A Language Modeling Approach for Temporal Information Needs

Klaus Berberich
Srikanta Bedathur
Omar Alonso
Gerhard Weikum

MPI–I–2010–5–001 January 2010

Authors' Addresses

Klaus Berberich
Max-Planck-Institut für Informatik
Campus E1 4
66123 Saarbrücken
Germany

Srikanta Bedathur
Max-Planck-Institut für Informatik
Campus E1 4
66123 Saarbrücken
Germany

Omar Alonso
Max-Planck-Institut für Informatik
Campus E1 4
66123 Saarbrücken
Germany

Gerhard Weikum
Max-Planck-Institut für Informatik
Campus E1 4
66123 Saarbrücken
Germany

Abstract

This work addresses information needs that have a temporal dimension conveyed by a temporal expression in the user's query. Temporal expressions such as "in the 1990s" are frequent, easily extractable, but not leveraged by existing retrieval models. One challenge when dealing with them is their inherent uncertainty. It is often unclear which exact time interval a temporal expression refers to.

We integrate temporal expressions into a language modeling approach, thus making them first-class citizens of the retrieval model and considering their inherent uncertainty. Experiments on the New York Times Annotated Corpus using Amazon Mechanical Turk to collect queries and obtain relevance assessments demonstrate that our approach yields substantial improvements in retrieval effectiveness.

Keywords

Temporal Expressions, Language Models, Crowdsourcing

Contents

1 Introduction

Many information needs have a temporal dimension as expressed by a temporal phrase contained in the user's query. Existing retrieval models, however, often do not provide satisfying results for such *temporal information needs*, as the following examples demonstrate:

- A sports journalist, interested in FIFA World Cup tournaments during the 1990s, issues the query fifa world cup 1990s. Documents such as the New York Times articles shown in Figure 1.1 would often not be found by existing retrieval models, despite their obvious relevance to the journalist's information need. Similarly, a document stating *France won the FIFA World Cup in 1998* or a document published in 1998 mentioning *FIFA World Cup final in July* would be missed. This is because existing retrieval models are not aware of the semantic connections between the temporal expressions "in 1998" and "in July" contained in the documents and the user's query temporal expression "1990s".

- A historian, doing research on Christianization, issues the query 13th century crusades. Documents with details on specific crusades, for instance, the Fourth Crusade that begun in 1202 would often not be among the retrieved results, unless they explicitly mention the 13th Century. Again, the reason is that existing retrieval models lack the knowledge about the semantic connections between temporal expressions like "from 1202 until 1204" and "in 1202" contained in documents and temporal expression "13th century" contained in the query.

Improving retrieval effectiveness for such temporal information needs is an important objective for several reasons. First, a significant percentage of queries has temporal information needs behind them – about 1.5% of web queries were found to contain an explicit temporal expression (as reported in [22]) and about 7% of web queries have an implicit temporal intent (as

WORLD CUP '98; Argentina Outlasts Short-Handed England

By JERE LONGMAN
Published: Wednesday, July 1, 1998

SIGN IN TO RECOMMEND
TWITTER

It began as a game tense with history and the possibility of violence and ended as the longest and most compelling match of the 1998 World Cup.

From sly nastiness to crampi
played on valiantly tonight t
periods of overtime. With the
second-round match as far a
of penalty kicks.

British fans were singing the
long stretches it appeared th
despite playing with only 10
Owen was the youngest play
scoring a goal for England w
the other.

WORLD CUP '98: France's Day of Soccer Glory Arrives; Upset of Brazil in World Cup

By JERE LONGMAN
Published: Monday, July 13, 1998

SIGN IN TO E-MAIL
SINGLE-PAGE
REPRINTS

The French soccer team, which had aroused an indifferent nation with its exceptional play over the past month, delivered an even more stunning accomplishment tonight, dominating favored Brazil to win
its first World Cu
3-0, France produ
championship ga

Zinedine Zidane,
goals of the comp
pair of corner kick
tournament, held

WORLD CUP '94; A 4th World Cup for Brazil

By LAWRIE MIFFLIN
Published: Monday, July 18, 1994

SIGN IN TO E-MAIL
REPRINTS
SHARE
ARTICLE TOOLS SPONSORED BY
IN THE PERFORMANCE OF A LIFETIME

"Making soccer history" was the theme United States organizers chose for the 15th World Cup, and yesterday Brazil played along, becoming the first nation to win the World Cup four times and the first to win it in a penalty-kick shootout. Brazil defeated Italy, also a three-time champion, in that one-on-one tie breaker after a scoreless tie in regulation time and 30 minutes of overtime.

As 94,194 people sat on the edge of their Rose Bowl seats in Pasadena, Calif., and some 2 billion television viewers worldwide grew white-knuckled, Brazil converted three penalty shots to Italy's two. Italy's star, Roberto Baggio, missed his team's final shot, sending it over the crossbar.

Figure 1.1: Documents published in New York Times relevant to the query fifa world cup 1990s but likely to be missed by existing retrieval models

reported in [21]). Notice that these numbers are based on general web queries – for *specific domains* (e.g., news or sports) or *expert users* (e.g., journalists or historians) we expect a larger fraction of queries to have a temporal information need behind them. Second, thanks to improved digitization techniques and preservation efforts, many document collections, including the Web, nowadays contain documents that (i) were *published a long time ago* and (ii) *refer to different times*. Consider, as one such document collection, the archive of the New York Times that covers the years 1851–2009. Articles in this archive provide a contemporary but also retrospective account on events during that time period. When searching these document archives, the temporal dimension plays an important role.

Temporal expressions are frequent across many kinds of documents and can be extracted and resolved at relative ease. However, it is not immediately clear how they should be integrated into a retrieval model. The key problem here is that the actual meaning of many temporal expressions is uncertain, or more specifically, it is not clear which exact time interval they actually refer to. As an illustration, consider the temporal expression "in 1998". Depending on context, it may refer to a particular day in that year, as in the above example about the FIFA World Cup final or to the year as a whole as in the sentence *in 1998 Bill Clinton was President of the United States*.

Our approach, in contrast to earlier work [10, 11, 17], considers this uncertainty. It integrates temporal expressions, in a principled manner, into a language modeling approach, thus making them first-class citizens of the retrieval model.

Contributions

In this work we make the following contributions: (i) a novel approach that integrates temporal expressions into a language model retrieval framework and (ii) a comprehensive experimental evaluation on the New York Times Annotated Corpus [3] and a snapshot of the English Wikipedia [7], as two real-world datasets, for which we leverage the crowdsourcing platform Amazon Mechanical Turk [1] to collect queries and obtain relevance assessments.

Organization

The rest of this report is organized as follows. Section 2 puts our work in context with existing related research. In Section 3 we introduce our model and notation. Section 4 describes how temporal expressions can be integrated into a language modeling approach. Conducted experiments and their results are described in Section 5. Finally, in Section 6, we conclude and point out promising open directions for future research.

2 Related Work

We now put our work in context with existing related research. The importance of temporal information for information retrieval is highlighted by Alonso et al. [8], who also mention the problem addressed in this chapter as one not yet satisfactorily supported by existing approaches. For our discussion of other related research, we broadly categorize it into the following three categories:

Time-Aware Retrieval Models

Li and Croft [20] and Dakka et al. [13] both propose language models that take into account publication times of documents, in order to favor, for instance, more recent documents. Kanahuba and Nørvåg [18] and de Jong et al. [14] employ language models to date documents, i.e., determine their publication time. Del Corso et al. [12] address the problem of ranking news articles, taking into account publication times but also their interlinkage. Jones and Diaz [16] focus on constructing query-specific temporal profiles based on the publication times of relevant documents. Thus, all of the approaches mentioned are based on the publication times of documents. None of the approaches, though, considers temporal expressions contained in the documents' contents.

Baeza-Yates [11] is the earliest approach that considers temporal expressions contained in documents for retrieval purposes. It aims at searching information that refers to the future. The proposed retrieval model is focused on confidences associated with statements about the future, thus favoring relevant documents that are confident about their predictions regarding a future time of interest. Kalczynski et al. [17] study the human perception of temporal expressions and propose a retrieval model for business news archives that takes into account temporal expressions. Arikan et al. [10] integrate temporal expressions into a language modeling approach but ignore the aspect of uncertainty. Metzler et al. [21], most recently, identify so-called implicitly temporal queries and propose a method to bias arbitrary ranking functions in favor of documents matching the user's implicit temporal intent – this work, in contrast, proposes a self-contained language modeling approach that seamlessly integrates temporal expressions.

Extraction of Temporal Expressions

The extraction of temporal expressions is a well-studied problem. We represent temporal expressions as quadruples to capture their inherent uncertainty – a formal representation that we adopt from Zhang et al. [25]. Koen and Bender [19] describe the Time Frames system that extracts temporal expressions and uses them to augment the user experience when reading news articles, for instance, by displaying a temporal context of concurrent events.

Several prototypes are available that make use of temporal expressions when searching the Web, most notably, Google's Timeline View [2] and TimeSearch [5]. Details about their internals, though, have not been published.

