Diachronic Analysis of Time References in News Articles

Adam Jatowt
University of Innsbruck
Dept. of Computer Science & DiSC
Innsbruck, Austria
jatowt@gmail.com

Antoine Doucet
University of La Rochelle
La Rochelle, France
antoine.doucet@univ-lr.fr

Ricardo Campos
LIAAD - INESCTEC
Ci2 - Polytechnic Institute of Tomar
Tomar, Portugal
ricardo.campos@ipt.pt

ABSTRACT

Time expressions embedded in text are important for many downstream tasks in NLP and IR. They have been, for example, utilized for timeline summarization, named entity recognition, temporal information retrieval, question answering and others. In this paper, we introduce a novel analytical approach to analyzing characteristics of time expressions in diachronic text collections. Based on a collection of news articles published over a 33-years' long time span, we investigate several aspects of time expressions with a focus on their interplay with publication dates of containing documents. We utilize a graph-based representation of temporal expressions to represent them through their co-occurring named entities. The proposed approach results in several observations that could be utilized in automatic systems that rely on processing temporal signals embedded in text. It could be also of importance for professionals (e.g., historians) who wish to understand fluctuations in collective memories and collective expectations based on large-scale, diachronic document collections.

CCS CONCEPTS

• **Information systems → Content analysis and feature selection**.

KEYWORDS

temporal expressions, news archives, temporal IR

ACM Reference Format:
Adam Jatowt, Antoine Doucet, and Ricardo Campos. 2022. Diachronic Analysis of Time References in News Articles. In *Companion Proceedings of the Web Conference 2022 (WWW '22 Companion), April 25–29, 2022, Virtual Event, Lyon, France.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3487553.3524671

1 INTRODUCTION

Time expressions are important signals in unstructured textual data and are used for a variety of NLP, IR and other related tasks. In news articles, for example, they define temporal scopes of described events and facts, being at least as informative as the publication dates of documents. They also play an important role in several NLP and IR tasks. Event detection and ordering [10, 24], timeline summarization [3, 7, 17, 22, 26, 31], event occurrence prediction [29], temporal clustering and information retrieval [2, 5, 6], question answering [19, 27] and named entity recognition [1, 20] are example tasks where using time embedded in text proved beneficial. Furthermore, it was revealed that a significant number of web queries contain explicit temporal expressions [32]. It is then important to study the characteristics governing the characteristics and distribution of temporal reference mentions in texts.

In this paper, we propose an analysis approach that aims at studying the interplay between temporal expressions embedded in text and document publication time within long-term news article collections. Our aim is to provide new observations that could be useful for NLP and IR tasks that utilize time expressions and to propose new analytical approaches that could be used for supporting collective memory studies. Our work has then two objectives: (a) to uncover new observations related to time references embedded in news articles based on large scale temporal document collections and (b) to propose a novel framework useful for analysis of changes in collective memories and future expectations[1]. The latter aim fits into the recent trend of Culturonomics [18] that looks into the way in which our culture evolved over time. Professionals such as sociologists, historians or journalists often need good understanding how our society referred to different time periods and how these references changed over time (e.g., which years were strongly remembered in different periods in the past and how these memories evolved). In particular, collective memory studies that investigate society-level memories and their triggers, and which have been increasingly mediated by quantitative approaches [4, 8, 14, 25], could benefit from the proposed analysis.

The literature describes several studies on the distribution and time horizon of temporal expressions embedded in text, which were often carried in the context of collective memory analysis. In [4] the authors combined content dates with topic models to uncover topics strongly associated with the remembrance of given years in relation to diverse countries. In the context of history-related tweets, it was demonstrated that the attention to the distant past is smaller than to the recent past and the recollections of past years tend to be driven by anniversaries [25]. Similar observation of memory triggers came from the analysis of Wikipedia edit histories [14]. Rizzo and Montesi [21] demonstrated through quantitative studies the temporal variant of Zipf's law by showing that the distributions of temporal expressions tend to be governed by the well-known relation between the rank and frequency. Jatowt et al. [12] have investigated and visualized in aggregate the prevalence and scope of past- and future-pointing temporal expressions in Twitter. An aggregate analysis of date mentions was also done

[1]Note that to keep the analysis manageable we conduct it based on year granularity, hence we neglect time expressions of finer granularities such as days or months.

