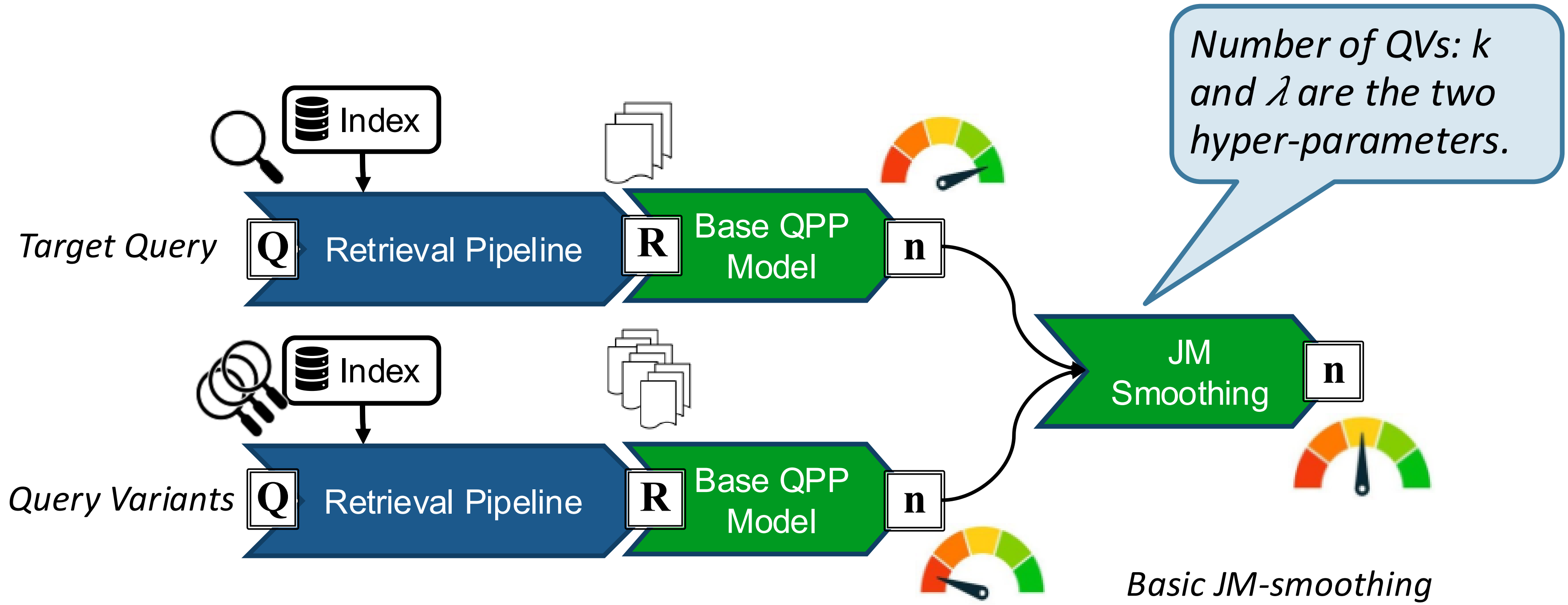


QPP based on Retrieved QVs

JM Smoothing in QV-based QPP (Zendel et al., 2019)

1. The prediction about QVs are interpolated into the final prediction with coefficient λ;

2. The contribution of each QV is in proportion with their RBO similarity to the target query.



Basic JM-smoothing Framework

QPP based on Retrieved QVs

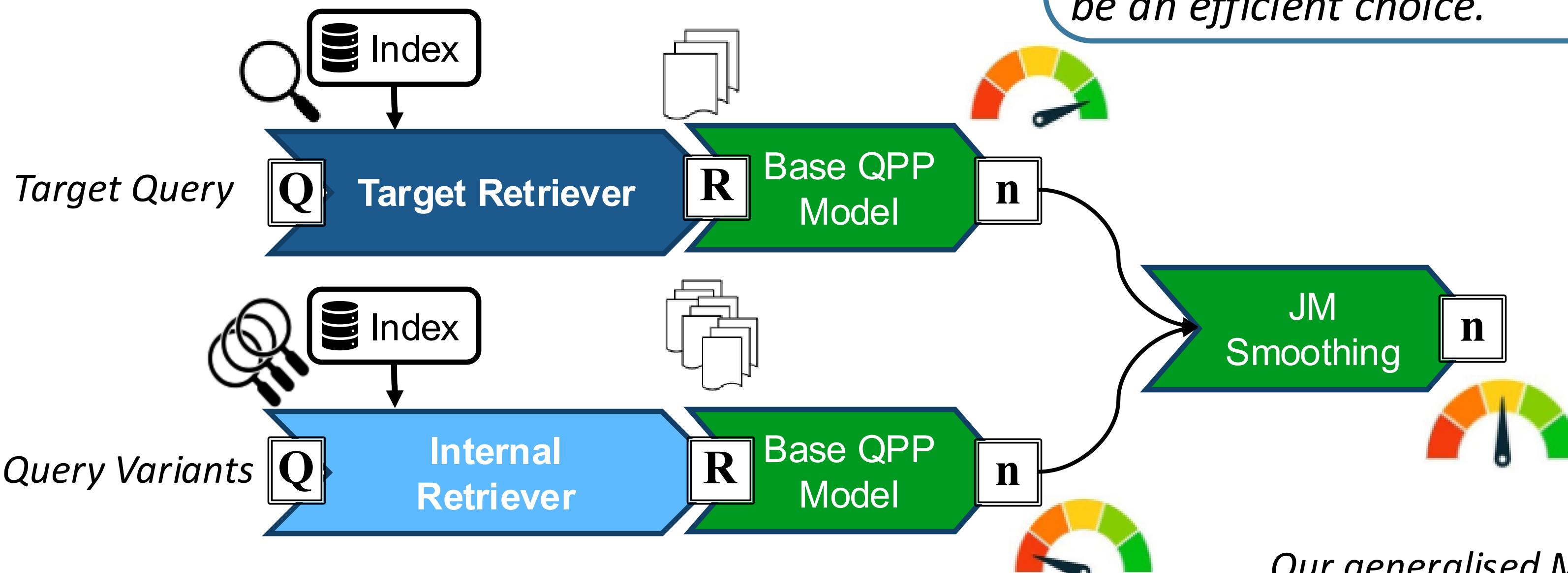
JM Smoothing in QV-based QPP (Zendel et al., 2019)