Crowdsourcing for IR Evaluation

Crowdsourcing platforms such as Amazon Mechanical Turk (AMT) are becoming a common tool for conducting experiments in information retrieval. AMT, as the best-known platform, allows requesters to publish so-called Human Intelligence Tasks (HITs), i.e., tasks that are hard for a computer but relatively easy for a human (e.g., determining the correct orientation of a photo). Apart from that, requesters can restrict the workers allowed to take up their HITs, for instance, based on the geographical location or depending on whether the worker passes a qualification test. On successful completion of a HIT, workers are paid a small reward that is typically below $0.10. For a discussion of benefits and guidelines on how to use crowdsourcing platforms for experiments in IR, we refer to Alonso et al. [9].

3 Model

We now lay out our formal model and our notation will be used in the following.

Time Domain & Temporal Expression Model

In this work, we apply a discrete notion of time and assume the integers $\mathbb{Z}$ as our *time domain* $\mathcal{T}$ with timestamps $t \in \mathcal{T}$ denoting the number of time units (e.g., milliseconds or days) passed (to pass) since (until) a reference time-point (e.g., the UNIX epoch). These time units will be referred to as *chronons* in the remainder. Our formal representation of temporal expressions is defined as:

Definition 3.1 (Temporal Expression) *A temporal expressions T is formally represented as a quadruple*

$$T = (\, tb_l,\, tb_u,\, te_l,\, te_u \,) \tag{3.1}$$

with tb_l, tb_u, te_l, $te_u \in \mathcal{T}$. The temporal expression T can refer to any time interval $[b, e]$ such that $b \in [tb_l, tb_u]$, $e \in [te_l, te_u]$ and $b \leq e$.

In our representation tb_l and tb_u are respectively a lower bound and upper bound for the begin boundary of the time interval – marking the time interval's earliest and latest possible begin time. Analogously, te_l and te_u are respectively a lower bound and upper bound for the end boundary of the time interval – marking the time interval's earliest and latest possible end time. Since the time interval is not necessarily known exactly, we hence capture lower and upper bounds for its boundaries. To give a concrete example, the temporal expression "in 1998" from the introduction is represented as

$$(\, \texttt{1998/01/01},\, \texttt{1998/12/31},\, \texttt{1998/01/01},\, \texttt{1998/12/31} \,)\,.$$

This representation thus captures the uncertainty inherent to many temporal expressions – a temporal expression T can refer to any time interval $[b, e]$

having a begin point $b \in [\, tb_l,\, tb_u \,]$ and an end point $e \in [\, te_l,\, te_u \,]$ along with the constraint $b \leq e$. We consider these time intervals thus as our *elementary units of meaning* in this work. In the remainder, when we refer to the temporal expression T, we implicitly denote the set of time intervals that T can refer to. Note that for notational convenience we use the format `YYYY/MM/DD` to represent chronons – their actual values are integers as described above.

Collection & Query Model

Let D denote our document collection. A document $d \in D$ is composed of its *textual part* d_{text} and its *temporal part* d_{time}. The textual part d_{text} is a bag of textual terms drawn from a vocabulary $\mathcal{V}$. The temporal part d_{time} is a bag of temporal expressions.

Analogously, a query q also consists of a textual part q_{text} and a temporal part q_{time}. We distinguish two modes of how we derive such a query from the user's input, which differ in how they treat temporal expressions extracted from the input. In the *inclusive* mode, the parts of the user's input that constitute a temporal expression are still included in the textual part of the query. In the *exclusive* mode, these are no longer included in the textual part. Thus, for the user input boston july 4 2002, as a concrete example, in the inclusive mode we obtain $q_{text} = \{\mathsf{boston}, \mathsf{july}, \mathsf{4}, \mathsf{2002}\}$, whereas we obtain $q_{text} = \{\mathsf{boston}\}$ in the exclusive mode.

4 Language Models for Temporal Information Needs

With our formal model and notation established, we now turn our attention to how temporal expressions can be integrated into a language modeling approach, and how we can leverage them to improve retrieval effectiveness for temporal information needs.

We use a query-likelihood approach and thus rank documents according to their estimated probability of generating the query. We assume that the textual and temporal part of the query q are generated independently from the corresponding parts of the document d, as captured in the following definition:

Definition 4.1 (Independent generation of query parts)

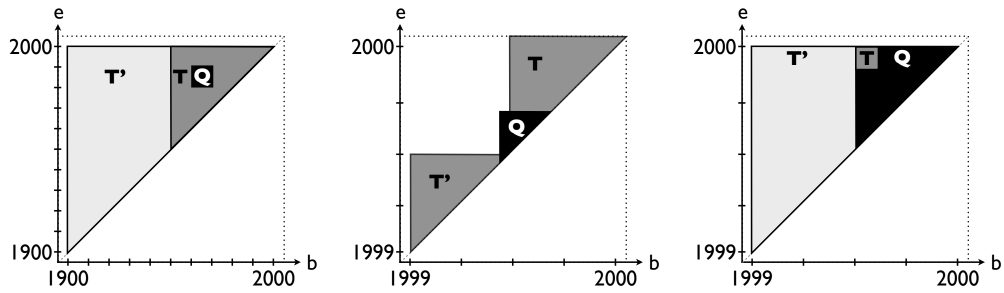
$$P(\, q \mid d \,) = P(\, q_{text} \mid d_{text} \,) \times P(\, q_{time} \mid d_{time} \,) \; . \tag{4.1}$$

The first factor $P(\, q_{text} | d_{text} \,)$ can be implemented using an existing text-based query-likelihood approach, e.g., the original Ponte and Croft model [23]. In our concrete implementation, which we describe in detail in Section 5, we employ a unigram language model with Jelinek-Mercer smoothing.

For the second factor in the above equation, we assume that query temporal expressions in q_{time} are generated independently from each other, i.e.,

$$P(\, q_{time} \mid d_{time} \,) = \prod_{Q \,\in\, q_{time}} P(\, Q \mid d_{time}) \; . \tag{4.2}$$

We use a two-step generative model to generate temporal expressions from a document d. In the first step, a temporal expression T is drawn at uniform random from the temporal expressions contained in the document. In the second step, a temporal expression is generated from the temporal expression T just drawn. Under this model, the probability of generating the query temporal expression Q from document d is defined as follows:



(a) Specificity (b) Coverage (c) Maximality

Figure 4.1: Three requirements for a generative model

Definition 4.2 (Generation of temporal expression from document)

$$P(\,Q \mid d_{time}) = \frac{1}{|d_{time}|} \sum_{T \in d_{time}} P(\,Q \mid T)\;. \tag{4.3}$$

In the rest of this section we describe two ways how the probability $P(\,Q \mid T\,)$ can be defined. Like other language modeling approaches, our model is prone to the zero-probability problem – if one of the query temporal expressions has zero probability of being generated from the document, the probability of generating the query from this document is zero. To mitigate this problem, we employ Jelinek-Mercer smoothing, and estimate the probability of generating the temporal expression Q from document d as

$$(1-\lambda)\cdot\frac{1}{|D_{time}|}\sum_{T \in D_{time}} P(\,Q \mid T) + \lambda\cdot\frac{1}{|d_{time}|}\sum_{T \in d_{time}} P(\,Q \mid T) \tag{4.4}$$

where $\lambda \in [0,1]$ is a tunable mixture parameter, and D_{time} refers to the temporal part of the document collection treated as a single very-large document.

Before giving two possible definitions of $P(\,Q \mid T\,)$, we identify the following requirements that any definition of $P(\,Q \mid T\,)$ must satisfy. Figure 4.1 illustrates these requirements – temporal expressions are represented a two-dimensional regions that encompass compatible combinations of begin point b and end point e.

Definition 4.3 (Specificity) *Given two temporal expressions T and T', we demand that*

$$|T \cap Q| = |T' \cap Q| \wedge |T| \leq |T'| \Rightarrow P(\,Q \mid T\,) \geq P(\,Q \mid T'\,)\;. \tag{4.5}$$

In other words, a query temporal expression is more likely to be generated from a temporal expression that closely matches it. Referring to Figure 4.1(a), the probability of generating Q (corresponding, e.g., to "from the 1960s until the 1980s") from T (corresponding, e.g., to "in the second half of the 20th century") is more than generating it from T' (corresponding, e.g., to "in the 20th century").

Definition 4.4 (Coverage) *Given two temporal expressions T and T', we demand that*

$$|T| = |T'| \wedge |T \cap Q| \leq |T' \cap Q| \Rightarrow P(\, Q \mid T\,) \leq P(\, Q \mid T'\,)\;. \tag{4.6}$$

In this requirement, we capture the intuition that a larger overlap with the query temporal expression is preferred. In Figure 4.1(b), the overlap of Q (corresponding, e.g., to "in the summer of 1999") with T (corresponding, e.g., to "in the first half of 1999") is more than the overlap with T' (corresponding, e.g., to "in the second half of 1999"). Therefore, the latter temporal expression is preferable and should have a higher probability of generating Q.