WWW '22 Companion, April 25–29, 2022, Virtual Event, Lyon, France

ACM ISBN 978-1-4503-9130-6/22/04…$15.00
https://doi.org/10.1145/3487553.3524671

over street names in [23]. Diachronic analysis of time references within temporal document collections was however less researched. In [9] the authors investigated diachronic changes of temporal expressions based on relative entropy in scientific writing ranging from 1665 to 2007. In this paper, we conduct frequency, semantic, and time-based analysis of content dates (i.e., year mentions) within a 20 years long collection of news articles. We employ a technique of graph-based embeddings and of temporal embeddings, and utilize named entities for representing years, as entities are the essence of news and are strongly associated with temporal signals [1, 20].

2 DATASET

The dataset is composed of New York Times articles that were published from Jan 1, **1981** to Mar 28, **2013** and which were crawled online, similarly to the procedure employed by Yao *et al.* [30]. In total, our document collection contains 282,412 news articles belonging to 4 decades. Named entities were extracted from the article content using the Stanford Natural Language toolkit for named entity extraction[2]. After lowercasing, there were 4,484,145 entity instances found, resulting in 706,163 unique entities. Entities were grouped into four types: `year`, `location`, `organization`, `person`.

All the entities, except years, were then further filtered to retain only those that appear at least 20 times. This resulted in 20,593 unique entities (there are about 17k, 19.5k, 19.3k, and 11.6k unique entities in the 1980s, 1990s, 2020s and 2010s decade, respectively) which were used for creating the co-occurrence graph further described in Sec. 5. In our analysis, we focus on the year references between **1900** and **2020**[3], based on the prior study in literature which revealed that time references of yearly granularity are most commonly used when referring to distant periods [13][4]. In the remainder of the paper we use an expression *content years* to denote years mentioned as temporal expressions in the content of news articles to distinguish them from publication years of these articles.

3 ANALYSIS

3.1 Frequency Analysis

We first look into how year mentions are distributed over time in our dataset. Fig. 1 shows the distributions of year mentions (called also content years) in each segment[5] of the dataset's interval (i.e., period from 1981 to 2013). For facilitating the visualization, we indicate with the same color the year mentions that belong to the same decade. First, we notice that the mentions of content years from the 1980s, 1990s, 2000s and 2010s are the most common in our dataset. Rizzo and Montesi [21] have already demonstrated that most of the content temporal expressions fall within the time scope of an underlying dataset. This can be indeed observed in Fig. 1 as the content years of 80s, 90s and 00s decades are the most common. What we can additionally see is that distant past decades (1900s-1930s), such as ones from around the beginning of the last century, are much less common, and they tend to appear relatively more uniformly over time than the years from the later decades (1940s-1970s), which are subject to sharper decreases.

The low frequency and low variance of frequencies of content years that point to the outside periods of our dataset (in our case these are content years of 1900s-1970s and of 2020s) is also evident in Fig. 2. Fig. 2 plots the so-called standardized variance of the frequency of content dates computed over the time period covered by our dataset. In particular, the vertical axis gives the coefficient of variation (also known as the relative standard deviation) which is defined as the ratio of standard deviation to the mean, both of which are calculated over 33 year-long segments of the dataset. The frequency of year mentions is shown on the other hand on the horizontal axis. For computing the variance, we first divided the dataset span into 33 year segments and then we measured the frequency of each unique content year mention in every segment. We could then compute content years' variances and plot Fig. 2 to see if we can corroborate the prior observations from Fig. 1.

To facilitate the visualization and comparison between Fig. 1 and Fig. 2, years of each particular decade are marked by the same distinctive color. The individual years in Fig. 2 are distinguished by labels representing their last-digits. For instance, the number 9 occurring at the very top of the figure with a dot in pink colour indicates the content date "2019" since pink is used to denote the last decade (2010s). We can observe that while this year has a rather moderate frequency, it is subject to the highest variation.

