Temporal and Geographic Oriented Event Retrieval for Historical Analogy

Kengo Fushimi[1] and Yasunobu Sumikawa[1]

Depat. of Computer Science, Takushoku University, 815-1 Tatemachi, Hachioji-shi, Tokyo, Japan
r88464@st.takushoku-u.ac.jp, ysumikaw@cs.takushoku-u.ac.jp

Abstract. There are many benefits to studying history. Recently, the study of support for learning history has emphasized the development of the ability to use knowledge of the past in an analogous manner when considering solutions to problems that arise in the present day, a process called historical analogy. Although previous studies have developed the ability of historical analogy using card games or datasets prepared in advance by experts, the situation is not ready for anyone to learn about a subject that is in line with their interests. In this study, we propose a Twitter chatbot that presents past events recorded on Wikipedia. When this chatbot receives a reply from a Twitter user, it analyzes the user's past tweets to collect geographic and temporal information of interest and returns past events that are close to those. We conducted experiments to confirm the effectiveness of our algorithm, and confirmed that the accuracy of our algorithm was approximately 30% points higher than that of the methods used in previous studies.

Keywords: History · Twitter · Wikipedia · Analogy

1 Introduction

The importance of learning history has been widely recognized. In fact, studies have been conducted on how to enhance historical analogy, which is the ability to use knowledge regarding historical events as an analogy for thinking about solutions to contemporary problems [15], to propose a search engine designed for history learning [11], and to implement a chatbot to support communication between people and history [19]. These systems that mediate between people and history are useful when searching for history that users want to know. However, they do not allow users to search for histories that can be used for historical analogies or for other similar histories that can be used as alternatives when the history required by the users is not available.

In this study, we propose a Twitter bot that outputs a historical event to enhance the historical analogy. The key idea of this bot is that the outputs are generated with calculations to satisfy the condition that the analogy works well. According to [10], there are two important conditions for effectively utilizing analogies. The first is the incompleteness. This indicates that, when generating a plausible inference from a source to a target, they cannot be exactly the same and cannot be completely different. The second is the explicit thinking ability pertaining to higher-order relations that exist between the source and the target.

To satisfy the above two conditions, the bot uses temporal and location information to create common aspects and higher-order relations between user input texts and historical events. The bot first takes the user text as a reply to the bot on Twitter. It then retrieves similar historical events that are close in time and location to the input text of the user. This bot infers the location from the input text and the user's past tweets if there is no explicit information in the text. It also obtains time information from the input timestamp. The bot calculates the ranking score for each historical event using these three pieces of information to measure the similarity between the input text and historical events. It then outputs the history with the highest score.

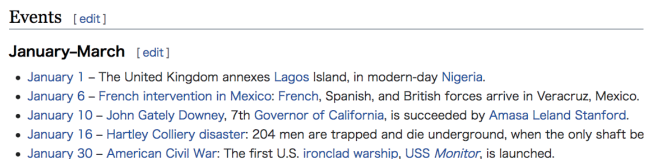
The remainder of this paper is organized as follows. Section 2 describes the related work. Section 3 details how this study collected data to output historical events using our bot and estimated the location. In section 4, we describe the proposed algorithm. The results of the experimental evaluation are presented in section 5 presents. The final section concludes the study and describes future work.

2 Related Works

Activities to create opportunities for the public to easily learn about history had already been carried out by Italian historian Guiccardini [14] in the 15th century. While historians' activities to encourage dialogue between non-researchers and history have been performed for a long time, research to determine the expertise of history that would influence the relationship between historians and the public began at the end of the 19th century [5]. In the early twentieth century, Rebecca Conard identified the use of history in activities other than teaching and proposed the value of making history relevant to the present [6]. Furthermore, history workshops have been held to help workshop organizers promote new practices in local and oral history [17].

