
dedication (optional)

Problem Description

Working title (?): Top-k relevant spatiotemporal retrieval on knowledge graphs using keyword search

Spatio-temporal keyword search on RDF graphs.

parts of the task:

Indexing:

In the paper "Top-k Relevant Semantic Place Retrieval on Spatial RDF Data"[1] a method using inverted index is proposed. This method is also used in "S. Elbassuoni and R. Blanco. Keyword search over RDF"[2] and will be used as a base for this thesis. The inverted index is made from the document description of all vertices. This will hopefully make it fast to retrieve vertices containing a keyword.

In paper [1] the keywords from a document description of a vertex is also proposed to be stored in a table, along with the spatial location, if it exists. I propose to extend this table to also include the temporal data, if it exists. This should make it possible to access the keywords and spatio-temporal data of a vertex during graph traversal. In the same paper all place nodes are also indexed in an R-tree to make incremental nearest place retrieval possible given a query location. Deciding on a data structure suitable for spatio-temporal data will be the first challenge to overcome in this thesis.

A proposed method is to substitute the r-tree used in [1] for a TPR-tree [3] though i do not yet know if this is necessary, since the TPR-trees are designed for moving objects, and the YAGO data set contains static objects(? I think, look more into this ?).

Traversing:// Start using traditional methods (e.g. BFS) and after the first working implementation start looking into more optimized traversal methods. The first part of the thesis will be creating the data structure, and proving that it is suitable for the task.

Ranking:

Like with traversing, this is not a priority. A proposed method is to traverse a subgraph, and rank based on the shortest path to hit all keywords, or hit most keywords within n-degrees. For spatial ranking a simple measure of euclidean distance between queried place, and closest place node can be used. For temporal ranking, the closest time can be used. The challenge in ranking is finding a method that takes all three aspects into account, this will however not be prioritized until a working data structure has been produced, and a traversal algorithm is implemented.

Problem description stuff...

Knowledge graphs allows for easy extraction of facts from an entity. The structure makes it more efficient than text processing, and the relations in the datasets makes multi degree relations between entities and facts more efficient than traditional table structures.

Currently there is little previous work done on spatial retrieval on knowledge graphs using keywords. There have been some (ref.) attempts ... Even less is done on spatiotemporal retrieval. Data collections such as, YAGO, holds both spatial and temporal on many different entities, making it possible to retrieve the entity data based on keyword searches.

To be able to do keyword search on RDF data the data needs to be indexed in some form, previously r-trees have been used (ref.) and graphs / sub-graphs must be traversed. Because the indexing for spatial RDF has been done previously, this thesis will not go too much in depth on that aspect. The temporal dimension has less research, so this thesis will explore this part of the problem in depth.

The task will also require some form of ranking to determine the top-k relevant results retrieved. There has also been done work on ranking RDF keyword search retrieval, so this thesis will not go in depth on the ranking.

In this project, we want to study spatio-temporal search on such graphs, aiming at developing indexes and algorithms for more efficient search.

[1] "Top-k Relevant Semantic Place Retrieval on Spatial RDF Data"

[2] S. Elbassuoni and R. Blanco. Keyword search over RDF

[3] Yufei Tao, Dimitris Papadias, and Jimeng Sun. 2003. The TPR*-tree: an optimized spatio-temporal access method for predictive queries

Summary

Write your summary here...

Preface

Write your preface here...

Table of Contents

Preface	3
Summary	i
Preface	ii
Table of Contents	iii
List of Tables	v
List of Figures	vii
Abbreviations	viii
1 Introduction	1
1.1 Equations	1
1.2 Figures	1
1.3 References	2
1.4 Tables	2
2 Literature Review	5
3 Basic Theory	7
4 Experiment	9
5 Analysis	11
6 Conclusion	13
Bibliography	15
Appendix	17

List of Tables

1.1 Table 1. 3

List of Figures

1.1	Pikachu.	2
-----	------------------	---

Abbreviations

Symbol = definition

Chapter 1

Introduction

1.1 Equations

To write an equation

```
\begin{eqnarray}\label{eq1}  
F = m \times a  
\end{eqnarray}
```

This will produce

$$F = m \times a \tag{1.1}$$

To refer to the equation

```
\eqref{eq1}
```

This will produce (1.1).

1.2 Figures

To create a figure

```
\begin{figure}[h!]  
  \centering  
    \includegraphics[width=0.5\textwidth]{fig/pikachu}  
    \caption{Pikachu.}  
  \label{fig1}  
\end{figure}
```

To refer to the figure



Figure 1.1: Pikachu.

```
\textbf{Fig. \ref{fig1}}
```

This will produce **Fig. 1.1**

1.3 References

To cite references

```
\cite{1,2,3}
```

or

```
\citep{1,2,3}
```

This will produce: ??? or (???), respectively.

1.4 Tables

To creat a table

```
\begin{table}[!h]
\begin{center}
\begin{tabular}{| l | l | l | l |}
\hline
\textbf{No.} & \textbf{Data 1} & \textbf{Data 2} & \\ \hline
1 & a1 & b1 & \\ \hline
\end{tabular}
\end{center}
\end{table}
```

```
2 & a2 & b2 \\ \hline
\end{tabular}
\end{center}
\caption{Table 1.}
\label{Tab1}
\end{table}
```

This will produce

No.	Data 1	Data 2
1	a1	b1
2	a2	b2

Table 1.1: Table 1.

To refer to the table

```
\textbf{Table. \ref{Tab1}}
```

This will produce **Table. 1.1**.

Chapter 2

Literature Review

Chapter 3

Basic Theory

Chapter 4

Experiment

Chapter 5

Analysis

Chapter 6

Conclusion

this is a test

Appendix

Write your appendix here...