

ChronoView: Visualization Technique for Many Temporal Data

Satoko Shiroy
Department of Computer Science
University of Tsukuba
Ibaraki, Japan
shiroy@iplab.cs.tsukuba.ac.jp

Kazuo Misue
Faculty of Engineering,
Information and Systems
University of Tsukuba
Ibaraki, Japan
misue@cs.tsukuba.ac.jp

Jiro Tanaka
Faculty of Engineering,
Information and Systems
University of Tsukuba
Ibaraki, Japan
jiro@cs.tsukuba.ac.jp

Abstract—This paper presents a method of visualizing data that contains temporal information, such as a human’s behavior and the time at which it occurs. A feature of the data is that each event may have one or more time-stamps. By analyzing this kind of data, we are able to find some behavioral patterns and obtain knowledge applicable to many fields, such as marketing research and security. We develop ChronoView, a visualization technique to support the analysis of data with temporal information. ChronoView represents an event with a set of time-stamps as a position inside a circle, similar to the dial of an analog clock. By representing each event as a position on a two-dimensional plane, we can simultaneously visualize many events and easily compare their occurrence patterns. We implement a tool based on ChronoView, which is enriched with additional functions and overcomes the drawbacks of the original system. A use case involving tweet data from Twitter illustrates the use and practicality of ChronoView.

Keywords—temporal information; event; time-stamp; periodicity

I. INTRODUCTION

In this paper, we describe a system for visualizing data that contains temporal information. The data consists of a set of pairs describing an event and its occurrence time. For example, we can consider a web access log as data with temporal information, because the log records which pages were accessed and at what time. Other examples include product sales histories, tweet data from Twitter, and incident logs at a call center. By analyzing these data, we can obtain some knowledge about the occurrences of events. For example, we may notice that some events occur at specific times. Such knowledge is useful in various fields such as market research, security, and lifelogging.

Generally, an event occurs once or more and then has one or more time-stamps. Suppose that access to a web page is an event. If a web page is accessed at 9 p.m., 10 p.m., and 11 p.m., the event (that is, the access to the web page) has three time-stamps.

Many visualization techniques for temporal information have been proposed [1]. Most of them represent each time-stamp individually, but do not represent sets of time-stamps. Therefore, it is not easy to find dependency on a specific time range or the periodicity of the event occurrences.

We propose a visualization technique to represent a set of time-stamps as positions on a two-dimensional plane. This positional representation enables us to visualize many events simultaneously and makes it easy to find their temporal features. Our visualization technique is based on a circle, like an analog clock, so we call it “ChronoView.” Visualization techniques based on a circle have been used for many years, and these techniques have recently been applied in various domains [2].

II. RELATED WORK

A. Visualization of Data with Temporal Information

Aris et al. aimed to treat unevenly spaced time series data and developed several visualization methods [3], and Buono et al. developed a visualization tool for the interactive analysis of time series data [4]. These tools enabled us to know when events occurred and to find the co-occurrence of several events by visualizing individual time-stamps. Krstajić et al. developed CloudLines to deal with time-based representations of large and dynamic event data sets in limited space [5]. Their representation adapts to the data by employing a decay function to let items fade away according to their relevance. However, it is difficult to find the periodicity of single events or compare the periodicities of a number of events.

Nguyen et al. developed visualization techniques for temporally referenced tags that exploit their text size and brightness [6]. Additionally, another variation uses the background color of tags and color-coded cells, as in a calendar. Using these techniques, we can easily compare occurrences of events if we focus on a specific date and time. However, it is difficult to understand the periodicity of the events and to compare their occurrences when focusing on the combined conditions of years, months, days, and time.

B. Circular and Spiral-Type Visualization

Misue proposed an Anchored Map representing bipartite graphs [7]. On an Anchored Map, nodes in one of two sets are placed on the circumference and nodes in the other set are placed at suitable positions to represent their relationships to adjacent nodes. Restricting nodes in one set