The purpose of this research was to encourage dialogue between history and Twitter users through chatbots. A chatbot is a program that can communicate with people using natural language processing techniques. It is defined by [4] as "an artificial construct designed to converse with human beings using natural language as input and output." Classic chatbots were designed to entertain people; however, in recent years, attempts have been made to improve the service quality [1, 7, 9]. Chatbots are also beginning to be used in education. There have been studies using chatbots for foreign language learning [12, 13]; however, in recent years, chatbots have been used not only for the study of other subjects (e.g., computer science [3]), but also for educational activities in general [2, 16].

A previous study used Twitter chatbot to encourage interaction with history [19]. This chatbot was designed to disseminate past events recorded on Wikipedia. The purpose of this research is to make people more familiar with


Events [edit]

January–March [edit]

- January 1 – The United Kingdom annexes Lagos Island, in modern-day Nigeria.
- January 6 – French intervention in Mexico: French, Spanish, and British forces arrive in Veracruz, Mexico.
- January 10 – John Gately Downey, 7th Governor of California, is succeeded by Amasa Leland Stanford.
- January 16 – Hartley Colliery disaster: 204 men are trapped and die underground, when the only shaft be
- January 30 – American Civil War: The first U.S. ironclad warship, USS *Monitor*, is launched.

Fig. 1. Examples of descriptions of events in Wikipedia.

the history of Twitter by automatically tweeting about past events that occurred in the same month and day or by spreading past events tweeted by other Twitter users by retweeting them. When another user replied to the chatbot in this previous study, it presented past events that included words in the sentence. At this point, if there are no appropriate past events in the DB, the chatbot simply replies "No relevant history." While the above chatbot was aimed at disseminating history, the chatbot in this study presented history with the aim of facilitating historical analogies. This difference in purpose makes it possible to present other histories that may be of interest to the user, even if there is no appropriate history in response to the user's reply.

3 Data Collection

In this section, we describe the data collected to implement our chatbot in detail.

3.1 History Event Data

Twitter allows users to post sentences described in 280 characters at a time; thus, we store the events described in short sentences in the DB. In addition, as Twitter is used worldwide, it is desirable to store the history of many countries and regions in a DB.

To meet the above criteria, in this study, we used Wikipedia. Wikipedia records past events in the year pages[1] from year 1 to the current year and the day pages[2] from January 1 to December 31. Fig. 1 shows an example of Wikipedia's events described in short texts. These events are listed in the "Events" section. We collected these data by using BeautifulSoup[3]. We collected past events and stored not only their description but also the date and related location information in our DB. The collected data covered events from all years spanning from 1 AD to 2019 AD. In total, the dataset contained the descriptions of 71,374 events. In Wikipedia, when an article describes each word in the text in detail, the word has a link to that article. To determine the location of each event, we retrieved the infobox of the linked article and name of the country or region in the Wikipedia category.

[1] E.g., `https://en.wikipedia.org/wiki/1945`

[2] E.g., `https://en.wikipedia.org/wiki/March_21`

[3] `https://www.crummy.com/software/BeautifulSoup/`

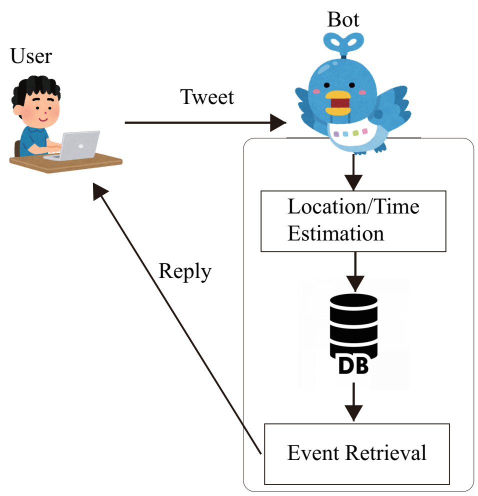


Fig. 2. System overview

3.2 Location Data

We used the latitude and longitude to determine geographic proximity. The collected texts of past events often use the names of the countries. To compare the geographic proximity between countries, we used the latitude and longitude of their capital cities. We collected the latitude and longitude of the capitals of each country from the Web[4].

