Experiments with Geo-Temporal Expressions filtering and query expansion at document and phrase context resolution

Jorge Machado
INESC-ID, Lisbon
Rua Alves Redol 9 Apartado 13069
1000-029 Lisboa
+351213100300
jorge.r.machado@ist.utl.pt

José Borbinha
INESC-ID, Lisbon
Rua Alves Redol 9 Apartado 13069
1000-029 Lisboa
+351213100300
jlb@ist.utl.pt

Bruno Martins
INESC-ID, Lisbon
Rua Alves Redol 9 Apartado 13069
1000-029 Lisboa
+351213100300
bruno.martins@ist.utl.pt

ABSTRACT

We describe an evaluation experiment on GeoTemporal Document Retrieval created for the GeoTime evaluation task of NTCIR 2010. GeoTemporal Retrieval aims at to improve retrieval results using Geographic and Temporal dimensions of relevance. To accomplish that task, systems need to extract geographic and temporal information from the documents, and then explore semantic relations among those dimensions within the documents. Since this is the first time the task is taking place our aim is to evaluate some basic techniques in order to set some research directions of our work. We aim to understand the relevance of temporal and geographic expressions for filtering purposes. The geographic expressions were extracted with Yahoo PlaceMaker and for temporal expressions we used the TIMEXTAG system. We experimented techniques using both the overall document and sentence resolutions, as also one mixed approach. We also used a query expansion mechanism in topics with no filters defined. We used the BM25 as retrieval model and preprocessed the topics with a semi-automatic methodology to create structures that let us create our filters and expansions. We learned that the sentence level is not a very good approach (but we got clues that probably the paragraph context resolution could improve the results) and the geographic and temporal expressions base filters had shown good performance.

Keywords

Geographic and Temporal Information Retrieval, Probabilistic Models, Multidimensional Retrieval Models, Hierarchical Indexes.

1. INTRODUCTION

This work was motivated by the GeoTime evaluation task that was part of NTCIR 8 workshop[1]. The GeoTime task aims to evaluate retrieval techniques focusing in geographic and temporal retrieval.

Geographic Information Retrieval (GIR) addresses the combination of usual Information Retrieval (IR) techniques with new techniques for addressing the geographic dimension of relevance. Previously proposed GIR systems can usually be broken down into three main processing stages: i) geographic entities extraction, ii) indexing documents with basis on the meaning behind the locations mentioned in the text, and ii) ranking the indexed documents with respect to geographic queries [5]. Regarding the fusion of semantic dimensions with raw text (see [7]), the most common techniques have been based on query expansion [9], filtering [8] and distance measures [4].

Temporal Information Retrieval (TIR) is a recent subject which addresses the combination of usual Information Retrieval (IR) techniques with new ones for addressing temporal dimension of relevance. The time is present everywhere, and every fragment of information has a potential value related with a specific temporal context. The value of time [3] creates a possibility of improving retrieval systems. Temporal Retreival systems could have four main stages, i) the temporal expressions extraction ii) temporal expressions normalization iii) indexing the expressions or a representation of them, and iv) the ranking of documents with respect to the temporal queries. However, the temporal features tend to be more complex to extract, since usually they have a more complex logic behind them. First of all is necessary to understand that temporal retrieval is not only focused on the date of the document, the date of document is typically used to increase relevance of recently updated documents [10] or to simple sort the results. The major challenge is to understand the meaning of the temporal expressions present the text. This kind of approach requires techniques for the extraction of temporal expressions, such as event ordering [2] and machine learning [1]. Representing temporal expressions can be done using schemas such as TIDES [11] or TIMEML [12]. Temporal Expressions can be modeled through a temporal logic [14], which includes temporal predicates and relations for events over time. Examples of predicates are: holds, occur, in, generates. Examples of relations could be: meets, equal, overlap, before, starts with and finish after. Identify temporal expressions is not an easy job, as they depend a lot from the native language grammar In GIR the names of places usually keep their representation with small variations from language to language, while in TIR there is the need to match more complex expressions.

The objective of this work was to understand what techniques might be more effective to merge or support some kind of fusion between Space Time, and Text. We aim to use traditional retrieval models, like BM25, but with no specific specializations for this first approach. Taking into account that temporal and geographic expressions can be normalized and indexed, our objective was to perform experiments on filtering and query expansion using several resolutions. We define resolution as the capability of index more fine grain texts, such as statements, in order to give relevance to the context between expressions. Using filters makes our experiment strongly dependent of the query processing phase,

[1] http://research.nii.ac.jp/ntcir/

which will set how to use an expression, as a filter or as an expansion feature.