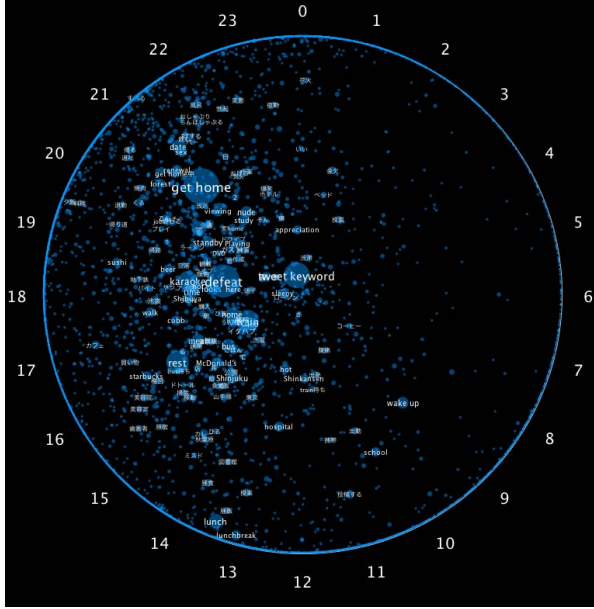


Figure 1. Overview of ChronoView.

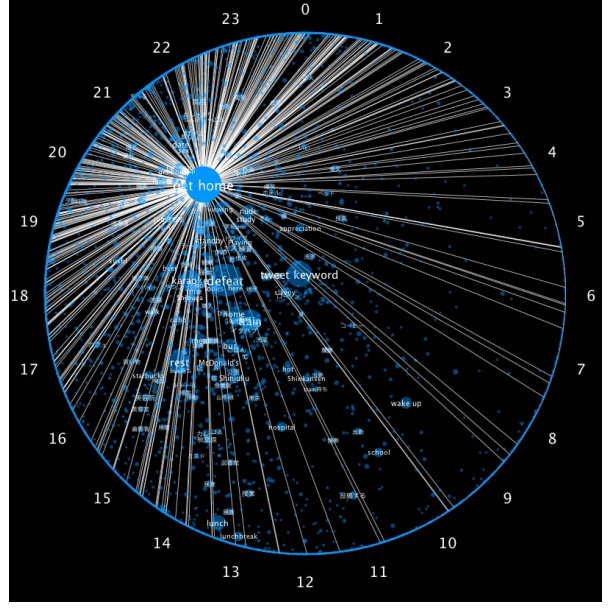


Figure 2. Example showing the distribution of events using the barycenter of time-stamps.

to the circumference makes grasping the relationships among them easier. ChronoView can be considered as a variation of an Anchored Map. However, while the Anchored Map is a technique for graph drawing, ChronoView is a technique for visualizing data with temporal information.

Carlis et al. developed spiral visualization to represent both the continuity of time and the periodicity of a week, a month, or a year [8]. This technique allows us to grasp the frequency and periodicity of events. Tominski et al. improved the existing spiral representation using a two-tone pseudo-coloring method and interaction [9]. With this technique, we can analyze multivariate data with some periodicity easily. However, it is not easy to compare many events when the number of spirals increases.

Dragicevic et al. developed SpiraClock, which shows events using an analogy of an analog clock but with a spiral representation [10]. This technique is useful for events over a few hours, but unreasonable for events in the longer term.

III. CHRONOVIEW

ChronoView is a technique for visualizing many events with one or more time-stamps. This technique represents each event as a position on a two-dimensional plane. Figure 1 shows an overview of ChronoView.

A. Representation of sets of time-stamps

First, we give a position on the circumference of a circle of radius r to each time-stamp. We give the twelve o'clock position on an analog clock to the time-stamp t_0 . Moving clockwise, we then place all time-stamps on the circumference according to the time elapsed from t_0 . If we

suppose that U is a set of all time-stamps, the positions of the time-stamps are expressed by the function $f_0 : U \rightarrow \mathcal{R}^2$.

$$f_0(t) = (r \cos \theta, r \sin \theta) \quad (1)$$

where

$$\theta = \frac{\pi}{2} - 2\pi \frac{t - t_0}{c}.$$

Suppose that T is a finite set of time-stamps, that is, $T = \{t_1, t_2, \dots, t_n\} \in 2^U$. We assign the barycenter of the positions of t_i ($i = 1, \dots, n$) to the set T . The positions of sets of time-stamps are expressed by the function $f : 2^U \rightarrow \mathcal{R}^2$.

$$f(T) = \frac{1}{|T|} \sum_{t \in T} f_0(t) \quad (2)$$

Thus, if an event e has a finite set of time-stamps T , e is placed at $f(T)$.

In Figure 1, 0:00 a.m. (midnight) is set to t_0 and a period c is 24 h. When c is 1 h, ChronoView represents a diagram of a one-hour period. When c is 168 h (i.e., seven days), it represents a diagram of a one-week period. In a similar way, ChronoView can represent a one-month or one-year period. However, we need a slight modification to create such diagrams, because the length of months and years are not actually constant.