4 Proposed Algorithm

In this section, we describe our algorithm for presenting history in order to enhance historical analogies. Fig. 2 presents an overview of the proposed algorithm. Our chatbot first takes reply texts from users. To collect the reply texts, we used the standard search API provided by Twitter as an official API[5]. It then extracts temporal and geographic words from the text to estimate the geographic and temporal interests of the user from the user's tweets. If these two pieces of information exist in the text, the bot uses them to retrieve historical events; otherwise, it infers information that is not in the text from the metadata and contextual information of the tweet. After retrieving the history close to this estimated geographic and time of interest from the DB, our chatbot replies to the user with the event that it thinks is the most appropriate.

In the remainder of this section, we describe algorithms for estimating and retrieving the geographic and temporal aspects of the events of interest.

---

[4] https://amano-tec.com/data/world.html

[5] https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/api-reference/get-search-tweets

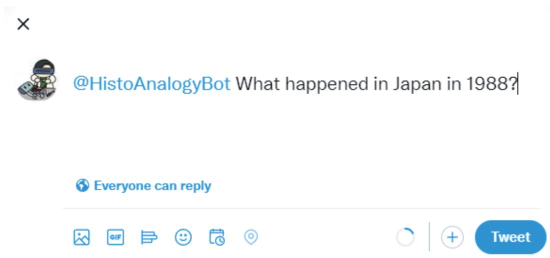


Fig. 3. Example of estimation of geographic interests

Fig. 4. Estimating geographic interests from past tweets

4.1 Estimation of Geographic and Temporal Interests

To present an appropriate past event in response to a reply from a user, the chatbot estimates the geographic and temporal interests of the user.

Geographic Interest Estimation We perform the geographic estimation in the following two steps. 1) The analysis of the reply text if it clearly includes the geographic words the user wants to know about. 2) The analysis of the user's timeline to infer the location of a particular interest from past tweets if it is not specified.

Figs. 3 and 4 show how our chatbot performs the above two analyses. Fig. 3 shows an example in which the Twitter user's replies explicitly contain the name of the country (Japan). In contrast, as shown at the top of Fig. 4, if there is no country name in the reply text, the chatbot collects past tweets of that user and analyzes the countries the user is interested in. In this example, as past tweets contain the words "Japan" and "Asia," this chatbot assumes that this user is interested in Japan.

To reveal the specified geographic information from the reply texts, we perform TAGME [8], a tool for named-entity detection, to extract country names. TAGME is a word sense disambiguation tools that links the extracted entities to Wikipedia articles. We extract the geographic information described in the infoboxes and Wikipedia categories of Wikipedia articles that are the result of TAGME. If the names of countries and capitals collected in Sec. 3.2 exist, they will be used as estimated results.

If there is no information regarding the location in the input text, we infer the location by applying TAGME to tweet texts posted by the user. If there are several location names in the TAGME results, we use the most frequent names.

Temporal Interest Estimation Temporal interest estimation analyzes which year the Twitter user has been interested in from the reply text as well as the geography. If there is no explicit year information in the reply text, we use the year with the smallest difference between the date of the event and that of the reply. This is because many Twitter users tend to be more interested in events close to the present day [18].

4.2 Event Retrieval

By applying the above estimation algorithms, we obtained geographic and temporal information in which Twitter users who replied may be interested. We use these pieces of information to retrieve the history from our DB. To determine the top 1 from the retrieval results, we also use importance score among the results. In the remainder of this section, we describe how to determine the importance of a history and the algorithm to determine the top 1 output by our chatbot.