In this paper we describe an experiment to set some directions in Geo-Temporal retrieval research. In section 2 we describe the entire experiment phase including the collection processing step of the articles from the test collection New York Times (2002-2005). In that section we present our first contribution, which is the statistical information about geographic and temporal expressions extraction, second we detail the documents processing and finally the topic processing. Later on section 3 we discuss the results and in section 4 we conclude this document and set some future directions.

2. EXPERIMENTAL HISTORY

Next section 2.1 details the characteristics of the collection, and in section 2.2 we detail the collection processing. The experiment consisted in 5 runs over 25 GeoTemporal topics. The topics are detailed in the Overview paper for the GeoTime task [13]. Topics consisted in questions mostly using the adverbs *when* and *where,* or providing some geographic references of temporal expressions in form of restrictions for the retrieved documents. In section 2.3 we detail our topic processing step.

We used 3 systems for our experiment. First, we extracted geographic entities using the online service Yahoo PlaceMaker[2]. For temporal expressions extraction we used the TIMEXTAG[3] tagger, developed at University if Amsterdam. The indexes were created with our tool LGTE[4], based on the Lucene text indexer with extensions for probabilistic models, geographic retrieval and hierarchical indexes. Below we detail our processing steps.

2.1 Collection Extraction Statistics

The GeoTime task used an English collection of news articles from New York Times published between January of 2002 and December of 2005, consisting of 315.417 documents. More details about the collection can be found in the Overview paper for the GeoTime task [13]. Geographic and temporal expressions were extracted using, respectively, PlaceMaker and TIMEXTAG. This section is dedicated to report the extracted data in order to validate the relevance of the geo-temporal information used in the experiment. In Tables 2 to 4 we report the extraction with Yahoo PlaceMaker, while in tables 5 to 9 we report the TIMEXTAG extraction of temporal expressions. In Table 1 we summarize the totals of documents with geographic places extracted.

Table 1 – Geo-Parsing General Statistics.

	Documents	%
Docs with Places	302695	95,97%
Docs with no Places	12722	0,30%
Docs Failed Anotation	0	0,00%
Docs	315417	100,00

[2] http://developer.yahoo.com/geo/placemaker/

[3] http://ilps.science.uva.nl/resources/timextag

[4] http://code.google.com/p/digmap/wiki/LuceneGeoTemporal

In Table 2 we show the place types distribution over the collection, as extracted by PlaceMaker.

Table 2 – Place types distribution over documents.

Woeid Types	Doc Frequency	References	%References
Town	1047125	1785315	42,75%
Country	419690	965972	23,13%
State	319410	577383	13,82%
POI	210048	307474	7,36%
Suburb	102924	149180	3,57%
County	79251	125312	3,00%
Colloquial	46198	66980	1,60%
Continent	32190	59625	1,43%
Supername	29234	39758	0,95%
ZIP	16604	17122	0,41%
LandFeature	10423	15729	0,38%
Airport	11048	14653	0,35%
Island	9038	12799	0,31%
HistoricalTown	5627	9528	0,23%
Ocean	7052	9475	0,23%
Sea	6321	8443	0,20%
Drainage	4617	6038	0,14%
LocalAdmin	2306	3604	0,09%
Miscellaneous	458	694	0,02%
HistoricalState	477	630	0,02%
Estate	356	460	0,01%
HistoricalCounty	216	317	0,01%
DMA	11	12	0,00%
Market	4	4	0,00%
Zone	2	2	0,00%
Total	2328440	4176509	100,00%

In Table 3 we present the Yahoo confidence degrees. The results show that more than 80% of the extracted places have a degree of confidence higher or equal to 7, in a scale of 1 to 10. This is a good indicator to use geographic entities in this collection.

Table 3 – Yahoo Place Maker confidence degree.

Yahoo Conf	Doc Frequency	Refs	% Refs
9	*1071096*	*1989415*	*47,63%*
8	*377001*	*693597*	*16,61%*
10	*314755*	*471296*	*11,28%*
7	*202086*	*354161*	*8,48%*
6	192948	338193	8,10%
5	72404	112156	2,69%
4	52738	81701	1,96%
3	41896	65548	1,57%
2	30541	49241	1,18%
1	12741	21201	0,51%
Total	2368206	4176509	100,00%

The totals for normalized WOEID (Where on Earth IDentifier) identifiers are found in Table 4. BelongTos are the places belonging to the tree of administrative parent regions for one given place, starting in a parent defined as the smaller administrative region containing the given place, following by the

smaller administrative region containing the parent of the parent, and so on.