For example, let us consider the case where $c = 24$ h. Figure 3 shows two cases. If an event occurs at 0:00, 8:00, and 16:00, the event is placed at the center of the circle. If another event occurs at 18:00 and 22:00, the event is placed

IV. DEVELOPMENT OF THE TOOL

We developed a visual tool to enable ChronoView to support the analysis of data with temporal information, using tweet data from Twitter as analysis objects. Figure 4 shows an overview of the tool.

A. ChronoView

Our visual tool places ChronoView in the center of the window. The tool provides three types of view: *all view*, *weekday-and-holiday view*, and *day view*.

All view shows all of the collected data, with each light-blue circle expressing an event. If we click on a circle, radial lines are drawn from the circle, as shown in Figure 2, allowing us to understand the actual occurrence time of the event.

In *weekday-and-holiday view* (see Figure 5), events are divided into two categories: events on weekdays and events on holidays. Each light-blue circle expresses an event on a weekday and each pink circle expresses an event on a holiday. In this view, we can examine the difference between occurrences on weekdays and holidays.

In *day view* (see Figure 6), events have been divided into seven categories, i.e., one for each day of the week. Events on different days are assigned a unique color. If we click on a circle, curved lines are drawn that connect circles corresponding to the same events. In this view, we can see how occurrence times for the same events differ across a week.

B. Detailed View

If we click on an event in ChronoView, detailed information about it is displayed in the sub-window on the right-hand side. When our target is tweet data from Twitter, the detailed information includes handle names, tweet text, tweet date, and so on.

C. Operation View of the Representation

The sub-window on the left-hand side provides three radio buttons for switching between the three view types in ChronoView and sliders for changing some view parameters. By operating the sliders, users can specify the maximum and minimum frequency of displayed events, and the opacity of the circles representing events.

If very many events are placed at a similar position, the view becomes so complicated that it is difficult to distinguish all events at the same time. In such cases, filtering the events by occurrence frequency is effective for retaining ease of observation.

If very many events with similar occurrence frequencies are placed at a similar position, filtering by frequency does not work well. In such cases, it is effective to adjust the opacity of the circles representing events. The more circles are crowded, the lower the opacity to be effective.

V. USE CASE

We analyzed tweet data from Twitter as a use case of ChronoView.

A. Data

We used tweets containing the keyword “now” in Japanese, such as “Cafe, now” and “Shinjuku, now.” “Now” is an English word, but Japanese tweeters often use it to express their current situation. Tweets containing “now” are considered as signs of casual, daily behavior. Therefore, analysis of such data is expected to be useful in the field of market research.

We collected tweet data using StreamingAPI from Twitter¹. This data contains the user name, date, day, time, and tweet text. We used MeCab² for a morphological analysis of tweets containing “now,” and picked out clauses before “now” that expressed human behavior. Table I shows some examples of tweet data.

We consider two sets of tweets. One set includes 18,584 tweets collected from 20th to 26th June 2011, which contains 6,318 different words. The other set includes 19,805 tweets collected from 13th to 19th September 2011. This data contains 9,494 different words. We analyzed and compared the events in both sets.

Our program of data collection using StreamingAPI displayed unstable behavior. Some parts of the data have not been completely retrieved, but we do not believe that this seriously influences the illustration of our use case. Of course, it is desirable that the data is complete to make the analysis more reliable.

B. Observations

From the ChronoView of tweet data from 20th June, we found that the size of the events “wake up” and “lunch” are almost the same. This means both events have similar frequencies. However, “wake up” is positioned farther from the circumference than “lunch.” We can infer from this that the occurrence time of “wake up” has a broad spread. On the other hand, “lunch” was placed near to the circumference, suggesting that this event is concentrated at a specific time.

“Wake up” was placed near a position from 6:00 to 12:00. This event occurred over a wide range of time. In contrast, “lunch” was concentrated between 11:00 and 14:00—obviously, “lunch” depends on a specific time.

According to the radial lines, we found that “wake up” had also occurred between 18:00 and 21:00 and “lunch” had also occurred between 16:00 and 19:00. From detailed information about these events, we were able to find some reasons for such outliers. These suggested that waking up late and eating lunch late were caused by jobs and studying.