Event Importance We determine the importance of an event by the number of references for it on Wikipedia as well as the chatbot spreading history [19]. We collected historical text from Wikipedia as described in Section 3.1. Wikipedia text may contain links to Wikipedia articles for the words in that text. We regard the sum of the number of references to these links as the importance of history. The formal definition is as follows:

$$Impr(evt) = \frac{\sum_{ett \in Entity(evt)} Link(ett)}{Max(\{\sum_{ett \in Entity(evt')} Link(ett)) \mid evt' \in E\})}$$

where evt denotes an event, ett denotes an entity, $Entity(evt)$ denotes the set of entities used in the description of argument evt, $Link(ett)$ denotes the function to calculate the number of references of ett given by the argument, and Max is a function that returns the maximum value from its argument.

Score Integration We integrate all scores obtained by applying the above algorithms to each history to output a history by our chatbot. To perform this integration, we calculate the ranking score for each history by using the following formula:

$$score(evt, t, Loc) = \alpha Temp(evt, t) + \beta Spatial(geo(evt), Loc) + \gamma imprt(evt) \quad (1)$$

where t is the result of temporal interest estimation, Loc is the result of geographic interest estimation, $Temp$ is a function that calculates the difference between the time when this algorithm is applied and the year when the event evt occurred, $Spatial$ is a function that calculates the distance between the regions where the event evt occurred and the user's interest, geo provides the geographic information of the given event, and $imprt$ is a function that calculates the importance of the event. $\alpha, \beta, and \gamma$ are the hyper-parameters. We set their values such that their sum is 1.

Algorithm1 shows the replying algorithm for the chatbot. Lines 1∼8 and 9∼10 estimate geographic and temporal information, respectively. The first step

Algorithm 1 Replying algorithm

Input: A replying tweet *reply*
Output: A past event

```
1: Function Replying(reply)
2:   // Geographic and temporal information collection
3:   Loc = ExtractGeoInfo(reply.text)
4:   //If the reply does not contain any country names
5:   if(Loc = ∅)
6:     // Geographic information estimation from past tweets
7:     Loc = GeoInference(reply.userID)
8:   end if
9:   // Temporal interest estimation
10:  t = TempInference(reply)
11:  // Loading history data
12:  HD = HistoryData(Loc)
13:  vals = []
14:  for e ∈ HD
15:    val = score(e)// Applying Eq. 1
16:    vals.append(val)
17:  idx=argmax(vals)
18:  return HD[idx]
```

in estimating geographic information is to extract only named entities from the reply text, and determine if they are names of countries or regions (3rd line). If we miss obtaining these names (5th line), we use Twitter's official API to retrieve the past tweets of the Twitter user who replied. Next, we estimate the temporal interests of the user from the replies using the same analysis as for the estimation of geographic information (lines 9∼10). After completing the above estimation analyses, we load the relevant historical data from the DB using geographic information. We then apply the Eq. 1 for each loaded history; we calculate its ranking score using temporal, geographic, and importance scores. Finally, we return the history with the highest ranking score so that our chatbot replies with a single history to the user.

5 Experimental Evaluations

5.1 Experimental setting

Dataset. Because our algorithm uses geographic information, the dataset must include tweets from several regions of the world for evaluation. In addition, because we will also use chronological information, we require a dataset that includes tweets posted at various times for the evaluation. However, there was no ground-truth dataset that met these conditions; thus, the authors created one. To collect tweet data, the tweets to be collected were made in English so that tweets from many countries could be collected. In addition, to collect tweets from a wide variety of periods, we limited the collection of tweets to official accounts of the following news organizations.

1. `TheHinduScience`: This is an Indian daily newspaper published in English by the Hindu Group. The headquarters are located in Chennai, Tamil Nadu, India. [6]
2. `BBCSport`: BBC Sport is the sports programming division of the British Broadcasting Corporation. The main office is located at Media City, UK, in Salford.[7]
3. `CBCPolitics`: CBC is the public broadcaster of Canada. It operates television and radio services collectively, forming a nationwide public broadcasting network.[8]
4. `CNNBusiness`: CNN is a U.S. cable and satellite news channel owned by CNN Worldwide, a division of WarnerMedia News Channel & Sports. CNNBusiness is a financial news and information website operated by CNN.[9]

We retrieved 25 tweets from each account, and collected 100 tweets. We checked whether the events output by this bot were the latest news articles in regions reported by each account or if they were the latest events that occurred close to the location of the account's company. Two people manually checked all the outputs of the bot to ensure that they were correct. One of the workers is a Ph.D. researcher specializing in machine learning.