Table 4 - Normalized WOEID's.

	Indexed Expressions	References
Place WOEID	70477	4176509
Administrative Scopes WOEID	2632	302695
Geographic Scopes WOEID	3752	302695
BelongTos WOEID	61299	58640147
All WOEID	138160	63422046

Tables 5 to 9 summarize the collection characteristics in terms of temporal expressions. Table 5 counts the number of temporal expressions (*timexes*) found in collection with TIMEXTAG system.

Table 5 – Temporal Expressions general statistics.

	Documents	%
Docs with Timexes	311490	98,75%
Docs with no Timexes found	3809	1,21%
Docs with Indexable Time Exprs	301235	95,50%
Docs with not Indexable Time Exprs	14182	4,50%
Docs Failed Annotation	118	0,04%
All Docs	315417	100,00

In Table 6 we display the formats of expressions normalized by TIMEXTAG vs. unknown expressions or impossible to normalize.

Table 6 - Normalized formats statistics.

Expression	Unique	Refs.	%
Y	5	229	0,01%
YY	31	11041	0,26%
YYY	80	60734	1,41%
YYYY	3916	18846655	17,44%
YYYY-MM	1297	318580	5,72%
YYYY-MM-DD	10041	2024089	34,53%
YYYY-Wn	34	100673	2,33%
UNKNOWN	not indexed	1652866	38,31%
Total References	15370	4314808	100,00

Duration expressions, or time periods, are represented in TIDES schema using the structure PnK, where n is the number of time periods which have passed and K represent days (D), months (M), years (Y) or weeks (W). This kind of expression is followed by an anchor which is a normalized date and finally a direction that define if the anchor marks the start or the end of the period. As an example consider the duration P2W with anchor 201001 and direction STARTING, this means that the period is the first 2 weeks of January 2010. In Table 7 we present all duration expressions that we were able to expand and index. Expressions "Week of the Year (YYYY-Wn)" are not included in this table. As we can see in the Table 7 this kind of expressions is very usual and in that sense research groups should address it providing new techniques to improve the retrieval.

Table 7 – Duration expressions expanded and indexed.

Expanded Timexes	Direction	Anchor Format	Timexes
PnD (Starting)	STARTING	YYYY-MM-DD	947
PnD (Ending)	ENDING	YYYY-MM-DD	1766
PnW (Starting)	STARTING	YYYY-Wn	1104
PnW (Ending)	ENDING	YYYY-Wn	3936
PnM (Starting)	STARTING	YYYY-MM	1700
PnM (Ending)	ENDING	YYYY-MM	6566
PnY (Starting)	STARTING	YYYY	6786
PnY (Ending)	ENDING	YYYY	50558
Unique Durations Found			5365
References			77781

In Table 8 we present the expressions of durations that were not used. Mainly are expressions representing time periods with not well defined limits. For example the expression "before 2010" is true for events that happened in 2008 but also for events that happened in 2000. This kind of expression was not addressed in our experiment, but is targeted of future research. Probably it is not a good idea to index this kind of expressions with temporal tokens expressing a date because the periods could easily cover very big time intervals. Index the document with temporal limits is probably a better choice but the topic processing must consider that fact and understand the user needs to map those needs to the equivalent query. In this case *a greater than* or *smaller than* query is a possibility.

Table 8 - Duration expressions not used.

Not Used Timexes	Direction	Anchor Format	Timexes
PnD (BEFORE)	BEFORE	YYYY-MM-DD	41880
PnD (AFTER)	AFTER	YYYY-MM-DD	1
PnD (NULL)	NULL	UNKNOWN	271
PnW (BEFORE)	BEFORE	YYYY-Wn	26129
PnW (NULL)	NULL	UNKNOWN	303
PnM (BEFORE)	BEFORE	YYYY-MM	31135
PnM (AFTER)	AFTER	YYYY-MM	1
PnM (NULL)	NULL	UNKNOWN	429
PnY (BEFORE)	BEFORE	YYYY	139020
PnY (AFTER)	AFTER	YYYY	3
PnY (NULL)	NULL	UNKNOWN	1069
Total References			240241

Table 9 summarizes the indexed expressions as key points which are expressions that could be directly normalized to a date using only the document date, generated expressions resultant from event ordering techniques, duration expressions and finally the documents publishing dates.

