dedication (optional)

Problem Description

Working title:
1: Top-k relevant spatiotemporal retrieval on knowledge graphs using keyword search
2: 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 possible method is using a slightly modified R-tree structure [3] to be able to index time dimension. An other possibility is the use of other tree structures, such as TPR-tree. The TPR-trees are designed for moving objects, so this might not be suitable for the task.

Traversing:
As a starting point, i propose using traditional methods (e.g. BFS) for traversal, and after the first working implementation start looking into more optimized methods. The first part of the thesis will be building an index, 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.

Work on problem description:

Knowledge graphs allows for easy extraction of facts from an entity, and the structure makes it efficient to find multi degree relations between different entities and facts. The task is to study indexing and search techniques on graphs containing spatio-temporal data, and propose effective ways to retrieve the data contained in the graph based on keyword searches

More thoughts/ notes:

Currently there is little previous work done on spatial retrieval on knowledge graphs using keywords. Some work is done on keyword search on RDF graphs, but very little that also takes spatial data into the search. Even less is done on spatio-temporal 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 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.

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

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

[3] R-Tree Based Indexing of General Spatio-Temporal Data

Summary

Write your summary here...

Preface

Write your preface here...

Table of Contents

List of Tables

List of Figures

Abbreviations

Symbol = definition

this is a test

Appendix

Write your appendix here...