

3D Search: Indexing and Ranking of Textual, Temporal and Spatial Features of Web Documents



Ali Khodaei and Cyrus Shahabi

Integrated Media Systems Center
University of Southern California

iCampus

iWatch

CT

Introduction

- Many applications are emerging in which data generated have **spatial** and **temporal** information.
- Users of such applications often need to query the system by providing requirements on a location and time as well as keywords

Robbery [August 24th 2011- August 26th 2011][near USC]

Problem Definition:

Given a number of search keywords, one or more locations and a time interval that user is interested in, a spatiotemporal web search finds and ranks the most textually, spatially and temporally relevant data objects according to query keywords, locations and time interval.

Application/Project

- **Web 2.0** applications



- **Web** documents

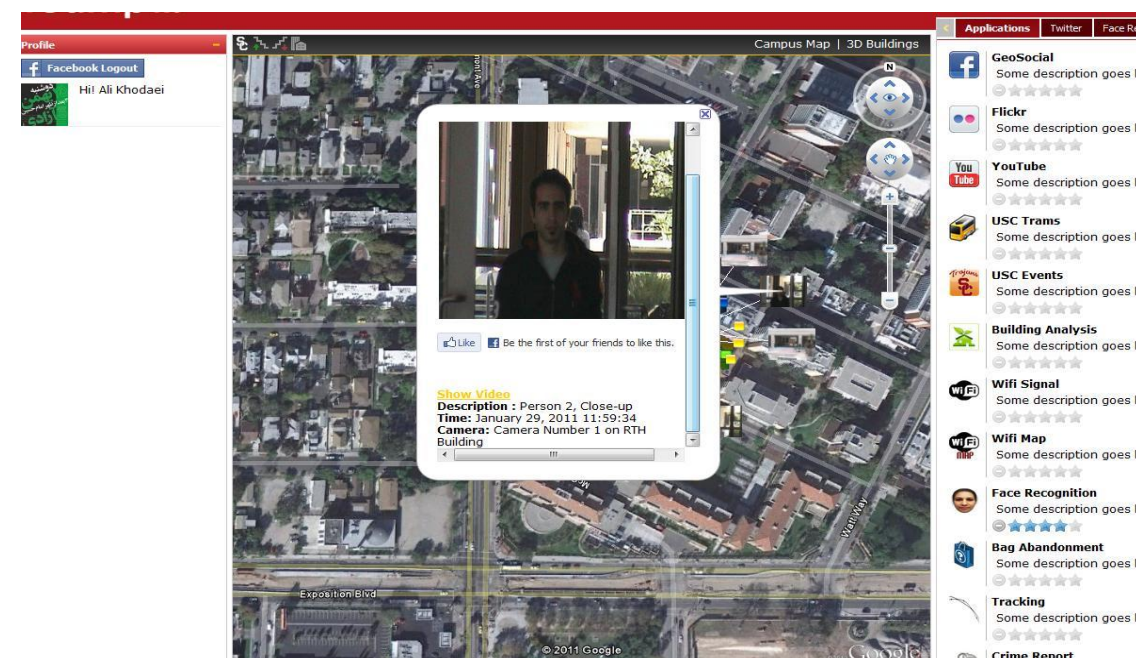
- News, events, biographies, ...

- **iWatch**

- Numerous heterogeneous

- surveillance data streams

- Video feeds, suspicious activities, crimes,



Challenges

- **Representation and Indexing**

- Time, space and text are three totally different data types requiring different index structures.
 - Conventional text engines are set-oriented while location indexes are usually based on two-dimensional spaces.
 - Neither is similar to temporal indexes : one dimensional and continuous space.
- A hybrid index structure should be able to simultaneously index all the temporal, spatial and textual features of the data objects and in a unified manner.