Table 9 – Indexed Temporal Expressions

	docs	refs	% to
Key Points	14687	2350436	50,93%
GenPoints	4	104	0,01%
Expanded from Durations	1389	1948788	42,23%
Total - T1	15370	4299328	93,17%
Document DateTime	1363	315417	6,83%
Total - T2 (include doc date)	15370	4614745	100,00

2.2 Collection Processing

Our experiment aimed to compare the filtering and query expansion approaches using geographic and temporal expressions in two different contexts, the document and the phrase. We started the processing by splitting the 315.417 documents into 11.702.480 sentences using LingPipe[5], a natural language processing tool. Next we created an index for documents and other one for sentences. For this task we used the LGTE hierarchical indexes extension for Lucene, which support mixed queries using document text and sentence text. Second, we parsed all the documents with Yahoo Placemaker to collect the places references and place types from the text. We performed a second passage through the extracted places to obtain the hierarchy of parents, known as "belongTos". We created an index of those references for each document, using identifiers WOEID, a unique reference space assigned by Yahoo Placemaker to identify any feature on Earth. For example the WOEID for Lisbon is 2346573, so we created the token "WOEID-2346573" to index it. After splitting the document into sentences we create another index to map the WOEID's to the obtained sentences crossing the text offset's given by LingPipe with those given by PlaceMaker.

In terms of temporal features, we extracted temporal expressions structured in TIDES schema using TIMEXTAG system. We addressed only the subset of timexes that we were able to normalize into metric dates. We addressed the following types of timexes: time points defined as dates expressed in natural language, generated points defined as relative dates expressed in natural language and normalized using an anchor which is also a generated or a time point, weeks defined with the number of the week in year, and finally anchored durations defined as time intervals that we were able to normalize and expand using an anchor which is a time point. We used day granularity to index all the normalized dates. We used the following format to index dates: YYYY[MM[DD]] (where Y is a digit for year, M for months and D for days). The usage of the parenthesis means not mandatory parts. This means that we indexed years such as "2005" or months such as November of 2005 using the token "200511". Period durations were normalized to the equivalent set of dates which could be days, months or years depending on the duration scope. For example a duration referring to the period from 1 January 2010 to 31 of January of 2010 was normalized to January of 2001 using the token "201001". A duration referring to the Week 4 of 2010 was normalized to the equivalent set of days from January 18 to January 24 using the tokens "20100118", "20100119", and so on until "20100124".

Tables 4 and 9 of the previous sub-section summarize the number of expressions indexed.

We created 6 groups of indexes: Contents, Sentences, WOEID's, WOEID's-Sentences, Timexes, Timexes-Sentences. The coverage of geographic hierarchies was done at index level using the belongTos every time a topic asks for some place inside another (e.g all cities in USA, we want documents where USA must be indexed in belongTos index). The coverage of hierarchic dates was done at query level using a wildcard * (e.g. earthquakes in 2002 results in the temporal filter 2002* that will retrieve all documents indexed with temporal expressions started by 2002)

[5] http://alias-i.com/lingpipe/index.html

2.3 Topic Processing

Each of the 25 topics of GeoTime had an identifier, a description and a narrative. We parsed the topics with a semi-automatic processor supervised by ourselves. We aim to split the topics in three dimensions of relevance: terms, places and times. The terms were obtained removing stopwords and reducing words with a stemming step implemented in a package of Lucene tool. We extracted places, place types and temporal expressions which were also removed from terms dimension. Semantically we also aimed to filter the user needs using restrictions in time and places dimensions. We defined a topic grammar and preprocessed the 25 topics to create a representation for each one. The grammar consisted in one filter chain of logic filters and a query part consisting in text, space and time terms. The filter chain aims to represent topic restrictions captured from the text of the topic. We used the description and the narrative to capture restrictions and query terms. Filters were structured into 4 types, as illustrated in Table 10.

Table 10 - Indexed tokens used in filters.

Features	Found values
woeidType	country, city, province
timeType	year, year-month, exact-date, any
place	Yahoo PlaceMaker WOEID references
time	Normalized Expressions found with TIMEXTAG

We captured place names, temporal expressions, place types and temporal expressions types. We considered geographic expressions all place references found in the topic by the system Yahoo PlaceMaker. Were also considered as restrictions expressions all types of references like "city" or "province" found near the text fragments considered user needs, explained below. We considered a time expression the set of all temporal references found by the TIMEXTAG tagger and restriction expressions referring to date formats like for example "the exact date" or "the date and month" found near the text fragments considered user needs. To filter exact dates or date restrictions found in topics we indexed all dates grounded by TIMEXTAG which could be relative dates like "last year", exact dates like "in December 2002" and durations like "between 2002 and 2005".