Definition 4.5 (Maximality) *The probability $P(\, Q \mid T\,)$ of generating Q from T should be maximal for $T = Q$, i.e.,*

$$T \neq Q \Rightarrow P(\, Q \mid T\,) \leq P(\, Q \mid Q\,)\;. \tag{4.7}$$

This requirement captures the intuition that the probability of generating a query temporal expression from a temporal expression matching it exactly must be the highest. As shown in Figure 4.1(c), the probability of generating Q (corresponding, e.g., to "in the second half of 1999") from itself should be higher than the probability of generating it from T (corresponding, e.g., to "from July 1999 until December 1999") or T' (corresponding, e.g., to "in 1999").

Uncertainty-Ignorant Language Model

Our first approach, further referred to as LmT, ignores the uncertainty inherent to temporal expressions. According to the following definition, a temporal expression T can only generate itself.

Definition 4.6 (LMT) *Let Q and T be temporal expressions,* LmT *defines the probability of generating Q from T as*

$$P(\, Q \mid T) = \mathbb{1}(T = Q)\;, \tag{4.8}$$

where $\mathbb{1}(T = Q)$ is an indicator function whose value assumes 1 iff $T = Q$ (i.e., $tb_l = qb_l \wedge tb_u = qb_u \wedge te_l = qe_l \wedge te_u = qe_u$).

The approach thus ignores uncertainty, since it misses the fact that a temporal expression T and a query temporal expression Q may refer to the same time interval, although $T \neq Q$.

Theorem 4.1 LmT *meets the requirements of specificity, coverage, and maximality defined above.*

Proof of Theorem 4.1 *We prove that specificity holds by showing the inverse direction*

$$\begin{aligned} P(Q \mid T) < P(Q \mid T') \quad & \Leftrightarrow \quad Q \neq T \wedge Q = T' \\ & \Leftrightarrow \quad (|T \cap Q| \neq |T' \cap Q| \wedge Q \neq T \wedge Q = T') \vee \\ & \qquad\ (|T \cap Q| = |T' \cap Q| \wedge Q \neq T \wedge Q = T') \\ & \Rightarrow \quad |T \cap Q| \neq |T' \cap Q| \vee |T| > |T'| \, . \end{aligned}$$

We prove that coverage holds by showing the inverse direction

$$\begin{aligned} P(Q \mid T) > P(Q \mid T') \quad & \Leftrightarrow \quad Q = T \wedge Q \neq T' \\ & \Leftrightarrow \quad (|T| \neq |T'| \wedge Q = T \wedge Q \neq T') \vee \\ & \qquad\ (|T| = |T'| \wedge Q = T \wedge Q \neq T') \\ & \Rightarrow \quad |T| \neq |T'| \vee |T \cap Q| > |T' \cap Q| \, . \end{aligned}$$

Finally, maximality holds for LmT*, since*

$$T \neq Q \Rightarrow P(Q \mid T) = 0 < P(Q \mid Q) = 1 \, .$$

□

Despite its simplicity the approach still profits from the extraction of temporal expressions. To illustrate this, consider the two temporal expressions "in the 1980s" and "in the '80s". Both share the same formal representation in our model, so that LmT can generate a query containing one of them from a document containing the other. In contrast, a text-based approach that does not pay special attention to temporal expressions, would not be aware of the semantic connection between the textual terms '80s and 1980s.

Uncertainty-Aware Language Model

As explained in the introduction, for many temporal expressions the exact time interval that they refer to is uncertain. Our second approach LmtU explicitly considers this uncertainty. In detail, we define the probability of generating Q from the document d as

$$P(\, Q \mid T\,) = \frac{1}{|Q|} \sum_{[q_b,\, q_e] \in Q} P(\, [q_b,\, q_e] \mid T\,) \, , \tag{4.9}$$

where the sum ranges over all time intervals included in Q. The approach thus assumes equal likelihood for each time interval $[q_b,\, q_e]$ that Q can refer to. Intuitively, each time interval that the user may have had in mind when uttering Q is assumed equally likely. Recall that $|Q|$ denotes the huge but finite total number of such time intervals.

The probability of generating the time interval $[q_b,\, q_e]$ from a temporal expression T is defined as

$$P(\,[q_b,\, q_e] \mid T\,) = \frac{1}{|T|}\, \mathbb{1}(\,[q_b,\, q_e] \in T\,)\;, \tag{4.10}$$

where $\mathbb{1}(\,[q_b,\, q_e] \in T\,)$ is an indicator function whose value is 1 iff $[q_b,\, q_e] \in T$. For T we thus also assume all time intervals that it can refer to as equally likely. Putting the two equations together we obtain

$$P(\,Q \mid T\,) = \frac{1}{|Q|} \sum_{[q_b,\, q_e] \in Q} \frac{1}{|T|}\, \mathbb{1}(\,[q_b,\, q_e] \in T\,)\;, \tag{4.11}$$

which can be simplified into the following concise definition of our uncertainty-aware language model.

Definition 4.7 (LMTU) *Let Q and T be temporal expressions,* LmtU *defines the probability of generating Q from T as*

$$P(\,Q \mid T\,) = \frac{|T \cap Q|}{|T| \cdot |Q|}\;. \tag{4.12}$$

Both Q and T are inherently uncertain. It is not clear which time interval the user issuing the query and author writing the document had in mind when uttering Q and T, respectively. Having no further information, our model assumes equal likelihood for all possible time intervals that Q and T respectively can refer to.

Theorem 4.2 LmtU *meets the requirements of specificity, coverage, and maximality defined above.*

Proof of Theorem 4.2 *For* LmtU *specificity and coverage follow immediately from Definition 4.7. To see that maximality holds, observe that $P(Q \mid Q) = 1/|Q|$ according to the above equation. Maximality then follows from the fact that*

$$T \neq Q \Rightarrow |T \cap Q|/|T| \leq 1\;.$$

□

Efficient Computation

For the practical applicability of this model, one important issue that needs addressing is the efficient computation of $P(\,Q\,|\,T\,)$ as defined above. Naïvely enumerating all time intervals that T and Q can refer to, before computing $|T \cap Q|$ is clearly not a practical solution. Consider again the temporal expression

$$(\,\texttt{1998/01/01},\,\texttt{1998/12/31},\,\texttt{1998/01/01},\,\texttt{1998/12/31}\,)\,.$$

For a temporal resolution with chronons corresponding to days the total number of time intervals that this temporal expression can refer to is $66,795$. When we make the granularity more fine-grained such that chronons correspond to hours, this number becomes $38,373,180$. Fortunately, there is a more efficient way to compute $P(\,Q\,|\,T\,)$, as we explain next.

Theorem 4.3 *The probability $P(\,Q\,|\,T\,)$ according to Definition 4.7 can be computed efficiently without enumerating all time intervals that Q respectively T can refer to.*

Proof of of Theorem 4.3 *We first show that $|T|$ can be computed efficiently for any temporal expression T. Let $T = (\,tb_l,\, tb_u,\, te_l,\, te_u\,)$ be a temporal expression, we distinguish two cases:*

(i) if $tb_u \leq te_l$ then $|T|$ can simply be computed as

$$(tb_u - tb_l + 1) \cdot (te_u - te_l + 1)\ ,$$

since any begin point b is compatible with any end point e, otherwise,

(ii) if $tb_u > te_l$ then $|T|$ can be computed as

$$|T| = \sum_{tb=tb_l}^{tb_u} (te_u - max(tb,\, te_l) + 1)\ , \tag{4.13}$$

which captures that only end points $e \geq max(b,\, te_l)$ are compatible with a fixed begin point b. Recall that we assume $tb_u > te_l$. This can be *simplified into a closed-form expression as follows:*

$$
\begin{aligned}
|T| &= \sum_{tb=tb_l}^{tb_u} (te_u - max(tb,\, te_l) + 1) \\
&= \sum_{tb=tb_l}^{te_l} (te_u - max(tb,\, te_l) + 1) + \sum_{tb=te_l+1}^{tb_u} (te_u - max(tb,\, te_l) + 1) \\
&= (te_l - tb_l + 1) \cdot (te_u - te_l + 1) + \sum_{tb=te_l+1}^{tb_u} (te_u - tb + 1) \\
&= (te_l - tb_l + 1) \cdot (te_u - te_l + 1) + \sum_{c=1}^{tb_u - te_l} (te_u - c - te_l + 1) \\
&= (te_l - tb_l + 1) \cdot (te_u - te_l + 1) \\
&\quad + (tb_u - te_l) \cdot (te_u - te_l + 1) - \sum_{c=1}^{tb_u - te_l} c \\
&= (te_l - tb_l + 1) \cdot (te_u - te_l + 1) \\
&\quad + (tb_u - te_l) \cdot (te_u - te_l + 1) - \; 0.5 \cdot (tb_u - te_l) \cdot (tb_u - te_l + 1)\ .
\end{aligned}
\tag{4.14}
$$

Let $Q = (\, qb_l,\, qb_u,\, qe_l,\, qe_u\,)$ be a query temporal expression. We can compute $|Q|$ using our preceding arguments. For computing $|Q \cap T|$ notice that each time interval $[b,\, e] \in Q \cap T$ fulfills $b \in [tb_l,\, tb_u] \cap [qb_l,\, qb_u]$ and $e \in [te_l,\, te_u] \cap [qe_l,\, qe_u]$. Therefore, $|T \cap Q|$ can be computed by considering the temporal expression

$$(\, max(tb_l, qb_l),\, min(tb_u, qb_u),\, max(te_l, qe_l),\, min(te_u, qe_u) \,)\ .$$

$\square$

Thus, we have shown that the generative model underlying LmtU allows for efficient computation. When processing a query with a query temporal expression Q, we need to examine all temporal expressions T with $T \cap Q \neq \emptyset$ and the documents that contain them. This can be implemented efficiently by keeping a small inverted index in main memory that keeps track of the documents that contain a specific temporal expression. Its lexicon, which consists of the known temporal expressions, can be organized using an interval tree to support the efficient identification of qualifying temporal expressions via interval intersection.

