Title	Extraction and Geographical Navigation of Important Historical Events in the Web
Author(s)	Yamamoto, Mitsuo; Takahashi, Yuku; Iwasaki, Hirotoshi; Oyama, Satoshi; Ohshima, Hiroaki; Tanaka, Katsumi
Citation	Lecture Notes in Computer Science, 6574, 21-35 https://doi.org/10.1007/978-3-642-19173-2_4 Web and Wireless Geographical Information Systems, Part of the Lecture Notes in Computer Science book series (LNCS, volume 6574), ISBN: 978-3-642-19172-5
Issue Date	2011
Doc URL	http://hdl.handle.net/2115/65543
Rights	The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-642-19173-2_4
Type	article (author version)
File Information	w2gis2011.pdf

Extraction and Geographical Navigation of Important Historical Events in the Web

Mitsuo Yamamoto[1], Yuku Takahashi[2], Hirotoshi Iwasaki[1], Satoshi Oyama[3], Hiroaki Ohshima[2], and Katsumi Tanaka[2]

miyamamoto@d-itlab.co.jp, ytakahas@dl.kuis.kyoto-u.ac.jp, hiwasaki@d-itlab.co.jp, oyama@ist.hokudai.ac.jp, ohshima@dl.kuis.kyoto-u.ac.jp, tanaka@dl.kuis.kyoto-u.ac.jp

[1] Denso IT Laboratory, Shibuya CROSS TOWER 28th Floor 2-15-1, Shibuya, Shibuya-ku, Tokyo 150-0002, Japan

[2] Graduate School of Informatics, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto, Kyoto 606-8501, Japan

[3] Graduate School of Information Science and Technology, Hokkaido University, Kita 14, Nishi 9, Kita-ku, Sapporo, Hokkaido 060-0814, Japan

Abstract. We propose techniques for achieving the geographical navigation of historical events described in Web pages as "Virtual History Tour". First, we develop a method for extracting information on the historical events from the Web and organizing it into a chronological table. Our method can effectively handle ambiguous cases – homonyms and multiple location names in a sentence – by using the number of co-occurrences among events, person names, location names, and addresses in the Web. Next, we propose a method for ranking historical entities according to their impacts at specific time and location. We extend the PageRank algorithm to calculate the temporal and spatial impacts of entities. Finally, we introduce our concrete application demonstrating how users can browse historical events through timeline and map interfaces.

1 Introduction

In recent years, there has been a huge amount of historical information placed on the Internet that can be accessed easily. Using this information, people can easily gather Web pages related to a certain historical person or happening.

We propose a method to aggregate such historical information in the Web and to generate a chronological table for a given historical person or happening. In particular, the proposed method tries to identify the location and the time of each event in the chronological table wherever possible. We also develop an application called "Virtual History Tour", which achieves the navigation of the extracted events on a map interface with automatic movements in chronological order. This is a novel means of providing information in which users can trace the life of a specified person or collect all the historical incidents that occurred at a specified location.

In order to achieve this service, we need a technique for extracting all the historical events from the Web and for determining their location and time by inferring from their descriptions that is often written in an ambiguous and symbolic way or is often missing. In addition, we need a technique to add "impact" on the output extracted from the Web information.

In this paper, we explain the techniques we developed and their applications as follows.

- A technique to extract information on historical incidents relative to a specific incident from the text available in the Web, and to classify it by time and location: in particular a technique for data inference with the help of "Collective Intelligence".
- A technique to compute the "Impact" of historical entities.
- A demonstration to the application called "Virtual History Tour",that we developed, to evaluate how available information can be used with the above techniques.

2 Related Work

There are several environments for the visualization of a historical event sequence. Our system is implemented on Google Earth[4], which is an application that allows for the easy creation of geographical applications. Google Map[5], a map service on the Web, provides an API that can be easily used to create many kinds of geographical Web services. Stoev et al. [18] provides a visual environment especially for the historical events.

There are several related techniques and applications that are proposed in the research field of Geographic Information Retrieval (GIR), which is defined by Larson [9, 11]. Strötgen et al. [19] proposed a method for extracting spacio-temporal information from a single document. They also developed a system to explore an extracted sequence of spacio-temporal information on a Google Map interface. While their target is a single document, our proposed method is to aggregate information from many documents available on the Web. Several other similar applications have been proposed by Larson's research group [12, 10, 5, 17]. They focus on historical biography and visualize a sequence of life events in time and space. Their system places the historical events of a target person on a Google Map interface. The historical events related to a location are explained by clicking a place marker. The explanation usually contains temporal information, but it is difficult to find an explanation that describes the events in chronological order. Geocoding is a technique for identifying the latitude and longitude for a given geographic document. There are several researches regarding this technique [1, 2]. We propose a method for extracting the location of an event by using an existing geocoding system.

