Disambiguating Implicit Temporal Queries for Temporal Information Retrieval Applications

Fraunhofer Portugal Challenge 2013

Participant Name: Ricardo Campos, PhD

Supervisor: Professor Gaël Dias (University of Caen Basse-Normandie)
Co-Supervisor: Professor Alípio Mário Jorge (University of Porto / LIAAD)

Fraunhofer AICOS offices
Porto, Portugal, October 30, 2013
Is Time a new Dimension of Information Retrieval?

INTRODUCTION

Main Idea

Objectives

Temporal IR Applications

Conclusions

Is Time a new Dimension of Information Retrieval?

- Emails
- Webpage’s, Blogs
- Newswire Articles
- Web Snippets
- Web Logs
- Digital Libraries

Ricardo Campos – Fraunhofer Portugal Challenge 2013

Disambiguating Implicit Temporal Queries for T-IR Applications
Is Time a new Dimension of Information Retrieval?

Due to its importance, **T-IR has become an active research area** over the last few years;

In general, T-IR aims to satisfy the user needs by combining the traditional notion of **document relevance** with the so-called **temporal relevance**;

Usually, this demands users to **explicitly specify a time span** together with the query:
Due to its importance, **T-IR has become an active research area** over the last few years;

In general, T-IR aims to satisfy the user needs by combining the traditional notion of *document relevance* with the so-called *temporal relevance*;

Usually, this demands users to **explicitly specify a time span** together with the query:
Due to its importance, **T-IR has become an active research area** over the last few years;

In general, T-IR aims to satisfy the user needs by combining the traditional notion of *document relevance* with the so-called *temporal relevance*;

Usually, this demands users to **explicitly specify a time span** together with the query:
Due to its importance, **T-IR has become an active research area** over the last few years;

In general, T-IR aims to satisfy the user needs by combining the traditional notion of *document relevance* with the so-called *temporal relevance*;

Usually, this demands users to **explicitly specify a time span** together with the query:
Disambiguating Implicit Temporal Queries

- Fifa World Cup Germany

Introduction
MAIN IDEA
Objectives
Temporal IR Applications
Conclusions

1974

2006
Disambiguate Implicit Temporal Queries

• Fifa World Cup Germany
• Avatar Movie

1974  2006  2009
Disambiguate Implicit Temporal Queries

- Fifa World Cup Germany
- Avatar Movie
- Margaret Thatcher

1979  1990  2012  2013
• Understand the **temporal nature** of any given **implicit temporal query** in a way that allows us to reach the temporal disambiguation of the query “on-the-fly”;
Goals

- **Understand** the **temporal nature** of any given **implicit temporal query** in a way that allows us to reach the temporal disambiguation of the query “on-the-fly”;

- Use this information to **improve search results exploration**, namely the **clustering** and **ranking** of search results;
Search engine user’s (having or not a temporal focus) whose issued queries may benefit from a temporal contextualization (e.g. “bp oil spill”, “Haiti earthquake”, “avatar movie”, etc);

- **Improve** the effectiveness of the results;
- **Satisfy** the user information needs;
- **Understand** the different temporal facets of the query;
- Give user’s an **historical perspective** of the query results.
Many (Temporal) Applications

Search Results Diversification

- Query Expansion
- Result Re-Ranking
- Temporal Web Image Retrieval
- Temporal Clustering
- Web Search Results Visualization
- Future Retrieval
GTE-Cluster implements a flat temporal clustering model that groups documents into the same cluster if they share a common year.

Haiti Earthquake
The first great earthquake mentioned in histories of Haiti occurred in 1564 in what was still the Spanish colony. It destroyed Concepción de la Vega.
GTE-Cluster implements a flat temporal clustering model that groups documents into the same cluster if they share a common year.

**2010**  
**2010 Haiti Earthquake**  
The *2010 Haiti earthquake* was a catastrophic magnitude 7.0 Mw earthquake, with an epicenter near the town of Léogâne....

**2011**  
**2011 Haiti Earthquake Anniversary**  
As of 2010 (see 1500 photos), the following major earthquakes have been recorded in Haiti. The 1\textsuperscript{st} one occurred in 1564.
GTE-Cluster implements a flat temporal clustering model that groups documents into the same cluster if they share a common year.

The following major earthquakes have been recorded in Haiti. The 1\textsuperscript{st} one occurred in 1564.

As of 2010 (see 1500 photos), the 2011 Haiti Earthquake Anniversary

---

**1500**

2011 Haiti Earthquake Anniversary

As of 2010 (see 1500 photos), the following major earthquakes have been recorded in Haiti. The 1\textsuperscript{st} one occurred in 1564.
Introduction

Main Idea

Objectives

Introduction

TEMPORAL IR APPLICATIONS

Conclusions

GTE-Cluster

Query: 11 de Setembro  
Search

☐ Check this box if you want to return only relevant dates

The red color means a rejected cluster whereas the blue one means a selected one.