Overall, we can see that years in the early decades (i.e., distant past) occur with quite low average frequency, thus confirming the intuition that distant past matters less than the near past. They also tend to be mentioned rather uniformly over the duration of our dataset (i.e., from 1981 to 2013) as indicated by their low relative standard deviation. Another thing to note is that both the mean frequency as well as the relative standard deviation tend to increase for the content years which are closer to the left boundary of the dataset (i.e., 1981) when following the timeline from the past to future. This trend reverses, however, for future-pointing content years, i.e., years after the dataset's right boundary (i.e., 2013) that are indicated in red. The future years again occur relatively less frequently and are subject to smaller variations across the dataset span compared to the years of 1980s-2010s. Quite high variation is on the other hand characterizing the years that fall within the range of the collection span (i.e., 1981 to 2013) which is likely due to the typical focus on freshness and recency in news articles.

Round years (i.e., the first years of each decade denoted by labels "0") have usually higher mean frequency than their nearby years. The round years also occur with a relatively low variance, usually lower than other years of the same corresponding decade. This is likely because they may serve as a kind of temporal landmarks or due to the occurrence of decade-denoting references like 1980s.

4 ENTITY CO-OCCURRENCE ANALYSIS

We now turn our attention to the analysis of the co-occurrence of year mentions with other entities. The plot in Fig. 3 shows the relation between the average frequency of content years computed over all the segments of the dataset (x-axis) and the total number of entities that these content years co-occur with in the dataset (y-axis). By looking at the figure one can notice the positive correlation between the content year's frequency and its level of co-occurrence

[2] https://nlp.stanford.edu/software/CRF-NER.html
[3] Note that since NYT dataset ends in 2013, our analysis also involves future pointing years that is years from 2013 to 2020.
[4] We decided to start from 1900 since the expressions pointing to years before 1900 were relatively rare in our dataset.
[5] We use yearly granularity hence the segments have unit length of 1 year.

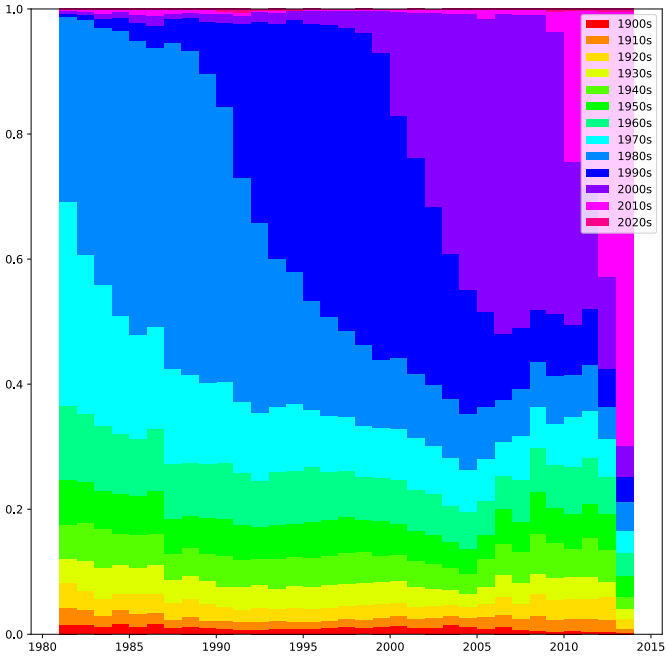

Figure 1: Distributions of decades corresponding to the content years in each year-long segment of the dataset span. Best viewed in color.

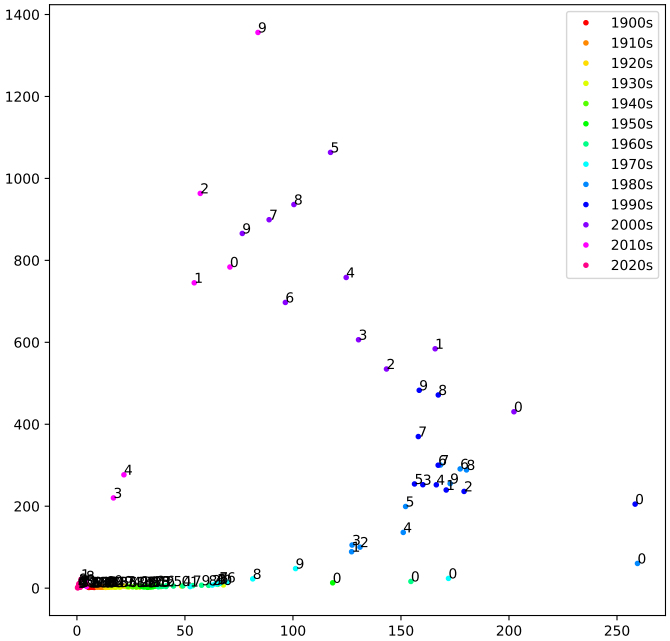

Figure 2: Average frequency of the content years (x-axis) vs. their coefficients of variation (y-axis) computed over the 33 year segments. A content year is indicated by its last digit and by the color of its corresponding decade. Best viewed in color.