We defined a question filter of the topic the set of all of the geographic and temporal expressions which occur near an adverb like "what", "where", "when", or the group "How long after/before", "How many time after/before" (e.g "In what province of China…"). We also considered restrictions those expressions declaring the user needs like for example "wants to find", "would like to know", "which one" and so on. Taking as example this topic "wants to know what **month and year**", for cases like this we considered month and year a restriction on the temporal expression type that should be of the kind YYYYMM. More examples of the used expressions are: "want to know the **country**", "want to know the **exact date**", "In what **city**", "In what **province** of **China**", "How long **after**", etc.

All terms found using the previous technique, including adverbs in questions, user references, places, times, places properties and time properties, were removed from the text fields description and narrative and placed in filters as geographic or temporal terms filters. Places' names and normalized dates

references not considered by the previous set of rules were removed from the terms fields description and narrative and placed in their own dimensions of relevance queries. We also removed the stopwords and punctuation characters. The follow example illustrates the topic GeoTime-0025 which was one of the most difficult topics to process:

```
<topic id="GeoTime-0025">
 <original>
   <desc>How long after the Sumatra earthquake did the tsunami hit Sri Lanka?
   </desc>
   <narr>The largest earthquake in recent times occurred off the coast of Sumatra in 2005. The earthquake caused a massive tsunami which spread across the Indian Ocean. The user would like to know how long it took the tsunami to reach Sri Lanka. An somewhat indefinite answer like 'a few days' is acceptable.</narr>
 </original>
 <originalClean>
   <desc>How long after Sumatra earthquake tsunami hit Sri Lanka</desc>
   <narr>largest earthquake recent times occurred coast Sumatra 2005 earthquake caused massive tsunami which spread across Indian Ocean how long took tsunami reach Sri Lanka</narr>
 </originalClean>
 <filterChain>
   <boolean type="AND">
     <term>
       <field>place</field>
       <value woeid="23424778">Sri Lanka</value>
     </term>
     <term>
       <field>place</field>
       <value woeid="12493166">Sumatra</value>
     </term>
     <term>
       <field>timeType</field>
       <value>any</value>
     </term>
   </boolean>
 </filterChain>
 <terms>
     <desc>earthquake tsunami hit</desc>
     <narr>largest earthquake recent times occurred coast earthquake caused massive tsunami spread across took tsunami reach</narr>
   </terms>
   <places>
     <term woeid="55959675">Indian Ocean</term>
     <term woeid="23424778">Sri Lanka</term>
     <term woeid="12493166">Sumatra</term>
   </places>
   <time>
     <term>2005</term>
   </time>
</topic>
```

We added to this topic the filter *timeType="any"* to consider any kind of temporal expression, even unknown ones, in terms of normalization. The places Sumatra and Sri Lanka were found near a question of kind "How long after" so were considered as filters.

Other topics like for example topic GeoTime-0014 question about places and dates result in a set of filters:

```
<topic id="GeoTime-0014">
 …
 <filterChain>
  <boolean type="AND">
    <term>
      <field>place</field>
      <value>Africa</value>
    </term>
    <term>
      <field>placeType</field>
      <value>country</value>
    </term>
    <term>
      <field>time</field>
      <value>2002</value>
    </term>
  </boolean>
</filterChain>
<terms>
  <desc>volcano erupt</desc>
  <narr>volcano erupted name volcano located</narr>
 </terms>
 <places>
  <term woeid="?">?</term>
 </places>
 <times>
  <term>?</term>
 </times>
</topic>
```

The original topic was:

"When and where did a volcano erupt in Africa during 2002?" narrative: *"date 2002 which volcano erupted Africa name volcano country located"*

For such topics we created a set of filters including a base filter to remove documents without temporal and geographic references, what is represented with the question mark.

2.4 Runs Description

We participated in GeoTime with 5 distinct strategies, all of them based on BM25. We used filters or query expansion when no filters were defined. If the topic requested places and/or dates, which was allays true, we also used a base filter to remove documents without geo and time expressions. If the extracted expressions were not considered filters, then they were used as query terms in special GeoTemporal indexes. The keywords component of the query was built using twice times the description, in order to increase its discriminatory power, and the narrative only once. In this sense, our term queries were composed by keywords, places and times using for that purpose several indexes in order to obtain an unique score. The BM25 model was used considering independent indexes. In first place we calculated the partial score of each term in its index, second we sum all term scores to obtain document score. We also used boost factors, detailed in 2.4.2 to 2.4.4, in each dimension of relevance. An issue that we must study in the future is the normalization of the scores per index, or more sophisticated techniques for fusion. Our approach for this experiment consisted in considering each index term independently. Meanwhile this approach could overstate the dimensions of relevance with more terms in the query. For example, if our query has 10 geographic names and 5 keywords, the geographic component of the score

will be, in theory, twice bigger than the keywords component. Also the independent scores depend of the index properties and it is not mathematically correct make a sum of scores calculated from different indexes. In this sense, if our approach is or not the best approach is subject of future research.