5 Experimental Evaluation

We next present the experimental evaluation of the proposed approaches.

Setup & Datasets

Methods under Comparison

In our experimental evaluation we compare the following methods:

- LM(γ) – Unigram language model with Jelinek-Mercer smoothing
- LMT-IN(γ, λ) – Uncertainty-ignorant method using inclusive mode
- LMT-EX(γ, λ) – Uncertainty-ignorant method using exclusive mode
- LMTU-IN(γ, λ) – Uncertainty-aware method using inclusive mode
- LMTU-EX(γ, λ) – Uncertainty-aware method using exclusive mode

Apart from our baseline LM, we thus consider all four combinations of (a) inclusive vs. exclusive mode (i.e., whether query terms constituting a temporal expression are part of q_{text}) and (b) uncertainty-ignorant vs. uncertainty-aware definition of $P(\,Q \mid T\,)$. The mixture parameters γ and λ control the Jelinek-Mercer smoothing used when generating the textual part and the temporal part of the query, respectively. We consider values in $\{0.25, 0.5, 0.75\}$ for each of them, giving us a total of 39 method configurations under comparison. Further, notice that our baseline LM, which is not aware of temporal expressions, always only considers q_{text} as determined using the inclusive mode, i.e., containing all terms from the user's input.

Implementation Details

We implemented all methods in Java 1.6 keeping data in an Oracle 11g database. All experiments described below were run on a single SUN V40z machine having four AMD Opteron CPUs, 16GB RAM, a large network-attached RAID-5 disk array, and running Microsoft Windows Server 2003. When processing the two document collections, we did not remove stopwords nor apply lemmatization/stemming. Temporal expressions were extracted using TARSQI [24]. TARSQI detects and resolves temporal expressions using a combination of hand-crafted rules and machine learning. It annotates a given input document using the TimeML [4] markup language. Building on TARSQI's output, we extracted range temporal expressions such as "from 1999 until 2002", which TARSQI does not yet support. Further, we added each article's publication date as an additional temporal expression. We map temporal expressions to our quadruple representation using milliseconds as chronons and the UNIX epoch (i.e., midnight of January 1, 1970) as our reference time-point.

Document Collections

We use two publicly-available document collections for our experimental evaluation, namely:

- *New York Times Annotated Corpus* [3] (NYT) that contains $1,855,656$ articles published in New York Times between 1987 and 2007.
- *The English Wikipedia* [7] (WIKI) as of July 7, 2009 that contains a total of $2,955,294$ encyclopedia articles.

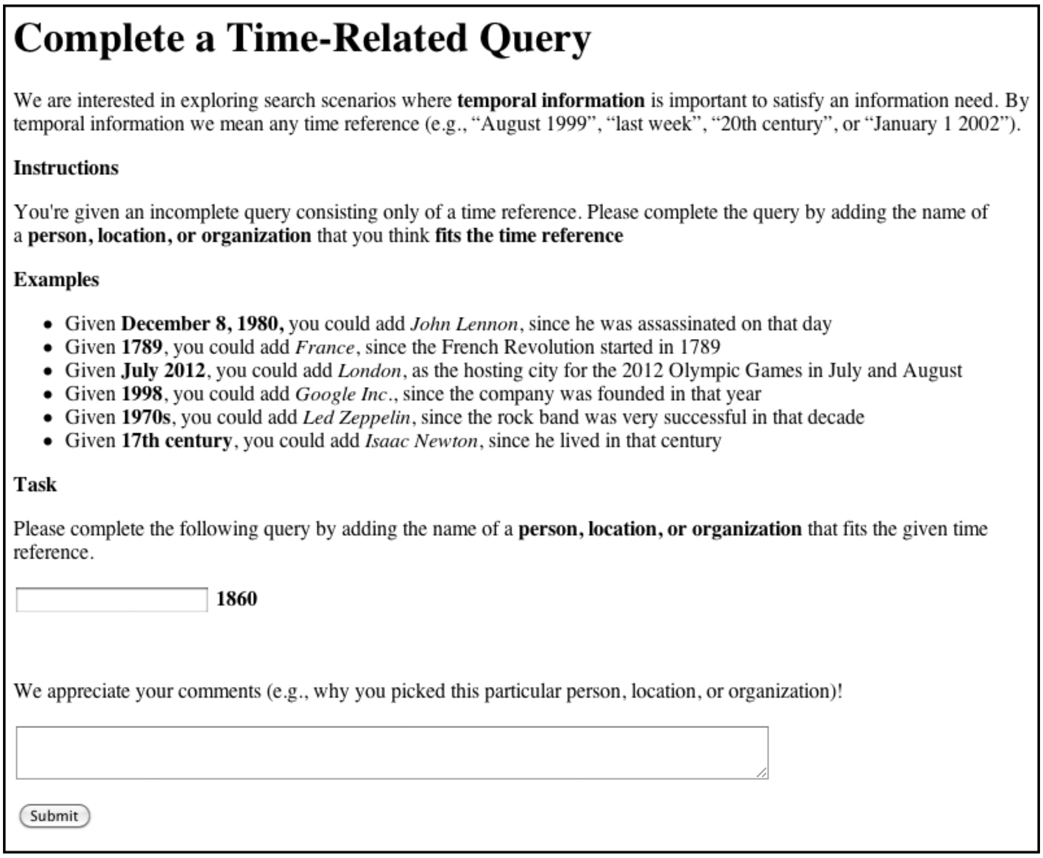
	NYT	WIKI
# Documents	1,855,656	2,955,294
Document Length (μ)	691.79	617.18
Document Length (σ)	722.88	1101.51
# Temporal Expressions per Document (μ)	6.35	12.91
# Temporal Expressions per Document (σ)	5.86	33.20

Table 5.1: Document collection statistics (with mean μ and standard deviation σ)

Table 5.1 shows additional statistics about the two datasets. From the figures we observe that the mean document length is similar for both datasets. Documents from WIKI, on average, contain more than twice as many temporal expressions as documents from NYT.

Queries

Since we target a specific class of information needs, query workloads used in benchmarks like TREC [6] are unemployable in our setting. Search-engine



Complete a Time-Related Query

We are interested in exploring search scenarios where **temporal information** is important to satisfy an information need. By temporal information we mean any time reference (e.g., “August 1999”, “last week”, “20th century”, or “January 1 2002”).

Instructions

You're given an incomplete query consisting only of a time reference. Please complete the query by adding the name of a **person, location, or organization** that you think **fits the time reference**

Examples

- Given **December 8, 1980,** you could add *John Lennon*, since he was assassinated on that day
- Given **1789**, you could add *France*, since the French Revolution started in 1789
- Given **July 2012**, you could add *London*, as the hosting city for the 2012 Olympic Games in July and August
- Given **1998**, you could add *Google Inc.*, since the company was founded in that year
- Given **1970s**, you could add *Led Zeppelin*, since the rock band was very successful in that decade
- Given **17th century**, you could add *Isaac Newton*, since he lived in that century

Task

Please complete the following query by adding the name of a **person, location, or organization** that fits the given time reference.