with entities per year. We can also observe that the degree of the co-occurrence with entities is the highest for the years that fall within the dataset's time span, while it gets lower for the years before and after that time span. This means the texts about the years before the collection's start date and the years after the collection's end have smaller numbers of co-occurring entities than the other dates. Also, as we see, round years have on average higher co-occurrence with entities than the other years in their respective decades, which corresponds to the observation from Fig. 2.

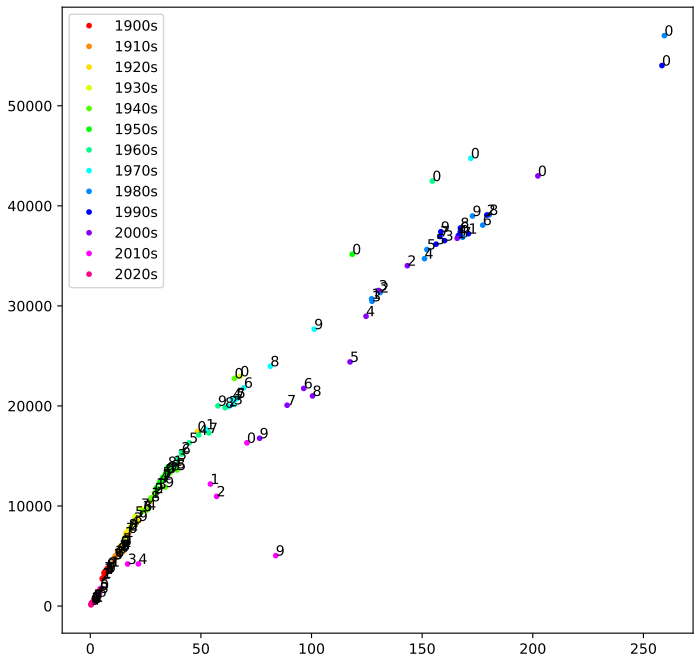

Figure 3: The total count of content years (x-axis) vs. the average number of unique entities that co-occur with them computed over yearly chunks of the dataset. A given year is represented by displaying its last digit and by the color of its corresponding decade. Best viewed in color.

5 SEMANTIC ANALYSIS OF CONTENT YEARS

We next embed years based on the entities they co-occur with. For this, we first create a graph $G = (V, E)$ such that a vertex $v_i \in V$ represents an entity (any entities including also content years which are considered as entities as well) and the weight of an edge $e_j \in E$ is determined by using Pointwise Mutual Information (PMI) measure of association computed at the document level[6]. We then use node2vec [11], a popular graph embedding approach, to compute year embedding vectors with dimensions set to 128[7]. Finally, we display in Fig. 4 the content year embeddings from our dataset on a 2D plot using t-SNE visualization [16].

As we can observe, the chronologically close years are generally located near each other, although this is rather less obvious for the distant past years. The past years are positioned on the bottom left, and these are followed by the years just before and within the dataset's time interval, which are situated around the lower middle of the figure. Future years appear to be separated from the rest of the pack and sit on the top right. The overall shape of the plot is quite interesting as the density (or spatial dispersion) of near years' seem to change along the time from less to more compact once we move from the past towards the present. The distant past (1900-1939) occupies a relatively large round area, while the more recent past (1940-1979) "becomes narrower" and the "present" (1980-2013) has already a quite elongated plot shape. One interpretation for this could be that years in the distant past tend to be quite diverse and often dissimilar from each other, and their chronological distance is less correlated with their semantic similarity. On the other hand, the semantic similarity between the more recent years is more governed by their chronological order, so that chronologically nearby years are close to each other. However, when the temporal distance increases, so does the dissimilarity, hence the correlation between the two is stronger for more recent years (hence the elongated plot).

[6]We have dropped edges whose entities co-occurred less than 15 times in the dataset.
[7]We use the following implementation: https://github.com/aditya-grover/node2vec/blob/master/src/main.py with default values (walk-length: 80, num-walks: 20, window-size: 10, etc.)