[4] Google Earth: http://earth.google.com/
[5] Google Maps: http://maps.google.com/

Information extraction is another research area related to our work. In particular, Web information extraction has become a large research field. For example, Paşca proposed methods to extract factual information from a large Web corpus [16] and query logs [15]. Although most of the research studies focus on *current* fact extraction, some focus on *historical* fact extraction. Garera et al. [4] proposed a method for biographic fact extraction such as "birthdate", "occupation", "nationality", and "religion". BioSnowball by Liu [13] is a method for gathering and integrating biographical facts. Their method can construct Wikipedia-stype pages for any person by aggregating the extracted facts. The values of the attributes of a person can change over time. When extracting the values from the Web, such changes can also be considered [14]. For temporal information extraction, not only the past events but also the future events can be extracted. Jatowt et al. [7] proposed a method for future event extraction from news articles or Web pages. They mainly focused on a method for extracting expressions that predict the future.

3 Historical Event Extraction

3.1 Extracting Historical Events from the Web

This section explains a method called "Historical Event Extraction" which extracts chronological events relevant to a specific person or happening, and classifies them according to time and location. In our previous work [8], we developed a method to gather chronological events about a specific person or happening and to estimate the time when each extracted event occurred. Adding to the previous work, we developed a method for estimating the location of an extracted event.

Fig. 1 shows the procedure for extracting historical events related to a specific person or happening. It consists of three steps as follows:

1. Collecting Web pages relevant to a specific person.
   Let a user specify a word that represents the target person or happening, which is used as an inquiry keyword on a conventional Web search engine such as Google[6] and Yahoo![7]. Then collect as many Web pages that as possible that hit on the query.
2. Adding chronological information to an extracted historical event.
   From the Web sites collected in step 1, search words or expression that can be a clue to adding chronological information, and modify these words or expressions in a predefined way.
3. Adding geographical information to an extracted historical event.
   Extract the name of the event and its location by applying segmentation to the information acquired in step 2, and determine the longitude and latitude where the event occurred.

[6] Google: http://www.google.com/
[7] Yahoo! Search: http://search.yahoo.com/

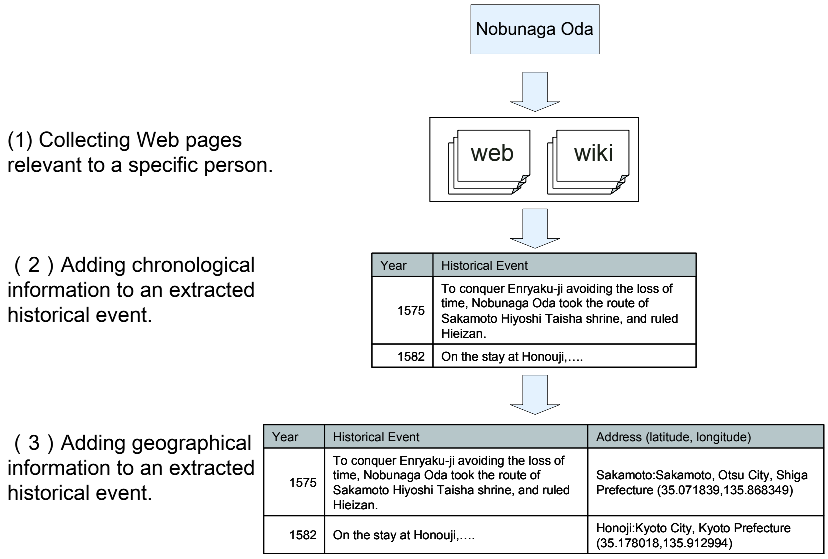

Fig. 1. Procedure for Chronological Event Extraction

Steps 1 and 2 are basically the same as the methods described in the previous work [8] except for some enhancements such as the use of better regular expressions to extract temporal information. These methods can efficiently aggregate and normalize temporal data that is written in a different style.

The step 3, the location estimation of historical events, is one of the contributions of this paper. In many cases, it is possible to determine the specific locations of the extracted events by using a method called "Location Estimation of Event with Collective Intelligence" The rest of this section describes and evaluates this method.

3.2 Location Estimation of Historical Event

Although the name of a certain location can vary depending on the time or era, we were able to safaly determine its location from the description safely by using other resources such as historical materials. However, there still exist some problems that need to be solved as follows.