Bellow we detail each strategy. The numbers at the runs were assigned considering the predicted priority, for that reason our baseline have the lower priority of 5 considering 1 the more sophisticated technique.

- INESC-EN-EN-05-DN - Our first run used documents contents index and the base filter to remove documents without geographic or temporal expressions.
- INESC-EN-EN-04-DN - Our second run used sentences index and the base filter to remove sentences without geographic or temporal expressions. The document position was defined by it first sentence in the retrieved list of sentences. Other sentences of that document were ignored.
- INESC-EN-EN-03-DN - Our third run used document content index, the base filter and the filters defined in topic processing, or query expansion if no filters were defined. Bellow we detail this run.
- INESC-EN-EN-02-DN - Our fourth run used the sentences index, a base filter to remove sentences without geographic or temporal expressions and filters defined in topic processing but at sentence level, or query expansion when no filters were defined in the topic.
- INESC-EN-EN-01-DN - Our fifth run used a linear combination of document content and sentences indexes. The base filter, the topic filters and the query expansion was made at document level. The linear combination factors are detailed bellow.

2.4.1 Filters

The base filter made use of one index created to mark geo-temporal documents. In runs 03-DN, 04-DN and 05-DN we used three geographic filters: the *places,* the *belongTos* and the *placeTypes* containing the types of places detailed in Table 2. For temporal expressions we used three indexes, *timeExpressionsFormat*, *timePoints* and *expandedTimeDurations*. The first one indexed the formats presented in Table 6, the second indexed the key points of Table 9 and the third one indexed expanded time expressions that we generated from time periods including weeks and expressions described in Table 7.

2.4.2 Query Expansion

In topics with zero filters and zero geographic and temporal expressions we used blind relevance feedback query expansion. We based our method in Rochio algorithm, with modifications to use multiple indexes. This technique is detailed in [15]. We considered the geographic indexes *belongTos* and *places* with relative weights of 0.3 and 0.7 respectively. This was done because we think that belongTos is good for filtering but not so good for expansion because they are extensions that were not in the original text. We used temporal indexes *timePoints* and *expandedTimeDurationswith* relative weights of 0.7 and 0.3 respectively. We used a maximum number of 15 terms in the first 5 documents with a decay factor of 0.15. The topics where we used query expansion were 3, 5, 7, 8, 9, 10 and 12 in runs 03-DN, 02-DN and 01-DN.

2.4.3 Geographic and Temporal Terms

When places and time expressions were not assigned to filters we used them as query terms with relative weights similar to those