1860

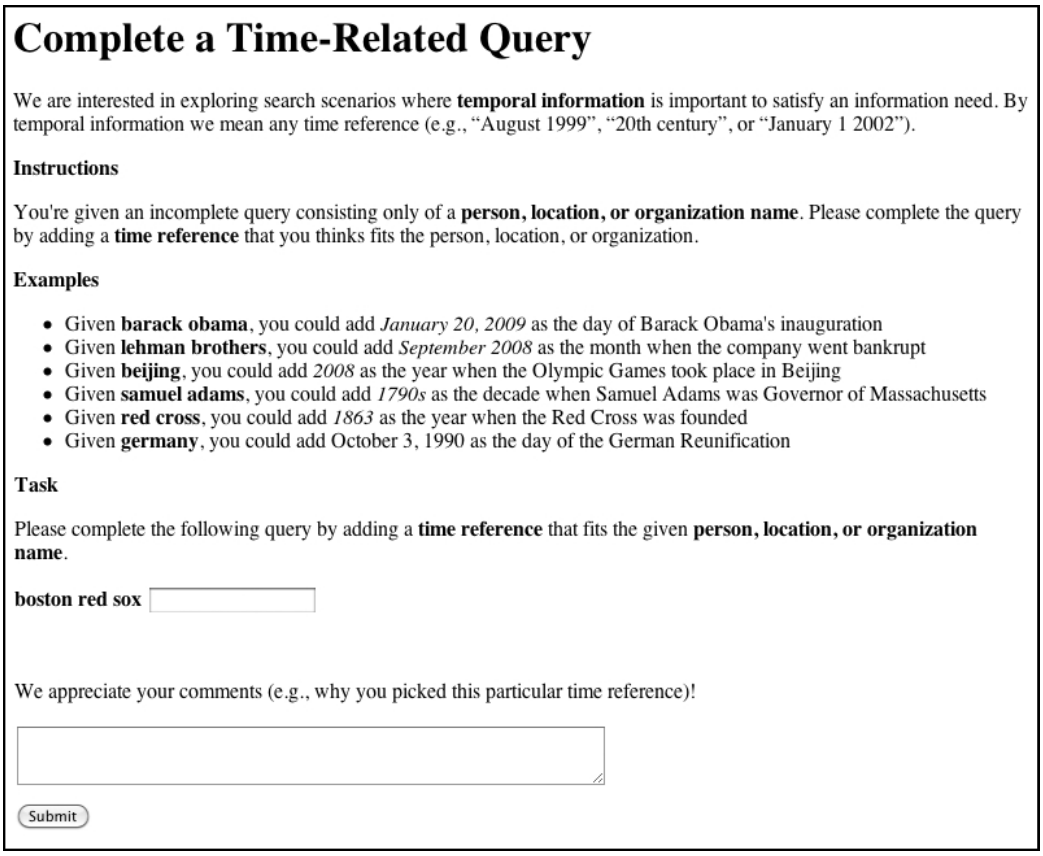
We appreciate your comments (e.g., why you picked this particular person, location, or organization)!

Submit

Figure 5.1: Amazon Mechanical Turk HIT to collect queries by letting users fill in an entity that fits a given temporal expression

query logs, on the other hand, as a second valuable source of realistic queries, are typically not publicly available. To assemble a query workload that captures users' interests and preferences, we therefore ran two user studies on AMT. In our first study, workers were provided with an entity related to one of the topics *Sports*, *Culture*, *Technology*, or *World Affairs* and asked to specify a temporal expression that fits the given entity. In our second study, users were shown a temporal expression corresponding to a *Day*, *Month*, *Year*, *Decade*, or *Century* and asked to add an entity related to one of the aforementioned topics. Figure 5.1 and Figure 5.2 show screenshots of our HITs. We asked users in both studies to comment on why they chose their particular answer. Examples of comments that we obtained are:

- boston red sox [october 27, 2004]: *Won 6th World Championship.*
- sewing machine [1850s]: *Isaac Singer invented the sewing machine, then*



Complete a Time-Related Query

We are interested in exploring search scenarios where **temporal information** is important to satisfy an information need. By temporal information we mean any time reference (e.g., "August 1999", "20th century", or "January 1 2002").

Instructions

You're given an incomplete query consisting only of a **person, location, or organization name**. Please complete the query by adding a **time reference** that you thinks fits the person, location, or organization.

Examples

- Given **barack obama**, you could add *January 20, 2009* as the day of Barack Obama's inauguration
- Given **lehman brothers**, you could add *September 2008* as the month when the company went bankrupt
- Given **beijing**, you could add *2008* as the year when the Olympic Games took place in Beijing
- Given **samuel adams**, you could add *1790s* as the decade when Samuel Adams was Governor of Massachusetts
- Given **red cross**, you could add *1863* as the year when the Red Cross was founded
- Given **germany**, you could add October 3, 1990 as the day of the German Reunification

Task

Please complete the following query by adding a **time reference** that fits the given **person, location, or organization name**.

boston red sox

We appreciate your comments (e.g., why you picked this particular time reference)!

Submit

Figure 5.2: Amazon Mechanical Turk HIT to collect queries by letting users fill in a temporal expression that fits a given entity

patented the motor for a sewing machine later in that decade.

- berlin [october 27, 1961]: *Tank standoff at Checkpoint Charlie.*
- chicago bulls [1991]: *The Bulls won the NBA Finals that year.*
- wright brothers [1905]: *The Wright brothers were starting out somewhere around that time.*

Among the queries obtained from our user studies, we selected the 40 queries shown in Figure 5.3. Queries are categorized according to their topic and temporal granularity, giving us a total of 20 query categories, each of which contains two queries.

	Sports	Culture
Day	boston red sox [october 27, 2004] ac milan [may 23, 2007]	kurt cobain [april 5, 1994] keith harring [february 16, 1990]
Month	stefan edberg [july 1990] italian national soccer team [july 2006]	woodstock [august 1994] pink floyd [march 1973]
Year	babe ruth [1921] chicago bulls [1991]	rocky horror picture show [1975] michael jackson [1982]
Decade	michael jordan [1990s] new york yankees [1910s]	sound of music [1960s] mickey mouse [1930s]
Century	la lakers [21st century] soccer [21st century]	academy award [21st century] jazz music [21st century]
	Technology	**World Affairs**
Day	mac os x [march 24, 2001] voyager [september 5, 1977]	berlin [october 27, 1961] george bush [january 18, 2001]
Month	thomas edison [december 1891] microsoft halo [june 2000]	poland [december 1970] pearl harbor [december 1941]
Year	roentgen [1895] wright brothers [1905]	nixon [1970s] iraq [2001]
Decade	internet [1990s] sewing machine [1850s]	vietnam [1960s] monica lewinsky [1990s]
Century	musket [16th century] siemens [19th century]	queen victoria [19th century] muhammed [7th century]

Figure 5.3: Queries categorized according to topic and temporal granularity

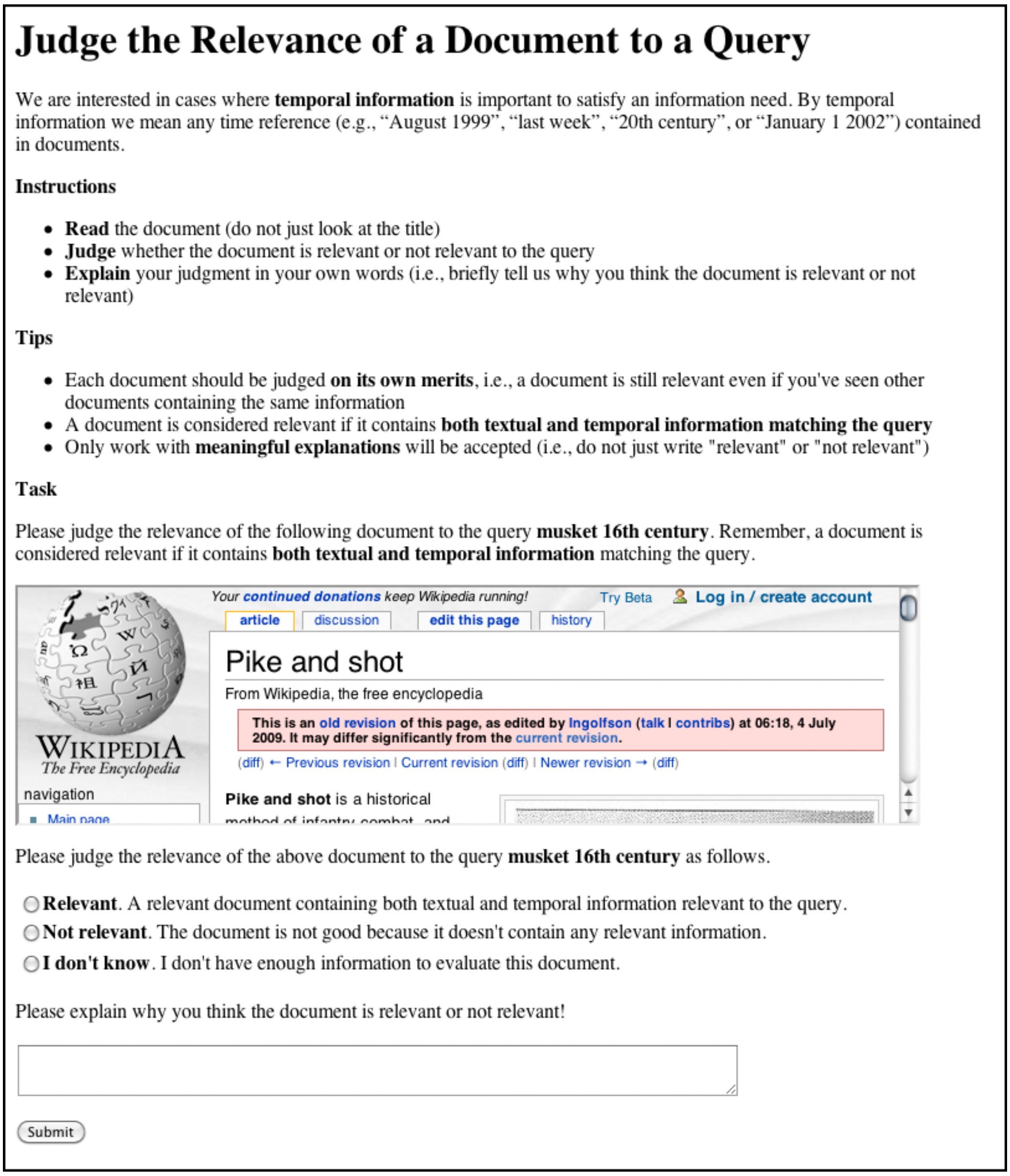
Relevance Assessments

Relevance assessments were also collected using AMT. Figure 5.4 shows a screenshot of our HIT. We computed top-10 query results for each query and each method configuration under comparison, pooled them, yielding a total of $1,251$ query-document pairs on NYT and $1,220$ query-document pairs on WIKI. Each of these query-document pairs was assessed by five workers on AMT. Workers could state whether they considered the document *relevant* or *not relevant* to the query. To prevent spurious assessments, a third option (coined *I don't know*) was provided, which workers should select if they had insufficient information or knowledge to assess the document's relevance. Further, we asked workers to explain in their own words, why the document was relevant or not relevant. We found the feedback provided through the explanations extremely insightful. Examples of provided explanations are:

- roentgen [1895]: *Wilhelm Roentgen was alive in 1895 when the building in New York at 150 Nassau Street in downtown Manhattan, NYC was built, they do not ever intersect other than sharing the same timeline of existence for a short while.*

- nixon [1970s]: *This article is relevant. It is a letter to the editor in response to a column about 1970s-era Nixon drug policy.*

- keith harring [february 16, 1990]: *The article does not have any information on Keith Harring, only Laura Harring. Though it contains*



Judge the Relevance of a Document to a Query

We are interested in cases where **temporal information** is important to satisfy an information need. By temporal information we mean any time reference (e.g., “August 1999”, “last week”, “20th century”, or “January 1 2002”) contained in documents.

Instructions

- **Read** the document (do not just look at the title)
- **Judge** whether the document is relevant or not relevant to the query
- **Explain** your judgment in your own words (i.e., briefly tell us why you think the document is relevant or not relevant)

Tips

- Each document should be judged **on its own merits**, i.e., a document is still relevant even if you've seen other documents containing the same information
- A document is considered relevant if it contains **both textual and temporal information matching the query**
- Only work with **meaningful explanations** will be accepted (i.e., do not just write "relevant" or "not relevant")

Task

Please judge the relevance of the following document to the query **musket 16th century**. Remember, a document is considered relevant if it contains **both textual and temporal information** matching the query.

Your ***continued donations*** *keep Wikipedia running!* Try Beta **Log in / create account**

article discussion **edit this page** history

Wikipedia
The Free Encyclopedia
navigation
Main page

Pike and shot

From Wikipedia, the free encyclopedia

This is an old revision of this page, as edited by Ingolfson (talk | contribs) at 06:18, 4 July 2009. It may differ significantly from the current revision.

(diff) ← Previous revision | Current revision (diff) | Newer revision → (diff)

Pike and shot is a historical

Please judge the relevance of the above document to the query **musket 16th century** as follows.

◯ **Relevant**. A relevant document containing both textual and temporal information relevant to the query.

◯ **Not relevant**. The document is not good because it doesn't contain any relevant information.

◯ **I don't know**. I don't have enough information to evaluate this document.

Please explain why you think the document is relevant or not relevant!

Submit

Figure 5.4: Amazon Mechanical Turk HIT to collect relevance assessments

the keywords Harring and 1990, the article is obviously not what the searcher is looking for.

Apart from that, when having to explain their assessment, workers seemed more thorough in their assessments. Per completely assessed query-document pair we paid \$0.02 to workers. For the relevance assessments on NYT, workers chose relevant for 33%, not relevant for 63%, and the third option (i.e., *I don't know*) for 4% of the total 6,255 relevance assessments. On WIKI, workers chose relevant for 35%, not relevant for 62%, and the third option (i.e., *I don't know*) for 3% of the total 6,100 relevance assessments. Relevance assessments with the last option are ignored when computing retrieval-effectiveness measures below. To measure the degree of agreement between assessors, we computed the Fleiss' κ statistic [15] of 0.36 and 0.40 on NYT and WIKI respectively, indicating a fair degree of agreement between assessors.

For the sake of reproducibility, annotated temporal expressions, queries, and relevance assessments are made available at:
`http://www.mpi-inf.mpg.de/~kberberi/ecir2010`

Experimental Results

We measure the retrieval effectiveness of the methods under comparison using Precision at k (P@k) and nDCG at k (N@k) as two standard measures. When computing P@k, we employ majority voting. A document is thus considered relevant to a query, if the majority of workers assessed it as relevant. When computing N@k, the average relevance grade assigned by workers is determined interpreting *relevant* as grade 1 and *not relevant* as grade 0.

	P@5	**N@5**	**P@10**	**N@10**
LM ($\gamma = 0.25$)	0.33	0.34	0.30	0.32
LM ($\gamma = 0.75$)	0.38	0.39	0.37	0.38
LMT-IN ($\gamma = 0.25$, $\lambda = 0.75$)	0.26	0.27	0.23	0.25
LMT-IN ($\gamma = 0.75$, $\lambda = 0.75$)	0.29	0.31	0.25	0.28
LMT-EX ($\gamma = 0.25$, $\lambda = 0.75$)	0.36	0.36	0.32	0.33
LMT-EX ($\gamma = 0.5$, $\lambda = 0.75$)	0.37	0.37	0.32	0.33
LMTU-IN ($\gamma = 0.25$, $\lambda = 0.75$)	0.41	0.42	0.37	0.37
LMTU-IN ($\gamma = 0.75$, $\lambda = 0.25$)	0.44	0.44	0.39	0.40
LMTU-EX ($\gamma = 0.25$, $\lambda = 0.75$)	0.53	0.51	0.49	0.49
LMTU-EX ($\gamma = 0.5$, $\lambda = 0.75$)	**0.54**	**0.52**	**0.51**	**0.49**

Table 5.2: Retrieval effectiveness overall on New York Times

	P@5	N@5	P@10	N@10
LM ($\gamma = 0.25$)	0.47	0.46	0.42	0.43
LM ($\gamma = 0.75$)	0.52	0.49	0.51	0.48
LMT-IN ($\gamma = 0.25$, $\lambda = 0.75$)	0.39	0.37	0.29	0.31
LMT-IN ($\gamma = 0.75$, $\lambda = 0.75$)	0.40	0.38	0.33	0.34
LMT-EX ($\gamma = 0.25$, $\lambda = 0.75$)	0.41	0.38	0.35	0.34
LMT-EX ($\gamma = 0.75$, $\lambda = 0.75$)	0.43	0.39	0.36	0.36
LMTU-IN ($\gamma = 0.25$, $\lambda = 0.75$)	0.50	0.48	0.44	0.43
LMTU-IN ($\gamma = 0.75$, $\lambda = 0.75$)	0.54	0.50	0.48	0.46
LMTU-EX ($\gamma = 0.25$, $\lambda = 0.75$)	0.57	0.51	0.54	0.50
LMTU-EX ($\gamma = 0.75$, $\lambda = 0.75$)	**0.60**	**0.53**	**0.56**	**0.51**

Table 5.3: Retrieval effectiveness overall on Wikipedia

Overall Retrieval Performance

Table 5.2 and Table 5.3 give retrieval-effectiveness figures computed using all queries and cut-off levels $k = 5$ and $k = 10$ on NYT and WIKI, respectively. For each of the five methods under comparison, the tables show the best-performing and worst-performing configuration with their corresponding values for the mixture parameters γ and λ.

The figures shown support the following observations: (i) on WIKI all methods achieve slightly higher retrieval effectiveness than on NYT, (ii) on both datasets the exclusive mode outperforms the inclusive mode for both both LMT and LMTU, (iii) LMT does not yield an improvement over the baseline LM but even reduces retrieval effectiveness, (iv) LMTU is at par with the baseline LM when the inclusive mode is used and outperforms it significantly when used with the exclusive mode. For LMTU-EX the worst configuration beats the best configuration of the baseline. Further, the worst and best configuration of LMTU-EX are close to each other demonstrating the method's robustness.

	Sports		Culture		Technology		World Affairs	
	P@10	N@10	P@10	N@10	P@10	N@10	P@10	N@10
LM	0.33	0.33	0.39	0.38	0.27	0.32	0.50	0.49
LMT-IN	0.36	0.36	0.25	0.30	0.10	0.15	0.30	0.30
LMT-EX	0.46	0.44	0.33	0.34	0.12	0.17	0.38	0.38
LMTU-IN	0.46	0.44	0.41	0.42	0.21	0.27	0.48	0.48
LMTU-EX	**0.67**	**0.58**	**0.47**	**0.49**	**0.29**	**0.34**	**0.60**	**0.57**

Table 5.4: Retrieval effectiveness by topic on New York Times

Retrieval Performance by Topic

For the best-performing configuration of each method (as given in Table 5.2 and Table 5.3), we compute retrieval-effectiveness measures at cut-off level $k = 10$ and group them by topic.