The context problem: Multiple different location names in a single sentence It is often the case where there are more than just one location name appears on a single sentence. For example, think of sentence below. *"To conquer Enryaku-ji avoiding the loss of time, Nobunaga Oda took the route of*

Sakamoto Hiyoshi Taisha shrine, and ruled Hieizan." [8] If you apply automated segmentation to this sentence and extracts all the words associated with the geographical word, you will get three location names: "loss", "Sakamoto", and "Hieizan". Here, "loss" is not location name, but "loss" and "Los" of *Los Angeles* is repsented in same way in Japanese, so it is extracted as location name. Obviously, "loss" was extracted by mistake, so we need to determine which of the remaining the candidates to choose.

The homonym problem: The same name that has different locations For example, if you search geocoding with the keyword "Sakamoto", you will get two hits such as (1) Sakamoto-chyou, Yasiro City and (2) Sakamoto Otsu City. If you are familiar with this event, you should correctly choose (1), but this is not always the case.

To overcome these problems, we propose a technique called "Location Estimation of Events with Collective Intelligence".

3.3 Location Estimation of Event with Collective Intelligence

In our estimation method, we focus on the name of the place and its address in the sentence. Here, we assume that the name of the place and its address appears within the same document in the Web that describes the historical person and event. The notation we use in this section is summarized as follows.

- T_i denotes one of the location names extracted from sentence that explain a historical event.
- A_j denotes a text expression of one of the addresses retrieved from geocoding with a location T_i.
- H denotes a set of people's names extracted from the sentence.
- E denotes a set of happenings' names extracted from the sentence.

The following formula is a scoring function that calculate relevance between a location name and an address based on our assumption.

$$\begin{aligned} Score(T_i, A_j) &= P(T_i|H,E) \cdot P(A_j|H,E) \\ &= \frac{N_{H,E}(T_i)}{N_{H,E}} \cdot \frac{N_{H,E}(A_j)}{N_{H,E}} \\ &= \frac{WebHitCount(T_i,H,E)}{WebHitCount(H,E)} \cdot \frac{WebHitCount(A_j,H,E)}{WebHitCount(H,E)} \end{aligned} \quad (1)$$

where $P(A_j|H,E)$ is the conditional probability of the address A_j among $N_{H,E}$ documents, and $P(T_i|H,E)$ is the conditional probability of the location T_i among $N_{H,E}$ documents. They can be calculated by $N_{H,E}$, $N_{H,E}(T_i)$, and $N_{H,E}(A_j)$. $N_{H,E}$ is defined as the number of the document that includes the

[8] *Nobunaga Oda* is one of the most famous feudal rulers in Japanese in the late 16th century. (http://en.wikipedia.org/wiki/Oda_Nobunaga)

Table 1. $P(T_i|H,E)$

T_i	$P(T_i\|H,E)$
loss	0.0192
Sakamoto	0.3728
Hieizan	0.8065

Table 2. $P(A_j|H,E)$

A_j	$P(A_j\|H,E)$
Shiga Prefecture	0.2753
Otsu City	0.2674
Sakamoto	0.3728
Kumamoto Prefecture	0.0412
Yashiro City	0.0015
California	0.0014
Los Angeles	0.0087

Table 3. Score

T_i \ A_j	Shiga Prefecture	Otsu City	Sakamoto	Kumamoto Prefecture	Yashiro City	California	Los Angeles
loss						0.0000	0.0002
Sakamoto	0.1026	0.0997	0.1389	0.0154	0.0001		
Hieizan	0.2220	0.2156	**0.3006**	0.0332	0.0012		

name of the person, H, and the event. $N_{H,E}(T_i)$ is the number of the document among $N_{H,E}$ documents that includes the name of the location. $N_{H,E}(A_j)$ is the number of the document among $N_{H,E}$ documents that includes the name of the address. We can obtain $N_{H,E}$, $N_{H,E}(T_i)$, and $N_{H,E}(A_j)$ by using a conventional Web search engine such as Google and Yahoo!, which returns *hit count* with a search result. Our implementation uses Google AJAX Search API[9]. Each of the words in T_i, A_j, H, and E is added to the keyword query as *AND* condition.

For the event sentence we show in Section 3.2, T_i can be either "loss","Sakamoto" or "Hieizan". When T_i is "loss", A_j can be "California, Los Angeles". H is "Nobunaga Oda", and E is empty.

Our proposed method is unique in the way that we estimate the location of the event by multiplying a term $P(T_i|H,E)$ with $P(A_j|H,E)$. Because each probability term is acquired independently in terms of T_i and A_j, this covers the case where the only address A_j appears but not the location name T_i appears in the document, and vice versa. After multiplication, we calculated Score defined by Eq.(1) above for each possible candidate address A_j. We selected the address A_j with the highest *Score* as the most likely address where the event occurred,along with its latitude and longitude. We show some examples of scores computed by the proposed method in Fig.3. In this example, the highest score is observed for T =Hieizan and A =Sakamoto, yielding the fact that Sakamoto Otsu City is most strongly related in the given set.