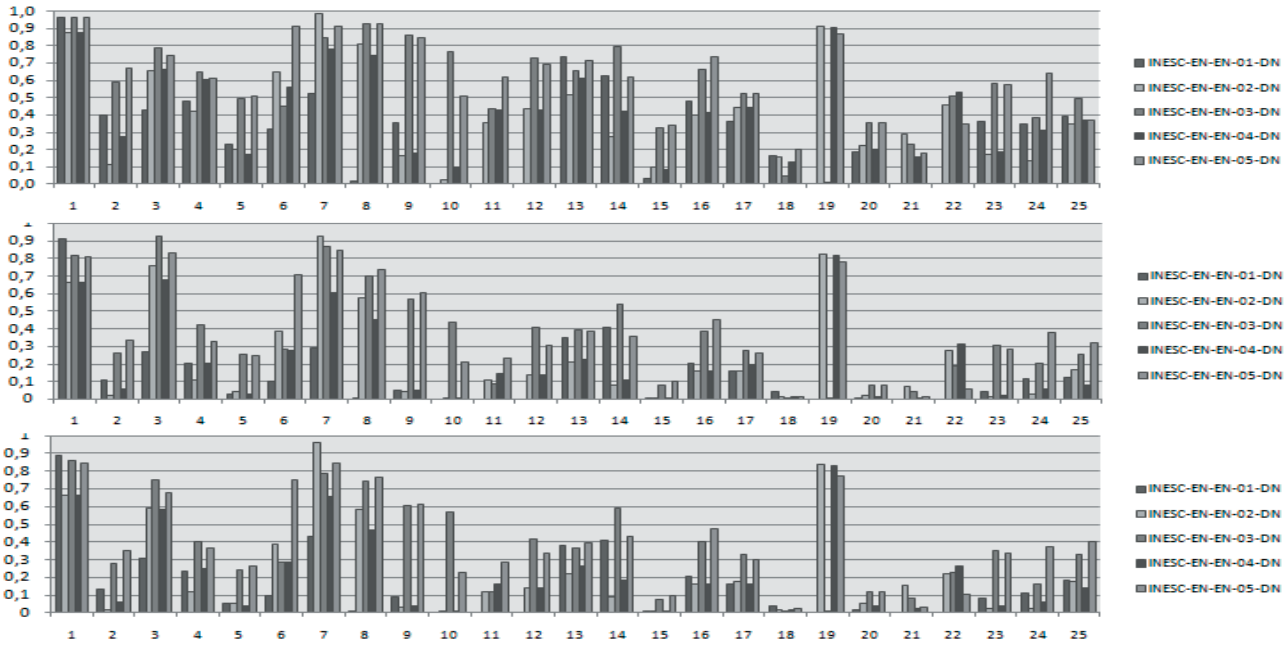

Figure 1 - Formal Results per Topic (from top to Bottom: NDCG, AP, Q)

used in query expansion. Time formats and place types were not used as query terms because they have only filtering purposes. Time durations have lower weight that time points. Once more belongTos were used but with lower priority. Example:

```
t_point:(2010*)^0.7 t_duration:(2010*)^0.3
g_place:(WOEID-2346573)^0.7 g_belongTos:(WOEID-2346573)^0.3
```

2.4.4 Combination Run

This run uses a mix of all others. The mix between sentences and contents was made using a linear combination with relative arbitrary weights of 0.7 to the sentence and 0.3 to the document content. Using the example of the previous subsections we will have:

```
(t_point:(2010*)^0.7 t_duration:(2010*)^0.3
g_place:(WOEID-213312)^0.7 g_belongTos:(WOEID-2346573)^0.3 )^0.3
(t_point_st:(2010*)^0.7 t_duration_st:(2010*)^0.3
g_place_st:( WOEID-2346573)^0.7 g_belongTos_st:( WOEID-2346573)^0.3 )^0.7
```

In the query note that the last part uses indexes ending with "*_st*", which means that we are using indexes at statement context.

3. RESULTS

We present the formal results means in Table 11. Our runs based on sentences produced poor results. The best run was the run with identifier INESC-EN-EN-05-DN that used only a base filter to remove documents without geographic and temporal expressions.

Table 11 – Formal Metrics Means

RUN	MAP	MQ	MNDCG
INESC-EN-EN-01-DN	0.137	0.153	0.2961
INESC-EN-EN-02-DN	0.232	0.233	0.4056
INESC-EN-EN-03-DN	0.352	0.364	0.5641
INESC-EN-EN-04-DN	0.213	0.222	0.4234
INESC-EN-EN-05-DN	0.387	0.407	0.6246

After analyzing the results topic by topic, we found that the sentences approach was a bad choice. The idea would be good if we had chosen the paragraph level. The majority of the relevant documents include our filtered terms in geographic and temporal dimensions of relevance but considering the paragraph level. At sentence granularity level many relevant results were omitted. Several of them have the relevant information in two followed sentences but not in the same. The sentence level seems to be very restrictive.

Looking to the special case of topic 25 we found a that the run 02-DN (filter at sentence level) return a relevant document given by our sentence. If we look to document 20050328.0205 we found the hit in Sentence 7: "Waves began hitting **Sri Lanka's** shores about two hours after the December quake; they struck **Sumatra within minutes**". The sentence 7 fulfills the two place filters and one temporal expression requirements. On other hand we found some problems in the follow fragments which are false positive examples took from the run INESC-EN-EN-01-DN results. These fragments were selected because the sentences have temporal expression beside they are not related with the topic question. A way to turn-around this problem is to consider only the duration expressions related with the sequence of two events, like:

*"... coastline of **Sri Lanka** ... no one will be thinking of marketing campaigns **for years** ..."*

*"... **Saturday**, U.S. helicopters ... to Sumatra's northwest coast, and cargo planes ... to **Sri Lanka**."*

This shows that our run using sentences and contents **(**01-DN) is very permissive in contrast with the runs 03-DN and 04-DN. The failure is due to the fact that we permit every cases giving high weights to the sentence occurrences of any term query. This makes think that the best option is between these two approaches.