1945: 0.0798515765757054

11 de setembro – Wikipédia, a enciclopédia livre
11 de setembro (AO 1945: 11 de Setembro) é o 254. dia do ano no calendário gregoriano (255. em anos bissextos). Faltam 111 para acabar o ano.
http://pt.wikipedia.org/wiki/11_de_setembro

1992: 0

11 de Setembro 1992 (11dSetembro1992) on Twitter
The latest from 11 de Setembro 1992 (@11dSetembro1992). Nasceu a garota mais perfeita, humilde, simpática e que me enxe de orgulho. Maria Gabriela de Faria Chacón ...
http://twitter.com/11dSetembro1992

2001: 0.6969767507789681

Ataques de 11 de setembro de 2001 – Wikipédia, a enciclopédia ...
Ataques ou atentados terroristas de 11 de setembro de 2001 (às vezes, referido apenas como 11 de setembro) foram uma série de ataques suicidas contra os Estados ...

11 De Setembro 2001 - World News
11 de setembro de 2001 - Jornal Nacional [1 bloco], 11 de Setembro 2001 - Documentário - Torres gémeas - world trade center "Recordando a Tragédia", 11 de ...
http://www.com/11_de_setembro_2001
Disambiguating Implicit Temporal Queries for T-IR Applications

Ricardo Campos – Fraunhofer Portugal Challenge 2013

**Main Idea**

**Objectives**

**Conclusions**

**Introduction**

**TEMPORAL IR APPLICATIONS**

GTE-Cluster

Query: [Search] □ Check this box if you want to return only relevant dates

The red color means a rejected cluster whereas the blue one means a selected one.
We seek to re-rank the results of implicit temporal queries so as to enhance the overall temporal part of the web search results;

Our aim is to retrieve, in the top list of results, documents that are not only topically relevant but that are also from the most important time periods;
Query: carlos paiao

The old id of the snippet is highlighted with red color followed by the GTERank value.

18: 2.15

Carlos Paiao Discography at Discogs
Singer and song-writer from Portugal, born 1 November 1957, died in a car accident 26 August 1988. He represented Portugal at the 1981 Eurovision Song Contest with ...
http://www.discogs.com/artist/Carlos+Pa%23A3o

32: 1.99

Carlos Paiao - Cinderela . Top download charts Portugal
Bio: Carlos Manuel de Marques Paiao (Coimbra, November 1, 1957 - Rio Maior, August 26, 1988) was a very popular singer and song-writer in Portugal.
http://pt.hitclipz.com/song/Carlos+Pa%23A3o%C3ADCinderela/V1q41X1NkBw/http%253A%252F%252Fwww.apple.com%252Fen%252Fchart%252Falbum%252Ff-cinderela%252F

1: 1.94

Carlos Paiao - Wikipedia, the free encyclopedia
Carlos Manuel de Marques Paiao (born in Coimbra on November 1, 1957 - August 26, 1988 in Rio Maior) was a very popular singer and song-writer in Portugal.
http://en.wikipedia.org/wiki/Carlos_Pa%23A3o

33: 1.87

Carlos Paiao - Wikipédia, a enciclopédia livre
Carlos Manuel de Marques Paiao (Coimbra, 1 de Novembro de 1957 - Rio Maior, 26 de Agosto de 1988) foi um cantor e compositor português. Licenciou-se em Medicina ...
http://pt.wikipedia.org/wiki/Carlos_Pa%23A3o
Disambiguating Implicit Temporal Queries for T-IR Applications

Ricardo Campos – Fraunhofer Portugal Challenge 2013

**Main Idea**

**Objectives**

**Conclusion**

**GTE_Rank: Temporal Re-Ranking**

- **Query:** [Input Field]
- **Search:** [Button]
- **Temporal Score:** 0.8

The old id of the snippet is highlighted with **red color** followed by the GTERank value.
Main Achievements

- Estimate the degree of relevance between candidate dates and a given query;
- Is able to identify top relevant dates for any given implicit query while filtering out non-relevant ones;
- Clusters Web search results;
- Testing each solution is possible through already public interfaces.
- Re-organizes them based on their temporal nature;
More Applications

Discover what the future will bring, mainly in texts of rumor nature

*Avatar 2* is a planned sequel to the 2009 film *Avatar* expected to be released in 2016........
More Applications

Discover what the future will bring, mainly in texts of rumor nature

Adopt an approach focused on the topic clustering dimension
More Applications

Discover what the future will bring, mainly in texts of rumor nature

Adopt an approach focused on the topic clustering dimension

Detect the period of time the query facets are related with
Thanks for your attention!

GTE-Cluster: http://wia.info.unicaen.fr/GTEAspNetFlatTempCluster_Server
GTE-Rank: http://wia.info.unicaen.fr/GTERankAspNet_Server