---

[9] Google AJAX Search API: http://code.google.com/intl/en/apis/ajaxsearch/

Table 4. Example of historical persons and events used in the experiment

Search keyword	Historican event name of location(bold)	Person H	Event name E
Yositsune Miyamoto	Yoritomo ordered samurais in **Kanto** to serve in **Kyoto** and forbided returning to **EastArea** on April,15h.	Yoritomo	serving at Kyoto
Ryoma Sakamoto	In 1866, on November, Sakamoto persuaded **Satsuma** Lord Saisuke 5th and **Choshu** Lord Heisuke Hirosawa on the agreement to establish a joint venture at **Shimonoseki**.	Heisuke Hirosawa	The establishment of a joint venture

Table 5. Results

	number of correct
Conventional method	17(42.5%)
Proposed method	25(62.5%)

3.4 Evaluation of Location Estimation

In order to see the validity of our proposed method, especially the method proposed in Section 3.3, we evaluated the precision of location estimation.(conducted on September, 30, 2009) in the following manner.

1. We picked up three well-known historical persons such as “Yoshitsune Miyamoto”, “Nobunaga Oda”, and “Ryoma Sakamoto”, and collected 40 historical events associated with these three people by performing a Google search.Here, historical events are chosen under the restriction in which both the name of the person and name appear in.
2. We estimated the locations of these 40 events based on our proposed method.
3. After having these events plotted on a map with their estimated locations, we asked a single examinee to evaluate the result subjectively.

The table bellow shows some of the events among the 40 events were used.

The evaluation result as shown in Table 5, appears to be 20% more precise than a conventional method. Table 6 show the results of the estimation of the location names of the event. As the data shows, we were still unable to estimate the 37.5% of location of events. We think this is because of the following reasons: failure of the sentence segmentation to extract the correct geographical words, and the domination of the major name of places over the remaining minor places, which can be mistakenly treated as a noise and discarded. This will lead to the conclusion that our method alone is not so suitable to deal with minor events where value N is small, and we are hoping to handle minor event case with other inference method to enhance its availability.

Table 6. Result of Estimation

Search keyword	Historican event, name of location(bold)	Estimated location
Yositsune Miyamoto	Yoritomo ordered samurais in **Kanto** to serve in **Kyoto** and for bided returning to **EastArea** on April	Kyoto Prefecture
Ryoma Sakamoto	In 1866, on November, Sakamoto persuaded **Satsuma** Lord Saisuke 5th and **Choshu** Lord Heisuke Hirosawa on the agreement to establish a joint venture at **Shimonoseki**.	Shimonoseki City

4 Computing "Impacts" of Historical Entities

The impact size of a certain historical entity changes with time. There are also geographical impacts, for example an event that influences only Japan can occur.

In this section, a historical entity means a historical happening, place, person, or item. We assume that a historical article on Wikipedia represent the related historical entity. In this section, we propose an approach that assesses the temporal and spatial impacts of historical entities by using the Wikipedia link structure.

4.1 Wikipedia Link Structure Analysis

Review of PageRank The main idea behind PageRank [3] is that a Web page is important if several other important Web pages point to it. This means that if page u has a link to page v, then the link implicitly confers some importance to v. How should we represent the value of conferred importance? Let $Rank(p)$ represent the degree of importance of page p, and let N_p represent the outdegree of page p. We can simply assume that all links are equal, so the link (u, v) confers $Rank(u)/N_u$ units of importance from u to v. This simple idea leads to the following recursion. If N is the number of pages, assign all pages the initial value $1/N$. Let B_v represent the set of pages that points to v. In each iteration, the ranks are shifted as follows:

$$\forall v \; Rank_{i+1}(v) = \sum_{u \in B_v} \frac{Rank_i(u)}{N_u} \tag{2}$$

We continue the iterations until all $Rank(p)$ stabilizes to within some threshold, although the convergence of the recursion depends on the link structure. We add the damping factor, α, to the rank propagation to guarantee the convergence. We define a new recursion, in which we add the constant value $(1-\alpha)/N$ to all pages:

$$\forall v \; Rank_{i+1}(v) = \alpha \sum_{u \in B_v} \frac{Rank_i(u)}{N_u} + \frac{1-\alpha}{N} \tag{3}$$

This modification improves the quality of PageRank by introducing the damping factor α, as well as facilitates to guaranteeing the convergence to a certain value for all pages.