- **Ranking**

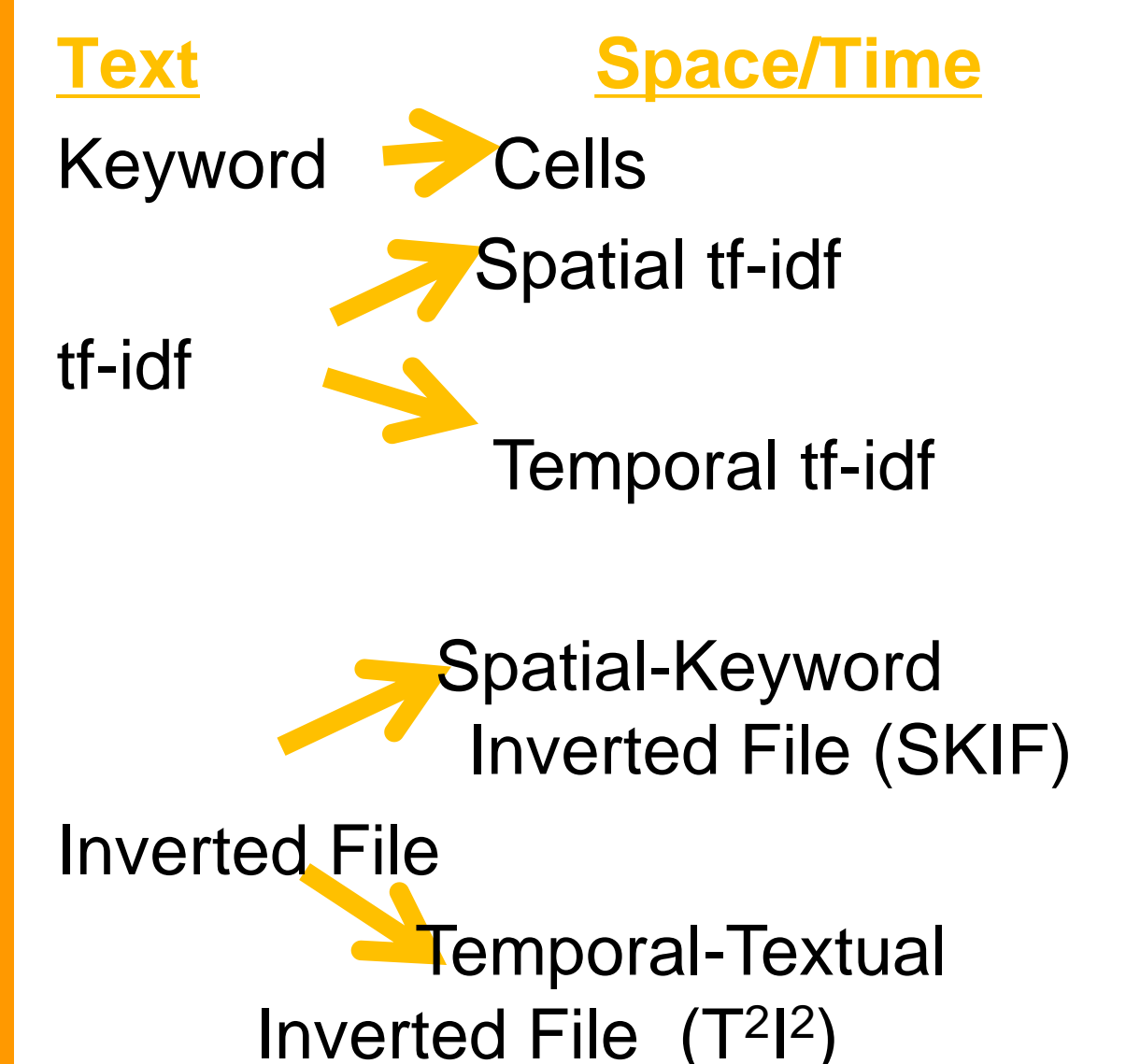
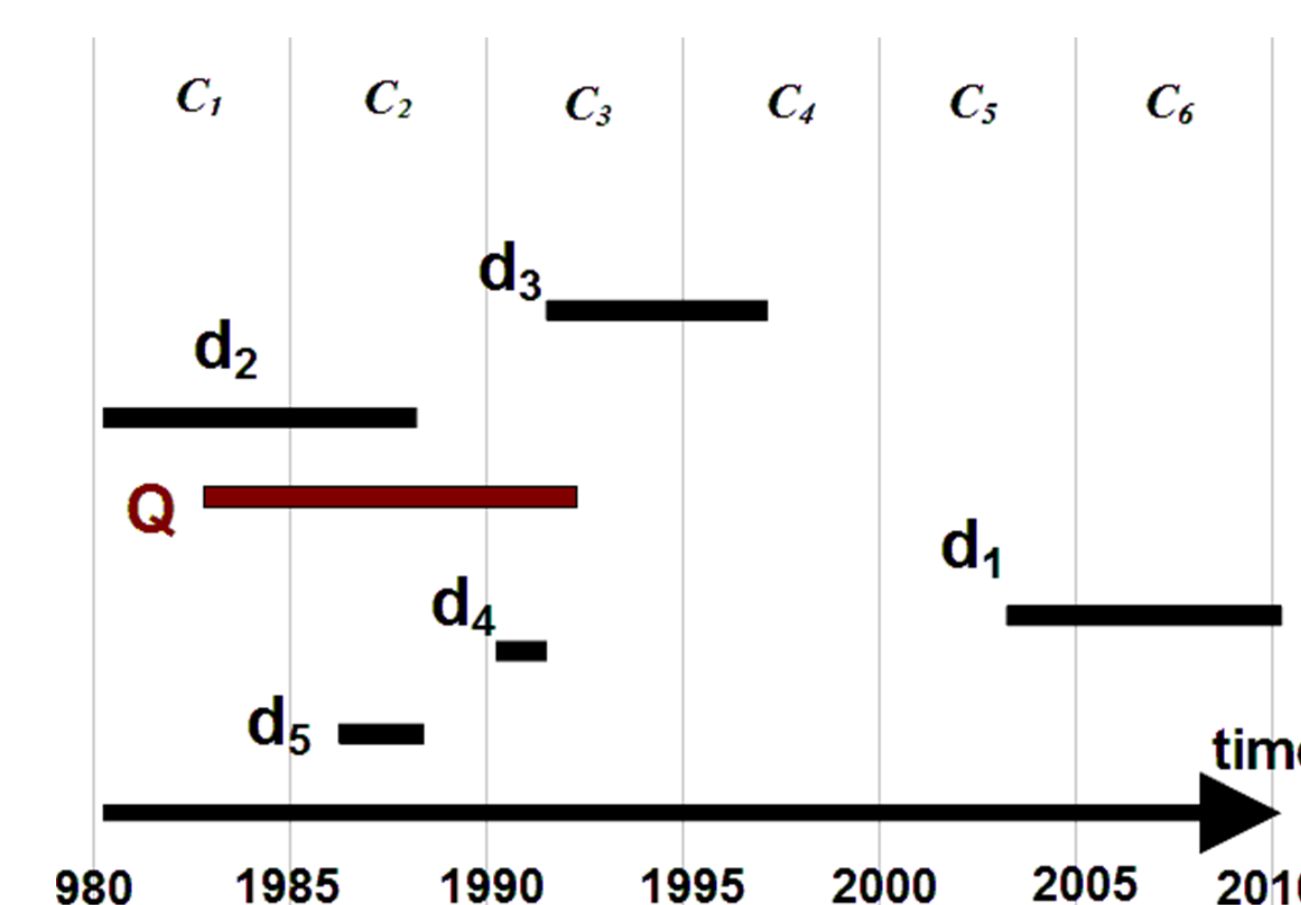
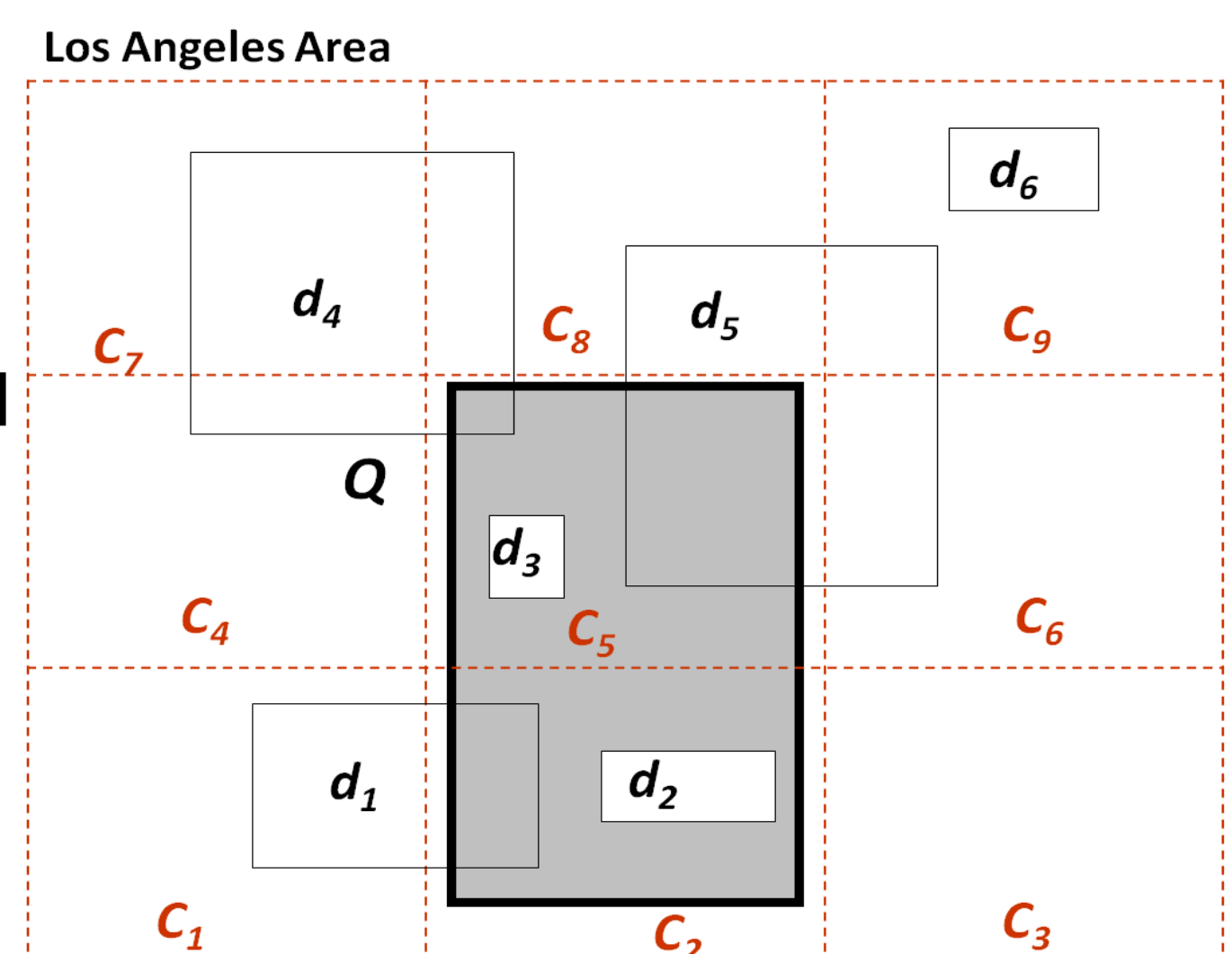
- How to measure the relevance of document to the query
 - Spatial relevance ?
 - Temporal relevance ?
- How to combine spatial relevance with temporal relevance and combine them with textual relevance accurately and seamlessly ?

Related Work

- No work on ranking and indexing of spatial, temporal and textual features of objects combined
- Textual-spatial index structures
 - Individual index structures / Hybrid index structures
- Textual-temporal index structures
 - Most approaches do not consider the temporal information in the documents' content for the relevance ranking and retrieval .

Approach

- Using same intuitions and concepts used in regular textual searches
- we define new concepts and parameters for spatial and temporal data.
- We represent space (and time) which is coherent and continuous in nature, as disjunct and set-oriented units of data
 - similar to the textual keywords



Experiments

- Performance
 - Superior **Efficiency**
 - *Metrics*: response time and number of I/Os
 - *Parameters*: Number of keywords, number of results (k), query timespan, query spatial length, weight (textual vs. non-textual)
- Relevance Ranking
 - Superior **Accuracy**
 - *Metrics*: R-precision, precision@k, nDCG@k
 - Several extensive user studies

Future Work

- Add social dimension
 - How to Represent, index and rank data with social features
 - How to combine it with spatio-temporal relevance ranking and indexing

Restaurant Socially relevant to me in Orange County
[Socially relevant: recommended by my friends, liked by my social network, linked by my followers, etc.]