1. The prediction about QVs are interpolated into the final prediction with coefficient λ;

2. The contribution of each QV is in proportion with their RBO similarity to the target query.

We employ an internal retriever which can be separated from the target one to retrieve documents for the QVs.

When the base predictor works poorly on the target retriever, separated internal retriever can be more reliable. Besides, it can be an efficient choice.



Our generalised Model

Research Questions

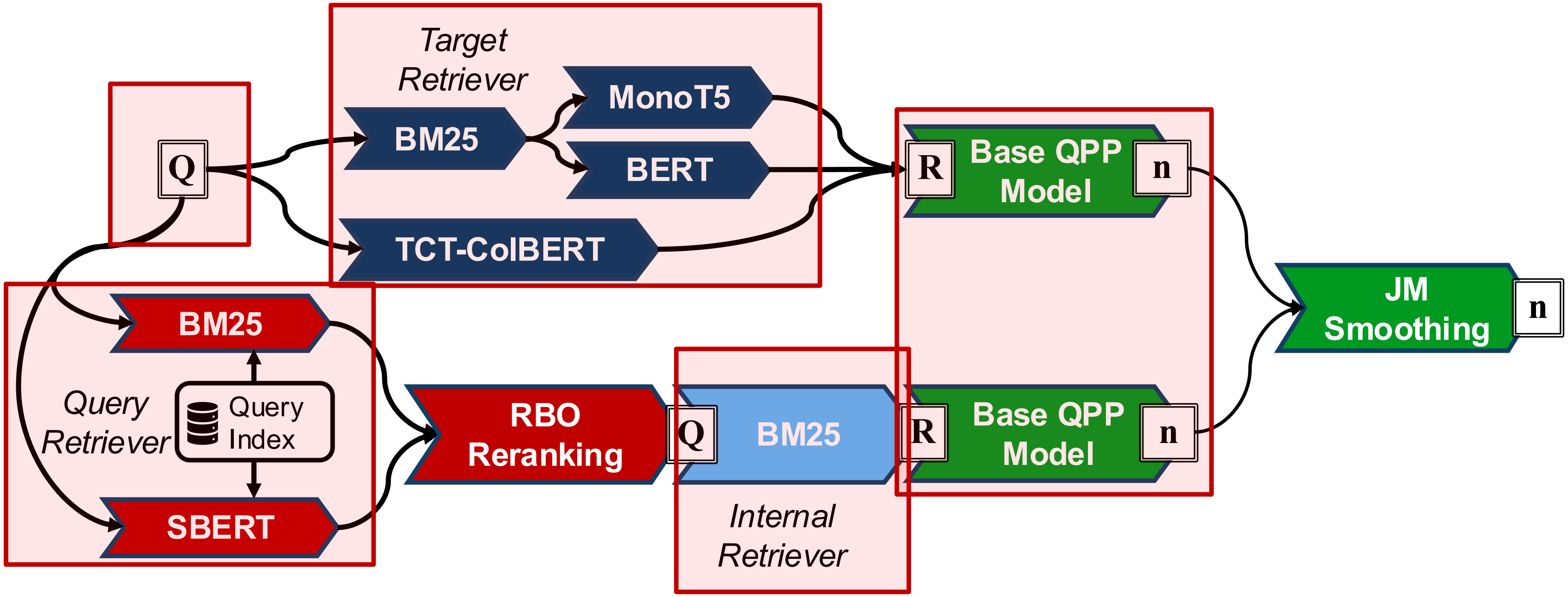
Regarding the effectiveness of QPP with retrieved QV.

RQ1: *Does QPP approaches with retrieved QVs outperform existing QV-based methods?*

Comparing between different configurations of the proposed QPP method.