Biased PageRank We can bias the PageRank computation to increase the effect of certain categories of pages by changing the added value $(1-\alpha)/N$ described in above. In particular, let D_j be the set of pages in a certain category c_j. We introduce a new damping factor d_{ij} for the page i:

$$d_{ij} = \begin{cases} 0 & i \notin D_j \\ \frac{1}{|D_j|} & i \in D_j \end{cases} \tag{4}$$

Then when computing the PageRank, we calculate as follows:

$$\forall v \; Rank_{i+1}(v) = \alpha \sum_{u \in B_v} \frac{Rank_i(u)}{N_u} + (1-\alpha) d_{vj} \tag{5}$$

In this recursion, pages that exactly fit a category and pages that are closely related get high scores.

The biased PageRank algorithm is also used in Topic-Sensitive PageRank [6]. For biasing, the 16 top-level categories of the Open Directory Project (ODP)[10] are used. The set of URLs included in each category is used as D_j.

We create the temporal and spatial categories for Wikipedia historical entities and discuss how D_j for category c_j is determined in Sections 4.2 and 4.3.

4.2 Temporal Impact Calculation

As discussed in Section 4.1, the basic idea implemented by a graph-based ranking model that includes PageRank is that a link of a Web page to another one has a social meaning (i.e. "voting", "recommendation", etc.) In addition, if a graph domain is specified, we are able to think on the domain specific meanings of the link. In our study, each node on a graph is a historical entity given by a Wikipedia article, so the links between these articles are not only a part of the World Wide Web, but also a meaningful relation between the two historical entities. We consider the linked relation as an impact of the history.

Our main idea is that a historical entity has a large impact on certain period if several other entities occurred at the same time have links to it. In our approach, we treat a certain period as a category c_j and a set of historical entities occurred in the period as D_j in order to use biased PageRank algorithm. We believe that the given score has good correspondence to the historical impact of the entity in c_j.

[10] Open Directory Project: http://www.dmoz.org/

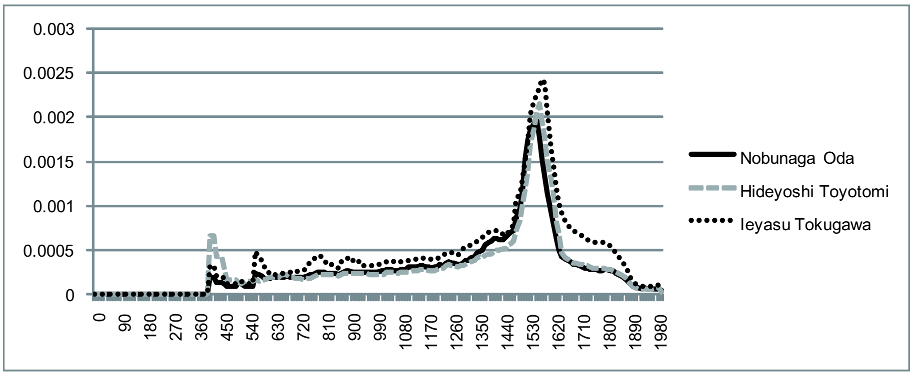

Fig. 2. Impacts of "Nobunaga Oda", "Hideyoshi Toyotomi", and "Ieyasu Tokugawa". The horizontal axis means A.D., not a category's index j. The vertical axis shows scale of biased PageRank.

4.3 Spatial Impact Calculation

Similar to the temporal categories created to calculate the temporal impacts of historical entities, we also create spatial categories to calculate the spatial impacts of them. For example, the geographic information of Wikipedia entries and the "history of somewhere" Wikipedia categories are particularly useful.

4.4 Experiments

As a dataset for our evaluation, we used 33,000 Japanese Wikipedia entries which all belong to subcategories of the "History of Japan" Wikipedia category. An element of the dataset represents a historical entity. There are 593,000 links among these pages.

We divide 0 to 2009 into 201 ranges at 10 intervals and name each ranges $c_0 \cdots c_{200}$, respectively. To be specific, the first range c_0 means 0 to 9, and the last one c_{200} is 2000 to 2009. D_j is a set of historical entities which occurred in c_j. For example, D_{194} includes "Pacific War" which broke out at 1941. We make a vector of 201 dimensions for all entities with these calculation.

The temporal impacts of the tree most important leaders: "Nobunaga Oda", "Hideyoshi Toyotomi", and "Ieyasu Tokugawa" in the Warring States period (16th century) are plotted in Fig.2. For all of them, the highest value is observed while they alive. During Edo period (1603-1868) "Tokugawa Ieyasu" has higher score than the other two. We think that this is the influence of his career; he established the Edo shogunate, which has great significance in Japanese history.

Fig.3 presents a comparison after the 19th century between the impacts of "Hiroshima Peace Memorial" and "Pacific War". The score of "Pacific War" reaches its peak around 1950; on the other hand the score of "Hiroshima Peace Memorial" is still increasing. We think that this is the influence of its designation as the UNESCO World Heritage Site in 1996.