6 TEMPORAL ANALYSIS

Finally, we would like to quantify the degree of drift that the semantics of each year underwent over the time span of the dataset. We thus need an approach that allows comparing the year embeddings computed over different decades of our dataset. To this regard, we create 4 graphs (each one for a different decade of data split) such that for each graph $G^i = (V^i, E^i)$ a vertex $v_j \in V^i$ represents all the entities including also years, and the weight of an edge $e_j \in E^i$ is determined by using Pointwise Mutual Information (PMI) measure of association between nodes computed on the document level in a similar way as in Sec. 5. What is important to note is that each graph is created based on data from a particular decade that the dataset spans over (i.e., 1980's, 1990's, 2000's and 2010's). The resulting graphs represent the co-occurrence data of all the entities in each of these four decades. We then use a method that relies on retraining the embeddings sequentially from the oldest decade (i.e., 1980s) to the latest decade (i.e., 2000s) [15] such that the model's parameters obtained after training on one decade (i.e., on graph G^i) are updated by subsequently training on the data from the next decade (i.e., on graph G^{i+1}). In this way, we generate 120 vectors that represent all the analyzed content years (i.e., ones from 1900

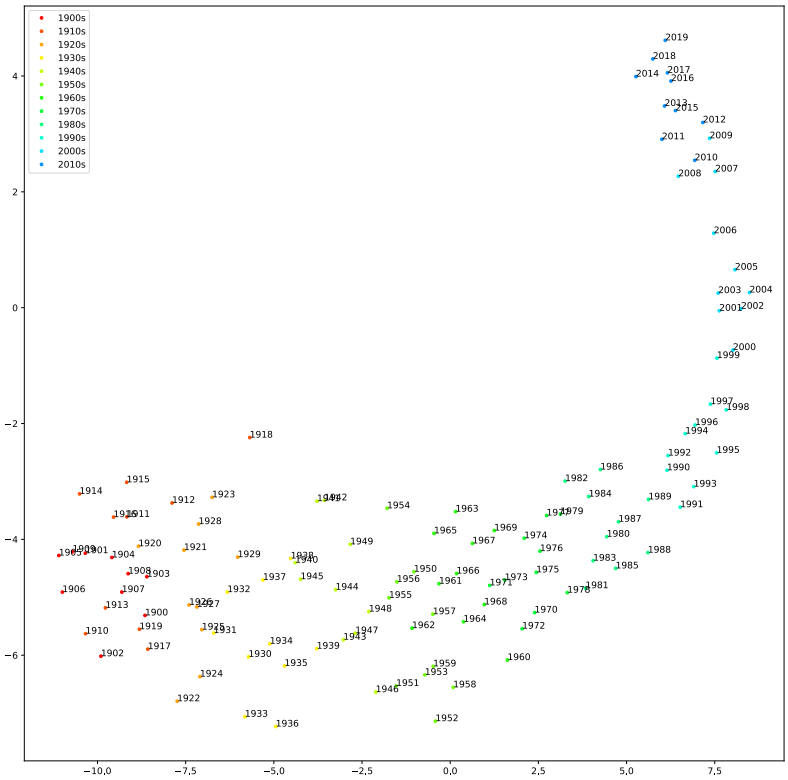

Figure 4: Similarities of year embeddings. Best viewed in color. 4

to 2020) in each decade of the dataset span ('80s, '90s, '00s and '10s). Hence in total there are 480 vectors as each of 120 content years is represented by 4 vectors where each one corresponds to one of 4 different decades covered by our dataset. We then use the Affinity Propagation (AP) clustering[8] to capture the similarities and differences between year embeddings trained over different temporal chunks of the dataset (i.e., four decades: '80s, '90s, '00s and '10s). AP is a clustering method based on a message passing mechanism that does not require setting the number of clusters. The number of clusters is decided automatically based on the input dataset. After applying AP to our 480 embedding vectors of years, we obtained 51 clusters which are shown by color coding in Fig. 5. Each cell of Fig. 5 corresponds to a particular content year (indicated as the row number from 1900 to 2020) that is embedded based on data from a given decade of the dataset's span (see the column labels: '80s, '90s, '00s and '10s). The numbers in cells and their colors correspond to particular clusters (the numbers represent cluster ids).