RQ2: *Compared with 1-hop QVs, are the 2-hop QVs more useful for QV-based QPP?*

Experimental Setup



- Experiments are conducted on MSMARCO passage corpus, using the TREC DL'19 and '20 test query sets.
- Target retrievers are **MonoT5**, **BERT** and **TCT-ColBERT**.
- Query retrievers are lexical **BM25** and semantic **SBERT**.
- Internal Retriever is **BM25**.
- Base QPP models are **NQC** and NQC-based **UEF**.

Experimental Setup

Experimented QPP Methods:

- with 1-hop retrieved QVs: **QV-$\boldsymbol{R}^1$-BM25**, **QV-$\boldsymbol{R}^1$-SBERT**;
- with 2-hop retrieved QVs: **QV-$\boldsymbol{R}^2$-BM25**, **QV-$\boldsymbol{R}^2$-SBERT**;

Baseline QPP Methods:

- Base Predictors: **NQC**, **UEF**;
- Existing QV-based QPP Methods: **QV-RLM, QV-W2V;**
- Supervised **BERTQPP** and its QV-based variant **BERTQPP-QV**.

Target Metric:

- **MAP@100**, **nDCG@10**.

Evaluation Method:

- Grid searching the optimal values, then averaged of 2-fold train-test split.

Comparisons of QPP Effectiveness

Terrier Team

University of Glasgow

Kendall's τ between estimation and ground truth

Target retriever	BM25>>MonoT5		BM25>>BERT		TCT-ColBERT	
k=1	**AP@100**	**nDCG@10**	**AP@100**	**nDCG@10**	**AP@100**	**nDCG@10**
NQC	0.1673	0.0274	0.1278	0.0391	0.3991	0.2618
QV-W2V	0.2685	0.2041	0.2395	0.1520	0.3923	0.2521
QV-RLM	0.3308	0.1848	0.3045	0.1460	0.3920	0.2848
QV-R^1-BM25	0.3520	0.1918	0.3669*	0.2081	0.3954	0.2611
QV-R^1-SBERT	0.3361	0.2000	0.3558*	0.2016	0.4177	0.2561
QV-R^2-BM25	0.3694*	0.2298	0.3678*	0.2111	0.3878	0.2539
QV-R^2-SBERT	0.4033*	0.2573*	0.4250*	0.2517*	0.4022	0.2614
BERTQPP	0.2277	0.1746	0.1827	0.1328	0.2238	0.1326
BERTQPP-QV	0.2432	0.1728	0.2051	0.1459	0.2529	0.1514

Strongest Baseline

Best-performing proposed method

RQ1: ***Does the QPP with retrieved QVs outperform existing QV-based methods?***

- Our proposed methods outperforms existing QV-based methods.
- The advantage is larger for predicting AP@100.

Comparisons of QPP Effectiveness

Terrier Team

University of Glasgow

Kendall's τ between estimation and ground truth

Compare QPP with 1-hop QVs And QPP with 2-hop QVs

Target retriever	BM25>>MonoT5		BM25>>BERT		TCT-ColBERT	
k=1	**AP@100**	**nDCG@10**	**AP@100**	**nDCG@10**	**AP@100**	**nDCG@10**
NQC	0.1673	0.0274	0.1278	0.0391	0.3991	0.2618
QV-W2V	0.2685	0.2041	0.2395	0.1520	0.3923	0.2521
QV-RLM	0.3308	0.1848	0.3045	0.1460	0.3920	0.2848
QV-R^1-BM25	0.3520	0.1918	0.3669*	0.2081	0.3954	0.2611
QV-R^1-SBERT	0.3361	0.2000	0.3558*	0.2016	0.4177	0.2561
QV-R^2-BM25	0.3694*	0.2298	0.3678*	0.2111	0.3878	0.2539
QV-R^2-SBERT	0.4033*	0.2573*	0.4250*	0.2517*	0.4022	0.2614
BERTQPP	0.2277	0.1746	0.1827	0.1328	0.2238	0.1326
BERTQPP-QV	0.2432	0.1728	0.2051	0.1459	0.2529	0.1514

RQ2: ***Compared with 1-hop QVs, are the 2-hop QVs more useful for QV-based QPP?***

- Utilising relevant information of 1-hop QVs to retrieve 2-hop QVs enhances QPP effectiveness.

Conclusions

Takeaways:

- Retrieved QVs can be leveraged in QV-based QPP, yielding better QPP effectiveness than the existing QV-based QPP methods.
- 2-hop QVs are more useful than 1-hop QVs in terms of enhancing QPP effectiveness.

Insight:

- The QVs that exist in a training resembles the real queries executed by users -> They can be valuable for QPP estimations.

Future work:

- Integrating retrieved QVs with LLM to generate QVs to further enhance QPP effectiveness in neural rankers.

University of Glasgow

Thanks for your Attention!

Github Repository

QPP effectiveness with varying k

- QPP with retrieved QV outperforms the existing QPP methods when a small number of QVs are leveraged;
- But when k is large, QV-RLM can outperform the proposed methods (still worse than the best results obtained when k=1 to 3.