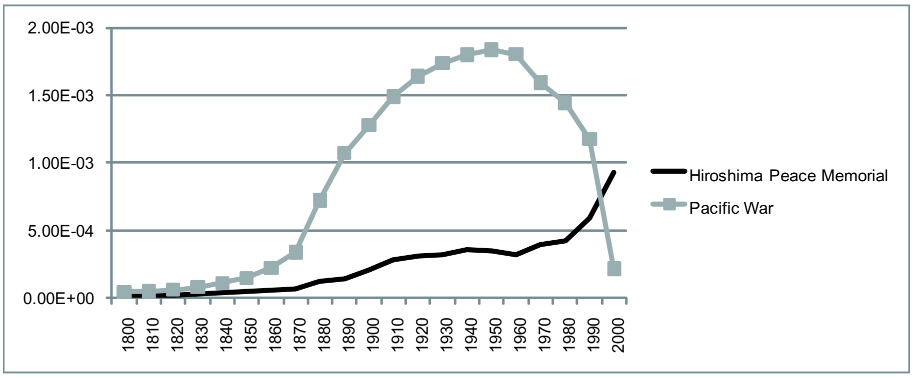

Fig. 3. Impacts of "Hiroshima Peace Memorial", and "Pacific War". Since these values are almost zero, the values of before the 19th century are not plotted in the figure.

5 Geographical Navigation of Historical Events

5.1 Application Overview

In this section, we explain the developed application, "Virtual History Tour", which enables users to explore the historical events that are organized and sorted by chronological and geographical order, having been collected from Web sites. This application is features the following functions

1. The user can trace back relevant events in chronological order that are associated with a specified person or event.
2. The user can have historical events plotted on various types of maps ranging from an ancient map to a satellite photo and view them.
3. The user can browse photos associated with a specified event that are collected automatically from Web sites.

5.2 System Structure and Procedure

Virtual History Tour is composed of the following three parts:

1. History Time Table Generator
   By using the technique "Historical Event Extraction", this generates a time table of events associated with a specified person and event is generated.
2. History Tour Planner
   This creates a tour plan along with the time table acquired by the Time Table Generator, together with photos collected from Web sites.
3. Browser
   This allows the user to browse the tour planned above.The current prototype works with Google Earth. Its procedure is shown in Fig.4.

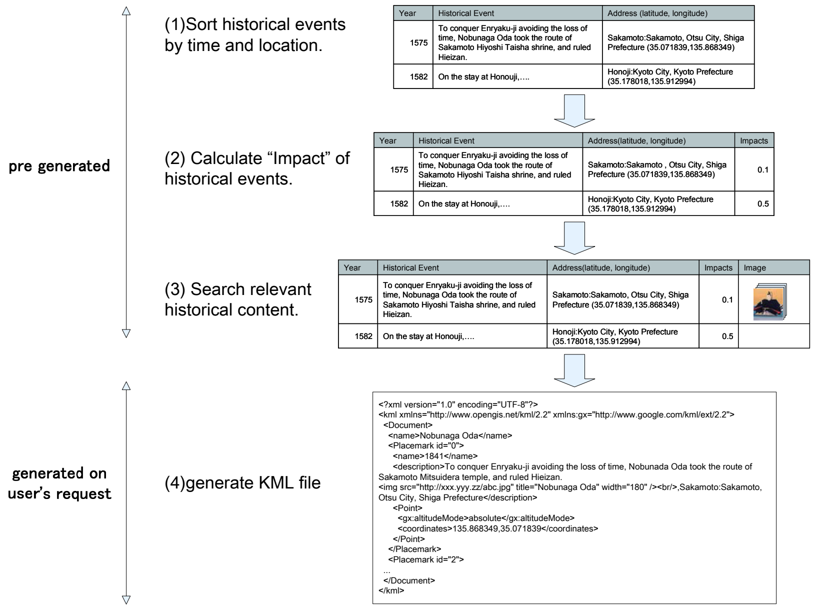

Fig. 4. Procedure for Historical Tour

5.3 Layout of Virtual History Tour

We show below, the result of tracing Ryouma Sakamoto.[11] Fig.5 shows the events associated with him that are plotted on the current map. Fig.6 shows the historical event details plotted on the old map. Fig.7 shows the superimposition of the Kyoto Edo-period old map on the current map.

5.4 Discussion

With our application, “Historical Event Extraction”, one can follow certain person's life along time axis and his geographical range over his life, and learn to what geographical and chronological extend a single person or event has spread its influence.

At last, we consider the further possible use of our applications. First, it can offer information about recommended historical famous spots to stop by on traveling planning site. With this service, use can be aware of all historical spots around his destination, he can make more flexible plan of palaces to go.