	Sports		Culture		Technology		World Affairs	
	P@10	**N@10**	**P@10**	**N@10**	**P@10**	**N@10**	**P@10**	**N@10**
LM	0.40	0.38	0.53	0.49	0.52	0.50	**0.57**	**0.54**
LMT-IN	0.28	0.29	0.33	0.34	0.33	0.31	0.36	0.40
LMT-EX	0.32	0.34	0.36	0.34	0.37	0.34	0.37	0.41
LMTU-IN	0.39	0.40	0.53	0.48	0.48	0.45	0.49	0.51
LMTU-EX	**0.54**	**0.48**	**0.57**	**0.51**	**0.61**	**0.54**	0.52	0.52

Table 5.5: Retrieval effectiveness by topic on Wikipedia

For NYT the resulting figures are shown in Table 5.4 and support our above observations. Thus, LMTU-EX consistently achieves the highest retrieval effectiveness across all topics. Further, we observe that all methods perform worst on queries from *Technology*. The best performance varies per method and measure.

Table 5.5 shows the resulting figures for WIKI. Here, LMTU-EX performs best on three of the four topics, but achieves retrieval-effectiveness scores slightly lower than those of the baseline LM on queries from *World Affairs*.

	Day		Month		Year	
	P@10	**N@10**	**P@10**	**N@10**	**P@10**	**N@10**
LM	0.35	0.38	0.42	0.40	0.65	0.59
LMT-IN	0.18	0.22	0.20	0.21	0.55	0.50
LMT-EX	0.26	0.28	0.24	0.25	0.58	0.55
LMTU-IN	0.33	0.36	0.47	0.46	0.59	0.56
LMTU-EX	**0.43**	**0.44**	**0.50**	**0.50**	**0.69**	**0.64**

	Decade		Century	
	P@10	**N@10**	**P@10**	**N@10**
LM	0.20	0.28	0.25	0.26
LMT-IN	0.23	0.30	0.20	0.24
LMT-EX	0.28	0.33	0.31	0.32
LMTU-IN	0.34	0.35	0.24	0.27
LMTU-EX	**0.56**	**0.54**	**0.36**	**0.35**

Table 5.6: Retrieval effectiveness by temporal granularity on New York Times

Retrieval Performance by Temporal Granularity

In analogy, we group retrieval-effectiveness measurements at cut-off level $k = 10$ by temporal granularity – again considering only the best-performing configuration of each method.

Table 5.6 gives the resulting figures for NYT. LMTU-EX consistently achieves the best retrieval performance. Apart from that, we observe significant variations in retrieval effectiveness across temporal granularities for the baseline LM. For queries that include a year, all methods achieve their best performance on NYT. The worst performance varies per method and measure.

	Day		Month		Year	
	P@10	**N@10**	**P@10**	**N@10**	**P@10**	**N@10**
Lm	0.35	0.37	**0.55**	**0.49**	**0.75**	**0.63**
LmT-IN	0.11	0.17	0.10	0.16	0.66	0.60
LmT-EX	0.11	0.18	0.10	0.16	0.70	0.60
LmtU-IN	**0.43**	**0.42**	0.44	0.48	0.66	0.60
LmtU-EX	0.34	0.38	0.45	0.46	0.71	0.61

	Decade		Century	
	P@10	**N@10**	**P@10**	**N@10**
Lm	0.50	0.50	0.38	0.39
LmT-IN	0.50	0.47	0.25	0.28
LmT-EX	0.55	0.51	0.31	0.33
LmtU-IN	0.51	0.49	0.30	0.30
LmtU-EX	**0.71**	**0.63**	**0.59**	**0.48**

Table 5.7: Retrieval effectiveness by temporal granularity on Wikipedia

For WIKI the resulting figures given in Table 5.7 show a less distinct picture. Thus, for queries containing a month or a year, the baseline Lm achieves the best retrieval effectiveness, although LmtU-EX is close behind. LmtU-EX clearly outperforms the baseline Lm for queries containing a decade or a century. Interestingly, for queries that include a day, LmtU-IN achieves the best performance.

Summary

Our experimental evaluation leads us to the following findings. When assessed on the whole of queries, LmtU consistently achieves superior retrieval performance on both datasets. The uncertainty-ignorant LmT model, in contrast, deteriorates retrieval performance in comparison to the baseline. For both methods, the exclusive mode of deriving the query from the user's input performs better than its inclusive counterpart. In summary, (i) considering the uncertainty inherent to temporal expressions is essential and (ii) excluding terms that constitute a temporal expression from the textual part of the query is beneficial.

6 Discussion & Outlook

In this work, we have developed a novel approach that integrates temporal expressions into a language model retrieval framework, taking into account the uncertainty inherent to temporal expressions. Comprehensive experiments on the New York Times Annotated Corpus and a snapshot of the English Wikipedia, as two publicly-available large-scale document collections, with relevance assessments obtained using Amazon Mechanical Turk showed that our approach substantially improves retrieval effectiveness for temporal information needs.

Outlook

Our focus in this work has been on temporal information needs disclosed by an *explicit* temporal expression in the user's query.

Often, as somewhat explored in [21], queries may not contain such an explicit temporal expression, but still have an associated *implicit* temporal intent. Consider a query such as bill clinton arkansas that is likely to allude to Bill Clinton's time as Governor of Arkansas between 1971 and 1981. Detecting and dealing with such queries is an interesting direction for future research.

Apart from that, temporal information contained in documents may be valuable, when trying to provide diverse query results. Even for queries that do not have a temporal intent behind them, the user profits from a result documents that *contain diverse temporal expressions*. Thus, for a query such as vincent van gogh, a set of result documents discussing different periods in the famous painter's life is preferable to a set of documents focused on his final years. Leveraging temporal expressions as a source for result diversification is another interesting direction for future research.

Bibliography

[1] Amazon Mechanical Turk
http://www.mturk.com.

[2] Google's Timeline View
http://www.google.com/experimental/.

[3] New York Times Annotated Corpus
http://corpus.nytimes.com.

[4] TimeML Specification Language
http://www.timeml.org.

[5] TimeSearch History
http://www.timesearch.info.

[6] Text REtrievel Conference
http://trec.nist.gov.

[7] Wikipedia
http://www.wikipedia.org.

[8] O. Alonso, M. Gertz, and R. Baeza-Yates. On the value of temporal information in information retrieval. *SIGIR Forum*, 41(2):35–41, 2007.

[9] O. Alonso, D. E. Rose, and B. Stewart. Crowdsourcing for relevance evaluation. *SIGIR Forum*, 42(2):9–15, 2008.

[10] I. Arikan, S. Bedathur, and K. Berberich. Time Will Tell: Leveraging Temporal Expressions in IR. In *WSDM (Late Breaking-Results)*, 2009.

[11] R. A. Baeza-Yates. Searching the future. In *Proceedings of ACM SIGIR Workshop MF/IR*, 2005.

[12] G. M. D. Corso, A. Gulli, and F. Romani. Ranking a stream of news. In *WWW '05: Proceedings of the 14th international conference on World Wide Web*, pages 97–106, New York, NY, USA, 2005. ACM Press.

[13] W. Dakka, L. Gravano, and P. G. Ipeirotis. Answering general time sensitive queries. In *CIKM '08: Proceeding of the 17th ACM conference on Information and knowledge management*, pages 1437–1438, New York, NY, USA, 2008. ACM.

[14] F. de Jong, H. Rode, and D. Hiemstra. Temporal language models for the disclosure of historical text. In *Humanities, computers and cultural heritage: Proceedings of the XVIth International Conference of the Association for History and Computing (AHC 2005)*, pages 161–168, Amsterdam, The Netherlands, September 2005. Royal Netherlands Academy of Arts and Sciences.

[15] J. L. Fleiss. Measuring nominal scale agreement among many raters. In *Psychological Bulletin* 76:323–327, 1971.

[16] R. Jones and F. Diaz. Temporal profiles of queries. *ACM Trans. Inf. Syst.*, 25(3):14, 2007.

[17] P. J. Kalczynski and A. Chou. Temporal document retrieval model for business news archives. *Inf. Process. Manage.*, 41(3):635–650, 2005.

[18] N. Kanhabua and K. Nørvåg. Improving temporal language models for determining time of non-timestamped documents. In *Research and Advanced Technology for Digital Libraries, 12th European Conference, ECDL 2008*, pages 358–370, 2008.

[19] D. B. Koen and W. Bender. Time frames: Temporal augmentation of the news. *IBM Systems Journal*, 39(3&4):597–616, 2000.

[20] X. Li and W. B. Croft. Time-based language models. In *CIKM '03: Proceedings of the twelfth international conference on Information and knowledge management*, pages 469–475, New York, NY, USA, 2003. ACM.

[21] D. Metzler, R. Jones, F. Peng, and R. Zhang. Improving search relevance for implicitly temporal queries. In *SIGIR '09: Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, pages 700–701, New York, NY, USA, 2009. ACM.

[22] S. Nunes, C. Ribeiro, and G. David. Use of Temporal Expressions in Web Search. In *Advances in Information Retrieval , 30th European Conference on IR Research, ECIR 2008*, pages 580–584, 2008.