It is interesting to observe that only for one year (1953) all the four vector representations of that year fall into the same cluster. Most of the time, embeddings of the same year that were derived on the basis of different decades are placed in different clusters meaning that the semantics of content dates tend to differ quite much in different temporal splits of the dataset. Only 13 years (i.e., only 11% of the studied years) have 3 or more of their representations (out of 4 possible) that belong to the same cluster. Interestingly, 37 years, which constitutes over 30% of all the analyzed content years, have each of their representations belonging to a different cluster. Also, we observe in Fig. 5 that clusters tend to be spaced column-wise rather than row-wise, and that they cover relatively coherent regions considering the rather long time span of content years (120 years). Overall, these findings indicate that the same year tends to be represented differently based on different temporal portions of dataset used, even if the year is in the distant past or far future w.r.t. the interval covered by this dataset.

[8]We utilize Scikit-learn implementation with the default parameters.

7 CONCLUSIONS

We presented in this paper a framework for analyzing temporal signals in diachronic text collections in a novel way focusing on the interplay between the content and publication dates. The analysis can be adapted to support investigating collective memories and collective expectations from large temporal document datasets. Finally, the observations we discuss help in better understanding the characteristics of time references embedded in text. We summarize them below:

- Time expressions that refer to the time interval of a temporal news collection are most common, have *higher average frequency* and are *subject to higher fluctuations* than temporal expressions which refer to the outside of that interval. They also tend to *co-occur with a larger number of entities.*
- Distant past years are *less similar to each other* than the more recent years are.

year	80s	90s	00s	10s
1900	4	4	19	19
1901	0	0	27	27
1902	0	0	12	12
1903	0	0	0	36
1904	4	4	21	21
1905	0	0	13	35
1906	0	49	49	49
1907	0	0	19	45
1908	0	0	13	13
1909	0	0	0	34
1910	4	4	0	17
1911	4	4	4	44
1912	4	0	18	48
1913	0	0	16	32
1914	0	4	15	40
1915	4	0	0	41
1916	0	0	16	39
1917	0	0	18	43
1918	4	0	18	38
1919	4	5	18	26
1920	4	0	16	5
1921	0	3	15	16
1922	0	5	16	15
1923	3	3	16	20
1924	5	5	18	15
1925	5	17	17	17
1926	3	5	16	2
1927	0	5	16	50
1928	3	3	16	29
1929	3	5	16	42
1930	3	3	16	15
1931	3	5	5	8
1932	5	3	18	46
1933	3	5	18	33
1934	3	16	5	47
1935	3	5	18	31
1936	5	5	18	18
1937	3	5	18	16
1938	3	18	18	18
1939	3	18	18	28
1940	5	2	16	18
1941	5	2	18	15
1942	5	2	18	30
1943	5	18	16	16
1944	5	5	16	25
1945	5	2	2	16
1946	2	8	16	18
1947	2	8	18	16
1948	5	18	16	16
1949	2	8	8	16
1950	8	2	16	16
1951	2	2	8	15
1952	2	2	2	18
1953	2	2	2	2
1954	2	2	8	37
1955	2	8	16	18
1956	2	2	2	15
1957	2	8	8	15
1958	2	8	18	15
1959	2	8	14	14
1960	1	8	15	11
1961	8	2	18	14
1962	2	8	15	8
1963	1	8	8	18
1964	2	1	8	15
1965	8	8	14	16
1966	2	8	8	15
1967	1	8	8	14
1968	8	1	15	15
1969	1	1	8	14
1970	1	7	16	15
1971	1	1	14	14
1972	1	1	14	11
1973	1	1	14	14
1974	1	8	15	11
1975	1	1	14	8
1976	1	14	8	16
1977	1	1	11	11
1978	1	7	7	16
1979	7	7	11	11
1980	7	7	11	11
1981	7	1	14	11
1982	7	7	7	11
1983	7	7	14	16
1984	7	7	14	11
1985	7	23	7	14
1986	7	7	7	11
1987	7	7	11	14
1988	7	7	7	15
1989	10	7	7	14
1990	10	7	11	16
1991	10	7	23	23
1992	10	20	23	14
1993	6	23	11	14
1994	6	23	23	23
1995	10	23	7	7
1996	6	10	20	11
1997	10	10	23	11
1998	6	10	23	23
1999	9	10	23	23
2000	10	10	23	23
2001	10	10	20	20
2002	6	6	23	23
2003	6	24	20	23
2004	9	24	20	14
2005	10	9	20	16
2006	9	9	20	20
2007	9	9	20	20
2008	6	24	24	20
2009	6	9	24	15
2010	10	24	24	20
2011	6	24	24	16
2012	9	9	24	16
2013	6	6	6	20
2014	9	9	9	24
2015	9	10	24	24
2016	6	9	9	24
2017	6	9	24	51
2018	6	9	24	51
2019	9	9	9	24
2020	6	22	22	22