This can be also applied to Car Navigation System, which offers drivers an unexpected history guide, by following a trait of historical figures, replacing the job done by bus guide concierge.

[11] The demonstration video is avairable in the following URL. http://www.dl.kuis.kyoto-u.ac.jp/˜ohshima/historicalnavi/demo.wmv

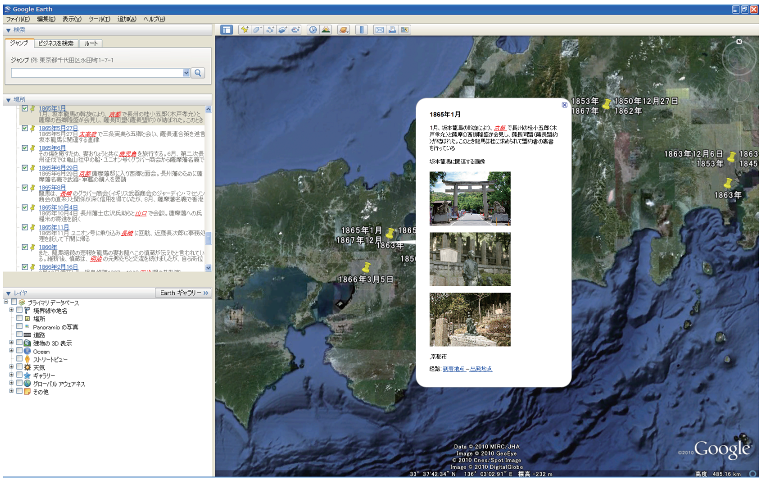

Fig. 5. Events associated with Ryoma Sakamoto plotted on the current map.

This guide feature can be further improved by adapting user's routine or preference, it lets a user decide what to learn next with his own, which helps user to learn history more pleasant way.

6 Conclusions

We proposed techniques for achieving the navigation of historical events described in Web pages. First, we developed a method for extracting chronological tables from the Web pages. The conventional method had difficulties in handling homonyms and multiple location names in a sentence. By using collective intelligence based on the number of co-occurrences among events, person names, location names, and addresses in the Web, our method can effectively overcome these difficulties. Next, we proposed a method for evaluating the impacts of historical entities. We extended the PageRank algorithm to calculate the temporal and spatial impacts of entities. Finally, we showed that our new application "Virtual History Tour" enables users to browse historical events through timeline and map interfaces. Our future work includes incorporating the temporal and spatial impacts of entities into history tours. We also plan to extend our system to other languages besides Japanese to see if it works.

Acknowledgements

This work was supported in part by the following projects and institutions: Grants-in-Aid for Scientific Research (Nos. 18049041, 21700105, and 21700106) from MEXT of Japan, a Kyoto University GCOE Program entitled "Informatics

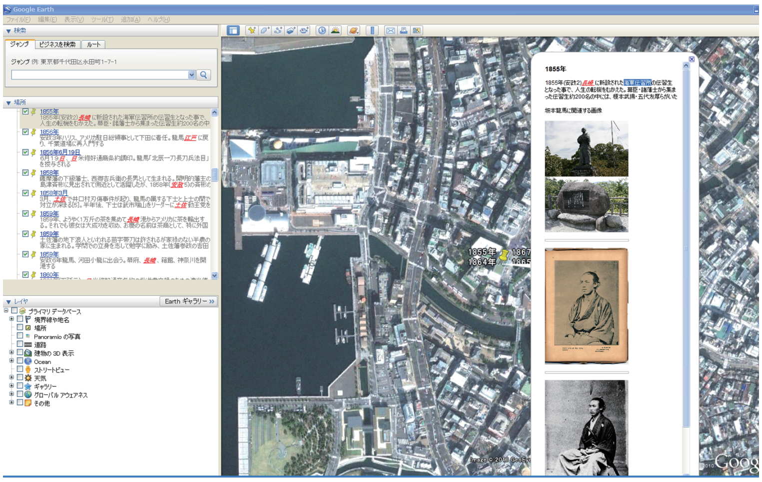

Fig. 6. Historical event detail plotted on the current map.

Education and Research for Knowledge-Circulating Society," and the National Institute of Information and Communications Technology, Japan.