[23] J. M. Ponte and W. B. Croft. A language modeling approach to information retrieval. In *SIGIR '98: Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 275–281, New York, NY, USA, 1998. ACM.

[24] M. Verhagen, I. Mani, R. Sauri, J. Littman, R. Knippen, S. B. Jang, A. Rumshisky, J. Phillips, and J. Pustejovsky. Automating Temporal Annotation with TARSQI. In *ACL 2005, 43rd Annual Meeting of the Association for Computational Linguistics*, 2005.

[25] Q. Zhang, F. M. Suchanek, L. Yue, and G. Weikum. TOB: Timely Ontologies for Business Relations. In *11th International Workshop on the Web and Databases, WebDB 2008*, 2008.

Below you find a list of the most recent research reports of the Max-Planck-Institut für Informatik. Most of them are accessible via WWW using the URL `http://www.mpi-inf.mpg.de/reports`. Paper copies (which are not necessarily free of charge) can be ordered either by regular mail or by e-mail at the address below.

Max-Planck-Institut für Informatik
– Library and Publications –
Campus E 1 4

D-66123 Saarbrücken

E-mail: `library@mpi-inf.mpg.de`

MPI-I-2009-RG1-005	M. Horbach, C. Weidenbach	Superposition for fixed domains
MPI-I-2009-RG1-004	M. Horbach, C. Weidenbach	Decidability results for saturation-based model building
MPI-I-2009-RG1-002	P. Wischnewski, C. Weidenbach	Contextual rewriting
MPI-I-2009-RG1-001	M. Horbach, C. Weidenbach	Deciding the inductive validity of $\forall\exists^*$ queries
MPI-I-2009-5-007	G. Kasneci, G. Weikum, S. Elbassuoni	MING: Mining Informative Entity-Relationship Subgraphs
MPI-I-2009-5-006	S. Bedathur, K. Berberich, J. Dittrich, N. Mamoulis, G. Weikum	Scalable phrase mining for ad-hoc text analytics
MPI-I-2009-5-005	G. de Melo, G. Weikum	Towards a Universal Wordnet by learning from combined evidenc
MPI-I-2009-5-004	N. Preda, F.M. Suchanek, G. Kasneci, T. Neumann, G. Weikum	Coupling knowledge bases and web services for active knowledge
MPI-I-2009-5-003	T. Neumann, G. Weikum	The RDF-3X engine for scalable management of RDF data
MPI-I-2009-5-002	M. Ramanath, K.S. Kumar, G. Ifrim	Generating concise and readable summaries of XML documents
MPI-I-2009-4-006	C. Stoll	Optical reconstruction of detailed animatable human body models
MPI-I-2009-4-005	A. Berner, M. Bokeloh, M. Wand, A. Schilling, H. Seidel	Generalized intrinsic symmetry detection
MPI-I-2009-4-004	V. Havran, J. Zajac, J. Drahokoupil, H. Seidel	MPI building model as data for your research
MPI-I-2009-4-003	M. Fuchs, T. Chen, O. Wang, R. Raskar, H.P.A. Lensch, H. Seidel	A shaped temporal filter camera
MPI-I-2009-4-002	A. Tevs, M. Wand, I. Ihrke, H. Seidel	A Bayesian approach to manifold topology reconstruction
MPI-I-2009-4-001	M.B. Hullin, B. Ajdin, J. Hanika, H. Seidel, J. Kautz, H.P.A. Lensch	Acquisition and analysis of bispectral bidirectional reflectance distribution functions
MPI-I-2008-RG1-001	A. Fietzke, C. Weidenbach	Labelled splitting
MPI-I-2008-5-004	F. Suchanek, M. Sozio, G. Weikum	SOFI: a self-organizing framework for information extraction
MPI-I-2008-5-003	G. de Melo, F.M. Suchanek, A. Pease	Integrating Yago into the suggested upper merged ontology
MPI-I-2008-5-002	T. Neumann, G. Moerkotte	Single phase construction of optimal DAG-structured QEPs
MPI-I-2008-5-001	G. Kasneci, M. Ramanath, M. Sozio, F.M. Suchanek, G. Weikum	STAR: Steiner tree approximation in relationship-graphs
MPI-I-2008-4-003	T. Schultz, H. Theisel, H. Seidel	Crease surfaces: from theory to extraction and application to diffusion tensor MRI
MPI-I-2008-4-002	D. Wang, A. Belyaev, W. Saleem, H. Seidel	Estimating complexity of 3D shapes using view similarity
MPI-I-2008-1-001	D. Ajwani, I. Malinger, U. Meyer, S. Toledo	Characterizing the performance of Flash memory storage devices and its impact on algorithm design
MPI-I-2007-RG1-002	T. Hillenbrand, C. Weidenbach	Superposition for finite domains

MPI-I-2007-5-003	F.M. Suchanek, G. Kasneci, G. Weikum	Yago : a large ontology from Wikipedia and WordNet
MPI-I-2007-5-002	K. Berberich, S. Bedathur, T. Neumann, G. Weikum	A time machine for text search
MPI-I-2007-5-001	G. Kasneci, F.M. Suchanek, G. Ifrim, M. Ramanath, G. Weikum	NAGA: searching and ranking knowledge
MPI-I-2007-4-008	J. Gall, T. Brox, B. Rosenhahn, H. Seidel	Global stochastic optimization for robust and accurate human motion capture
MPI-I-2007-4-007	R. Herzog, V. Havran, K. Myszkowski, H. Seidel	Global illumination using photon ray splatting
MPI-I-2007-4-006	C. Dyken, G. Ziegler, C. Theobalt, H. Seidel	GPU marching cubes on shader model 3.0 and 4.0
MPI-I-2007-4-005	T. Schultz, J. Weickert, H. Seidel	A higher-order structure tensor
MPI-I-2007-4-004	C. Stoll, E. de Aguiar, C. Theobalt, H. Seidel	A volumetric approach to interactive shape editing
MPI-I-2007-4-003	R. Bargmann, V. Blanz, H. Seidel	A nonlinear viseme model for triphone-based speech synthesis
MPI-I-2007-4-002	T. Langer, H. Seidel	Construction of smooth maps with mean value coordinates
MPI-I-2007-4-001	J. Gall, B. Rosenhahn, H. Seidel	Clustered stochastic optimization for object recognition and pose estimation
MPI-I-2007-2-001	A. Podelski, S. Wagner	A method and a tool for automatic veriication of region stability for hybrid systems
MPI-I-2007-1-003	A. Gidenstam, M. Papatriantafilou	LFthreads: a lock-free thread library
MPI-I-2007-1-002	E. Althaus, S. Canzar	A Lagrangian relaxation approach for the multiple sequence alignment problem
MPI-I-2007-1-001	E. Berberich, L. Kettner	Linear-time reordering in a sweep-line algorithm for algebraic curves intersecting in a common point
MPI-I-2006-5-006	G. Kasnec, F.M. Suchanek, G. Weikum	Yago - a core of semantic knowledge
MPI-I-2006-5-005	R. Angelova, S. Siersdorfer	A neighborhood-based approach for clustering of linked document collections
MPI-I-2006-5-004	F. Suchanek, G. Ifrim, G. Weikum	Combining linguistic and statistical analysis to extract relations from web documents
MPI-I-2006-5-003	V. Scholz, M. Magnor	Garment texture editing in monocular video sequences based on color-coded printing patterns
MPI-I-2006-5-002	H. Bast, D. Majumdar, R. Schenkel, M. Theobald, G. Weikum	IO-Top-k: index-access optimized top-k query processing
MPI-I-2006-5-001	M. Bender, S. Michel, G. Weikum, P. Triantafilou	Overlap-aware global df estimation in distributed information retrieval systems
MPI-I-2006-4-010	A. Belyaev, T. Langer, H. Seidel	Mean value coordinates for arbitrary spherical polygons and polyhedra in $\mathbb{R}^3$
MPI-I-2006-4-009	J. Gall, J. Potthoff, B. Rosenhahn, C. Schnoerr, H. Seidel	Interacting and annealing particle filters: mathematics and a recipe for applications
MPI-I-2006-4-008	I. Albrecht, M. Kipp, M. Neff, H. Seidel	Gesture modeling and animation by imitation
MPI-I-2006-4-007	O. Schall, A. Belyaev, H. Seidel	Feature-preserving non-local denoising of static and time-varying range data
MPI-I-2006-4-006	C. Theobalt, N. Ahmed, H. Lensch, M. Magnor, H. Seidel	Enhanced dynamic reflectometry for relightable free-viewpoint video
MPI-I-2006-4-005	A. Belyaev, H. Seidel, S. Yoshizawa	Skeleton-driven laplacian mesh deformations
MPI-I-2006-4-004	V. Havran, R. Herzog, H. Seidel	On fast construction of spatial hierarchies for ray tracing
MPI-I-2006-4-003	E. de Aguiar, R. Zayer, C. Theobalt, M. Magnor, H. Seidel	A framework for natural animation of digitized models
MPI-I-2006-4-002	G. Ziegler, A. Tevs, C. Theobalt, H. Seidel	GPU point list generation through histogram pyramids