Topics 15 and 18 return poor average precision results for all participants because for this topic were evaluated about 1000 documents, more documents than for the rest of the topics which the majority have about 500 evaluated documents each.

4. CONCLUSIONS

Our runs were not very successfully but we can set some future directions for GeoTemporal retrieval work. We learned that the sentence level is not a very good approach, but we have clues that the paragraph context resolution could probably improve the results of statements runs and possible reach the best run. Using geographic and temporal expressions base filters shows good performance, what reinforces our idea that probably it is possible to use these features in order to improve the results. Future work needs to address the topic processing because that could make the real difference independently of the technique used after that step. Temporal expressions extraction and representation is a very deep domain of research and many work need to be done in order to cover all cases; meanwhile we think that those kinds of features are strongly dependent of topic processing and interpretation.

5. ACKNOWLEDGEMENTS

This work was partially supported by FCT (INESC-ID multiannual funding) through the PIDDAC Program funds. We would also like to thanks to Professor Pável Calado and students Ricardo Vaz and Flávio Esteves, from the INESC-ID / Technical University of Lisbon at IST, for their valuable contribution in the assessments for the GeoTime English Task.

6. REFERENCES

[1] D. Ahn, J. van Rantwijk, and M. de Rijke. A Cascaded Machine Learning Approach to Interpreting Temporal Expressions. In: Proceedings NAACL-HLT 2007, April 2007.

[2] E. Saquete, R. Muñoz , P. Martínez-Barco. Event ordering using TERSEO system. Special issue: Application of natural language to information systems (NLDB04). Pages: 70 – 89, 2006.

[3] Omar Alonso, Michael Gertz, Ricardo Baeza-Yates. On the value of temporal information in information retrieval. ACM SIGIR Forum. Volume 41, Issue 2 (December 2007).

[4] Jorge Machado, Bruno Martins e José Borbinha, "LGTE: Lucene Extensions for Geo-Temporal Information Retrieval". GIIW, ECIR, Toulouse, 2009.

[5] T. Mandl et al. "An evaluation resource for Geographical Information Retrieval". In Proceedings of the 6th International Conference on Language Resources and Evaluation, 2008.

[6] F. Gey, at al. "GeoCLEF 2006: the CLEF 2006 Cross-Language Geographic Information Retrieval Track Overview". In A. Nardi, C. Peters, and J. L. Vicedo, editors,

Cross Language Evaluation Forum: Working Notes for the CLEF 2006 Workshop

[7] Ray R. Larson. “Cheshire at GeoCLEF 2008: Text and Fusion Approaches for GIR”, GeoCLEF 2008, CLEF 2008, Aarhus, Denmark, September 17-19, 2008.

[8] D. Ferres, H. Rodríguez. “TALP at GeoCLEF 2007: Results of a Geographical Knowledge Filtering Approach with Terrier”. CLEF, Budapest, Hungary, 2007

[9] Nuno Cardoso, Patrícia Sousa and Mário J. Silva, "The University of Lisbon at GeoCLEF 2008" GeoCLEF 2008, CLEF 2008, Aarhus, Denmark, September 17-19, 2008.

[10] Sato, N., Uehara, M., and Sakai, Y.: Temporal Ranking for Fresh Information Retrieval. Proceedings of the 6th International Workshop on Information Retrieval with Asian Languages. Sapporo, Japan, 2003, 116–123.

[11] Ferro Lisa, Inderjeet Mani, Beth Sundheim and George Wilson. TIDES Temporal Annotation Guidelines. Version 1.0.2 MITRE Technical Report, MTR 01W0000041. 2001.

[12] Roser Saur´ý, Jessica Littman, Bob Knippen, Robert Gaizauskas, Andrea Setzer, and James Pustejovsky. TimeML Annotation Guidelines. Version 1.2.1. January 31, 2006

[13] Fredric Gey†, Ray Larson, Noriko Kando, Jorge Machado, Tetsuya Sakai. NTCIR-GeoTime Overview: Evaluating Geographic and Temporal Search. NTCIR, Tokyo, Japan, 2010.

[14] James F. Allen. Towards a general theory of action and time. Artificial Intelligence 23, 2, July 1984.

[15] Jorge Machado, Bruno Martins and José Borbinha. Experiments with N-Gram Prefixes on a Multinomial Language Model versus Lucene’s off-the-shelf ranking scheme and Rocchio Query Expansion (TEL@CLEF Monolingual Task). ECDL/CLEF, Corfu, Greece, 2009.