References

1. Angel, A., Lontou, C., Pfoser, D.: Qualitative geocoding of persistent web pages. In: Proc. of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS 2008). pp. 1–10 (November 2008)
2. Borges, K.A.V., Laender, A.H.F., Medeiros, C.B., Jr., C.A.D.: Discovering geographic locations in web pages using urban addresses. In: Proc. of the 4th ACM Workshop on Geographical Information Retrieval (GIR 2007). pp. 31–36 (November 2007)
3. Brin, S., Page, L.: The anatomy of a large-scale hypertextual web search engine. In: Proc. of the 7th International Conference on World Wide Web (WWW 1998). pp. 107–117 (April 1998)
4. Garera, N., Yarowsky, D.: Structural, transitive and latent models for biographic fact extraction. In: Proc. of the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2009). pp. 300–308 (March 2009)
5. Gey, F., Shaw, R., Larson, R., Pateman, B.: Biography as events in time and space. In: Proc. of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS 2008). pp. 1–2 (November 2008)
6. Haveliwala, T.H.: Topic-sensitive PageRank: A context-sensitive ranking algorithm for web search. IEEE Transactions on Knowledge and Data Engineering (TKDE 2003) 15(4), 784–796 (July/Aug 2003)
7. Jatowt, A., Kanazawa, K., Oyama, S., Tanaka, K.: Supporting analysis of future-related information in news archives and the web. In: Proc. of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2009). pp. 115–124 (June 2009)

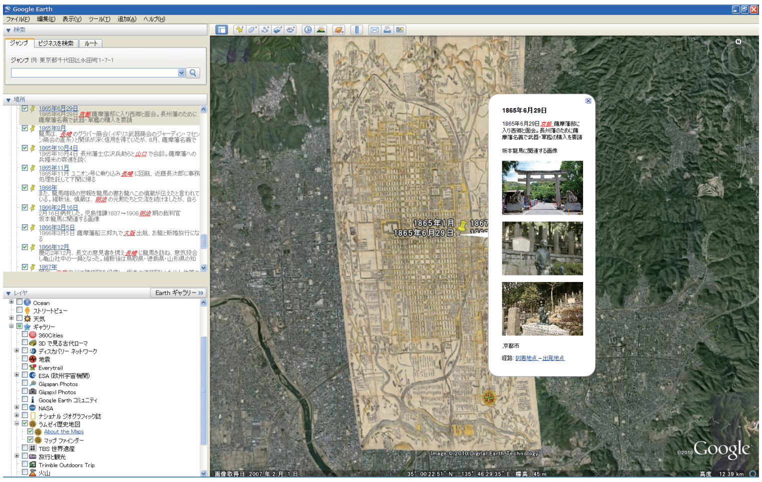

Fig. 7. Superimposition of Kyoto Edo-period old map to current map.

8. Kimura, R., Oyama, S., Toda, H., Tanaka, K.: Creating personal histories from the web using namesake disambiguation and event extraction. In: Proc. of the 7th International Conference on Web Engineering (ICWE 2007). pp. 400–414 (July 2007)
9. Larson, R.R.: Geographic information retrieval and spatial browsing. GIS and Libraries: Patrons, Maps and Spatial Information pp. 81–124 (April 1996)
10. Larson, R.R.: Geographic ir and visualization in time and space. In: Proc. of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2008). p. 886 (July 2008)
11. Larson, R.R., Frontiera, P.: Geographic information retrieval (gir): Searching where and what. In: Proc. of the the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2004). p. 600 (July 2004)
12. Larson, R.R., Shaw, R.: Mapping life events: Temporal and geographic context for biographical information. In: Proc. of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2009). pp. 471–472 (June 2009)
13. Liu, X., Nie, Z., Yu, N., Wen, J.R.: Biosnowball: Automated population of wikis. In: Proc. of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2010). pp. 969–978 (July 2010)
14. Oyama, S., Shirasuna, K., Tanaka, K.: Identification of time-varying objects on the web. In: Proc. of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2008). pp. 285–294 (June 2008)
15. Paşca, M.: Organizing and searching the world wide web of facts – step two: Harnessing the wisdom of the crowds. In: Proc. of the 16th International Conference on World Wide Web (WWW 2007). pp. 101–110 (May 2007)
16. Paşca, M., Lin, D., Bigham, J., Lifchits, A., Jain, A.: Organizing and searching the world wide web of facts – step one: the one-million fact extraction challenge. In: Proc. of the 21st National Conference on Artificial Intelligence (AAAI 2006). pp. 1400–1405 (July 2006)

17. Petras, V., Larson, R.R., Buckland, M.: Time period directories: A metadata infrastructure for placing events in temporal and geographic context. In: Proc. of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2006). pp. 151–160 (June 2006)
18. Stoev, S.L., Feurer, M., Ruckaberle, M.: Exploring the past: a toolset for visualization of historical events in virtual environments. In: Proc. of the the ACM Symposium on Virtual Reality Software and Technology (VRST 2001). pp. 63–70 (November 2001)
19. Strötgen, J., Gertz, M., Popov, P.: Extraction and exploration of spatio-temporal information in documents. In: Proc. of the 6th ACM Workshop on Geographic Information Retrieval (GIR 2010). pp. 1–8 (February 2010)