Baselines. We compared our algorithm with the following three algorithms.

- HistoChatbot: This calculates the similarity of text and returns one history [19]
- Temp: This uses only temporal interest information proposed in this study
- Imprt: This uses only the event importance proposed in this study

Evaluation Criteria. We assessed whether the sentences resulting from the application of the baselines and the proposed algorithm were past events that were close to the region of each account. In addition, because our algorithm is triggered by replies from Twitter users, we confirmed that the past events output by our algorithm occurred even before each tweet in the dataset was posted.

5.2 Results

Tab.1 shows all the results of the evaluation. At the beginning, the sixth column (All) that represents the accuracy for all results shows that the proposed method performed the best.

[6] `https://mobile.twitter.com/TheHinduScience`
[7] `https://mobile.twitter.com/BBCSport`
[8] `https://mobile.twitter.com/CBCPolitics`
[9] `https://mobile.twitter.com/CNNBusiness`

Table 1. Accuracies

	`TheHinduScience`	`BBCSport`	`CBCPolitics`	`CNNBusiness`	All
HistoChatbot	56%	52%	50%	60%	54%
Temp	0%	0%	20%	24%	11%
Imprt	16%	64%	72%	56%	52%
Proposed	**80**%	**92**%	**88**%	**76**%	**84**%

Next, we check the results for each Twitter account used in the evaluation. The proposed algorithm achieved the best accuracy for all four accounts. In particular, for the results of `TheHinduScience` and `BBCSport`, the accuracy was about 25∼30% higher than that of the baselines.

Because our algorithm does not use text similarity, next, we analyzed the results of HistoChatbot, which uses text similarity. For this analysis, we used a tweet about the Zika virus tweeted by `TheHinduScience`[10]. When the text of the tweet was replied to HistoChatbot, we got "1/28/2016. January 28 - The World Health Organization announces an outbreak of the Zika virus ." as the output result. In contrast, when the same text was input into our algorithm, we got "November 27 - Indian Prime Minister Jawaharlal Nehru appeals to the United States and the Soviet Union to end nuclear testing and to start nuclear disarmament, stating that such an action would "save humanity from the ultimate disaster." as a result. The output result of HistoChatbot was the same as the input text; the output event was related to the Zika virus. However, the output results of our algorithm returned an event in which the Prime Minister of India claimed to be working to protect the health of humanity. Thus, we can see that HistoChatbot gave better output results when we wanted to know the past events related to a specific topic. However, our algorithm gave better results when we wanted to know the events in a specific region.

6 Conclusions

In this study, we have proposed an algorithm for retrieving past events that are close to the user's geography and time of interest. This algorithm is applied to users who reply to the Twitter chatbot created in this study. This algorithm first analyzes the location and time the user is interested in from the text of replies and the user's past tweets. Using the geographic and temporal analysis results, the algorithm calculates a ranking score for each past event collected from Wikipedia according to the importance of the event in addition to its geographical and temporal proximity. Finally, the algorithm outputs the event with the highest ranking score.

Future Work. Future work will *identify how our bot is effective for inquiry-based history-learning*. By using this bot, it is possible to determine history that

[10] `https://mobile.twitter.com/TheHinduScience/status/1459566755718877190`

the user did not expect. By analyzing how this affects inquiry-based learning of history, it becomes possible to examine the effectiveness of bots as a new history learning environment.

Acknowledgments. This work was supported in part by MEXT Grant-in-Aids (#19K20631).