Figure 5: Clusters returned by AP clustering indicated by cluster IDs listed in each cell (from 1 to 51) and distinctive colors.

- Chronological order *plays stronger role in inter-year similarities* for years within the dataset interval or close to it, than for other more distant years.
- The *same years have different semantics in different temporal splices of the dataset* that used for computing their embeddings, even if these years belong to the distant past or the far future.

The above findings could be incorporated in various applications that utilize temporal expressions. For example, they could be applied for normalizing the frequencies of temporal signals found in texts or for estimating their relative importance. It is expected that a year in a distant past should be on average less frequent than any more recent year, hence if the frequency of the two years is similar, we should assign the higher importance to the more distant year. In another example, event-to-event linking in timeline generation [22] could incorporate the expected similarity change between descriptions of events from different years depending on the distance between these years. Further, QA systems that use temporal expressions to locate correct answers (e.g., [27, 28]) over long-span archival news collections could make use of expected distributions of these temporal expressions over different time segments.

ACKNOWLEDGMENTS

Ricardo Campos was financed by the ERDF – European Regional Development Fund through the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 and by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia within project PTDC/CCI-COM/31857/2017 (NORTE-01-0145-FEDER-03185). This funding fits under the research line of the Text2Story project.

REFERENCES

[1] Prabal Agarwal, Jannik Strötgen, Luciano Del Corro, Johannes Hoffart, and Gerhard Weikum. 2018. Dianed: time-aware named entity disambiguation for diachronic corpora. In *ACL*. 686–693.

[2] Omar Alonso, Michael Gertz, and Ricardo Baeza-Yates. 2007. On the value of temporal information in information retrieval. *SIGIR Forum* 41, 2 (2007), 35–41.

[3] Omar Alonso, Michael Gertz, and Ricardo Baeza-Yates. 2009. Clustering and exploring search results using timeline constructions. In *CIKM*. 97–106.

[4] Ching-man Au Yeung and Adam Jatowt. 2011. Studying how the past is remembered: towards computational history through large scale text mining. In *CIKM*. 1231–1240.

[5] Ricardo Campos, Gaël Dias, Alípio M Jorge, and Adam Jatowt. 2014. Survey of temporal information retrieval and related applications. *ACM Computing Surveys (CSUR)* 47, 2 (2014), 1–41.

[6] R. Campos, A. M. Jorge, G. Dias, and C. Nunes. 2012. Disambiguating Implicit Temporal Queries by Clustering Top Relevant Dates in Web Snippets. In *WI*. 1–8.

[7] R. Campos, A. Pasquali, A. Jatowt, V. Mangaravite, and A. Jorge. 2021. Automatic Generation of Timelines for Past-Web Events. In *The Past Web. Exploring Web Archives*, D. Gomes, E. Demidova, J. Winters, and T. Risse (Eds.). Springer, Chapter 18, 225–242.

[8] James Cook, Atish Das Sarma, Alex Fabrikant, and Andrew Tomkins. 2012. Your two weeks of fame and your grandmother's. In *WWW*. 919–928.

[9] Stefania Degaetano-Ortlieb and Jannik Strötgen. 2017. Diachronic variation of temporal expressions in scientific writing through the lens of relative entropy. In *International Conference of the German Society for Computational Linguistics and Language Technology*. Springer, 259–275.

[10] Leon Derczynski. 2017. *Automatically Ordering Events and Times in Text*. Vol. 677. https://doi.org/10.1007/978-3-319-47241-6

[11] Aditya Grover and Jure Leskovec. 2016. Node2Vec: Scalable Feature Learning for Networks. 855–864.