References

1. Agus Santoso, H., Anisa Sri Winarsih, N., Mulyanto, E., Wilujeng saraswati, G., Enggar Sukmana, S., Rustad, S., Syaifur Rohman, M., Nugraha, A., Firdausillah, F.: Dinus intelligent assistance (dina) chatbot for university admission services. In: 2018 International Seminar on Application for Technology of Information and Communication. pp. 417–423 (2018)
2. Auccahuasi, W., Santiago, G.B., Núñez, E.O., Sernaque, F.: Interactive online tool as an instrument for learning mathematics through programming techniques, aimed at high school students. pp. 70–76. ICIT '2018, Association for Computing Machinery, New York, NY, USA (2018)
3. Benotti, L., Martnez, M.C., Schapachnik, F.: A tool for introducing computer science with automatic formative assessment. IEEE Transactions on Learning Technologies **11**(2), 179–192 (2018)
4. Brennan, K.: The managed teacher: Emotional labour, education, and technology. Educational Insights **10**(2), 55–65 (2006)
5. Cauvin, T.: The rise of public history: An international perspective. Historia Crítica No.40 **68**, 3–26 (04 2018)
6. Conard, R.: The pragmatic roots of public history education in the united states. The Public Historian **37**(1), 105–120 (2015)
7. Dibitonto, M., Leszczynska, K., Tazzi, F., Medaglia, C.M.: Chatbot in a campus environment: Design of lisa, a virtual assistant to help students in their university life. In: Kurosu, M. (ed.) Human-Computer Interaction. Interaction Technologies. pp. 103–116. Springer International Publishing, Cham (2018)
8. Ferragina, P., Scaiella, U.: Tagme: On-the-fly annotation of short text fragments (by wikipedia entities). In: Proceedings of the 19th ACM International Conference on Information and Knowledge Management. pp. 1625–1628. CIKM '10, Association for Computing Machinery, New York, NY, USA (2010)
9. Griol, D., Molina, J.M., Callejas, Z.: Incorporating android conversational agents in m-learning apps. Expert Systems **34**(4), e12156 (2017)
10. Holyoak, K.J., Thagard, P.: Mental Leaps: Analogy in Creative Thought. MIT Press (1980)
11. Ikejiri, R., Sumikawa, Y.: Developing world history lessons to foster authentic social participation by searching for historical causation in relation to current issues dominating the news. Journal of Educational Research on Social Studies **84**, 37–48 (2016), (in Japanese)
12. Jia, J.: Csiec: A computer assisted english learning chatbot based on textual knowledge and reasoning. Knowledge-Based Systems **22**(4), 249–255 (2009)
13. Katchapakirin, K., Anutariya, C.: An architectural design of scratchthai: A conversational agent for computational thinking development using scratch. IAIT '2018, Association for Computing Machinery, New York, NY, USA (2018)
14. Knevel, P.: Public history: the european reception of an american idea? Levend Erfgoed **6**, 4–8 (2009)

15. Lee, P.: Historical literacy: Theory and research. International Journal of Historical Learning, Teaching and Research **5**(1), 25–40 (2005)
16. Pérez, J.Q., Daradoumis, T., Puig, J.M.M.: Rediscovering the use of chatbots in education: A systematic literature review. Computer Applications in Engineering Education (2020)
17. Samuel, R.: 'editorial introduction', in samuel, ed., history workshop: A collectanea 1967–1991 (oxford: Hw 25, 1991) , iii; colin ward, 'fringe benefits' , new society (November 1989)
18. Sumikawa, Y., Jatowt, A.: Analyzing history-related posts in twitter. Int. J. Digit. Libr. **22**(1), 105–134 (2021)
19. Sumikawa, Y., Jatowt, A.: Designing chatbot systems for disseminating history-focused content in online social networks. IEICE TRANSACTIONS on Information and Systems (Japanese Edition) **J104-D**(5), 486–497 (2021), (in Japanese)