[12] Adam Jatowt, Émilien Antoine, Yukiko Kawai, and Toyokazu Akiyama. 2015. Mapping Temporal Horizons: Analysis of Collective Future and Past related Attention in Twitter. In *Proceedings of the 24th Int. Conference on World Wide Web, WWW 2015, May 18-22, 2015*. ACM, 484–494.

[13] Adam Jatowt and Ching-man Au Yeung. 2011. Extracting collective expectations about the future from large text collections. In *CIKM*. 1259–1264.

[14] Nattiya Kanhabua, Tu Ngoc Nguyen, and Claudia Niederée. 2014. What triggers human remembering of events? A large-scale analysis of catalysts for collective memory in Wikipedia. In *JCDL*. IEEE, 341–350.

[15] Yoon Kim, Yi-I Chiu, Kentaro Hanaki, Darshan Hegde, and Slav Petrov. 2014. Temporal analysis of language through neural language models. *ACL* (2014), 61.

[16] Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of machine learning research* 9, Nov (2008), 2579–2605.

[17] S. Martschat and M. Katja. 2018. A temporally sensitive submodularity framework for timeline summarization. In *CoNLL*. 230–240.

[18] Jean-Baptiste Michel, Yuan Kui Shen, et al. 2011. Quantitative analysis of culture using millions of digitized books. *science* 331(6014) (2011), 176–182.

[19] Marius Pasca. 2008. Towards Temporal Web Search. In *SAC*. 1117–1121.

[20] Shruti Rijhwani and Daniel Preotiuc-Pietro. 2020. Temporally-informed analysis of named entity recognition. In *ACL*. 1–13.

[21] Stefano Giovanni Rizzo and Danilo Montesi. 2017. Quantification of time in digital libraries: temporal Zipf's law. In *IDEAS*. ACM, 143–152.

[22] Julius Steen and Katja Markert. 2019. Abstractive Timeline Summarization. In *the 2nd Workshop on New Frontiers in Summarization*. 21–31.

[23] Jannik Strötgen, Rosita Andrade, and Dhruv Gupta. 2018. Putting Dates on the Map: Harvesting and Analyzing Street Names with Date Mentions and Their Explanations. In *Proceedings of JCDL'18*. ACM, New York, NY, USA, 79–88.

[24] Jannik Strötgen and Michael Gertz. 2012. Event-centric search and exploration in document collections. In *JCDL*. 223–232.

[25] Yasunobu Sumikawa, Adam Jatowt, and Marten During. 2018. Digital History Meets Microblogging: Analyzing Collective Memories in Twitter. In *JCDL* (Fort Worth, Texas, USA). ACM, 213–222.

[26] Giang Tran, Mohammad Alrifai, and Eelco Herder. 2015. Timeline summarization from relevant headlines. In *ECIR*. Springer, 245–256.

[27] Jiexin Wang, Adam Jatowt, Michael Färber, and Masatoshi Yoshikawa. 2020. Answering Event-Related Questions over Long-Term News Article Archives, Vol. 12035. Springer, 774–789.

[28] Jiexin Wang, Adam Jatowt, and Masatoshi Yoshikawa. 2021. ArchivalQA: A Large-scale Benchmark Dataset for Open Domain Question Answering over Historical News Collections. *SIGIR* (2021).

[29] Jiexin Wang, Adam Jatowt, and Masatoshi Yoshikawa. 2021. Event Occurrence Date Estimation based on Multivariate Time Series Analysis over Temporal Document Collections. In *SIGIR '21*. ACM, 398–407.

[30] Zijun Yao, Yifan Sun, Weicong Ding, Nikhil Rao, and Hui Xiong. 2018. Dynamic word embeddings for evolving semantic discovery. In *WSDM*. 673–681.

[31] Yi Yu, Adam Jatowt, Antoine Doucet, Kazunari Sugiyama, and Masatoshi Yoshikawa. 2021. Multi-TimeLine Summarization (MTLS): Improving Timeline Summarization by Generating Multiple Summaries. In *Proceedings of the 59th ACL/IJCNLP 2021*. ACL, 377–387.

[32] Ruiqiang Zhang, Yuki Konda, Anlei Dong, Pranam Kolari, Yi Chang, and Zhaohui Zheng. 2010. Learning recurrent event queries for web search. In *Proceedings of EMNLP*. 1129–1139.