We compared the data from 20th June with data from 13th September. We found that daily events such as “wake

¹<https://dev.twitter.com/docs/streaming-api>

²<http://mecab.sourceforge.net/>

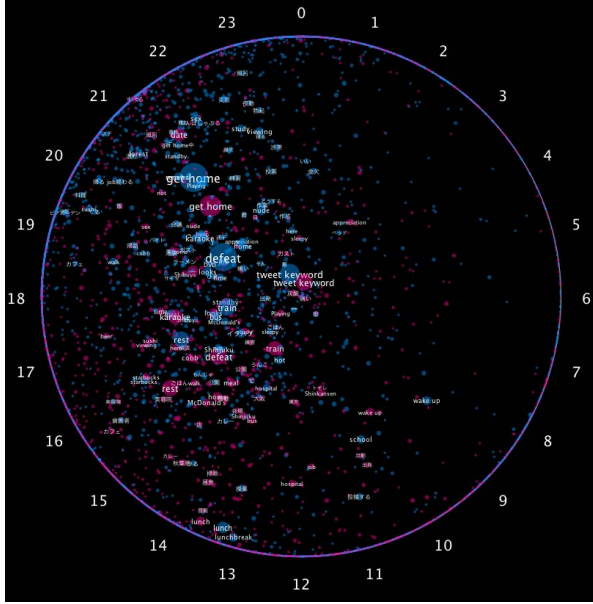


Figure 5. Events on weekdays and holidays are distinguished by colors.

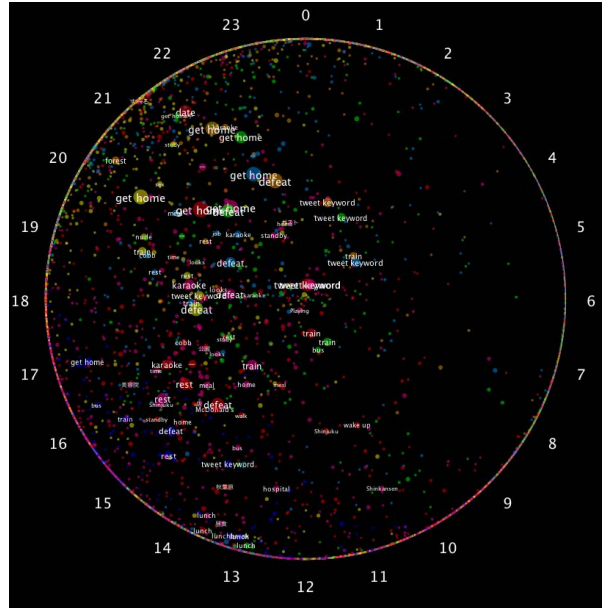


Figure 6. Events on different days of the week are distinguished by colors.

Table I
EXAMPLES OF DATA INCLUDING TEMPORAL INFORMATION.

User	Date	Day	Time	Tweet	Event
white_luc	20/06/2011	Mon	17:30:24	Staying in cafe, now.	cafe
yamahakusyon	20/06/2011	Mon	17:33:09	Got home, now.	get home
adajmdap	20/06/2011	Mon	17:35:12	Staying in Shinjuku	Shinjuku

up,” “lunch,” and “get home” were placed at almost the same positions in both sets of data. However, the “wake up” event has a noticeable feature. In the data from 20th June, this event occurred later on holidays than on weekdays, but in the 13th September data, the event occurs later on weekdays than on holidays. Thus, the rhythm of people’s lives might have changed.

By comparing the data from 20th June with data from 13th September, we found some events occurred frequently only in a term. These events are found in the data of one period, while they are not found in the other period.

We found the event “cobb” at the left-center of the view in the 20th June data, and inferred from its position that the event had occurred between 12:00 and 21:00. When we clicked the circle, radial lines were seen to be concentrated from 12:00 to 15:00 and from 17:00 to 21:00. We may conclude that this event occurred at lunchtime or dinnertime for many people.

From the day view, we found that the event had occurred on Thursday, Friday, and Saturday. By studying the detailed view of the “cobb” event, we came to understand that a major hamburger shop had used Twitter to promote their new product.

C. Discussion

We could understand some rough features of the events using ChronoView. Examples of these features are not only how often events occurred, but also whether an event depends on a specific time and when an event has occurred. Even if two or more events have almost the same frequency, we can find some differences between them, like their dependency on a specific time and periodicity.

We were able to actually understand the occurrence time, and the distribution of the time about an event, by adding radial lines from the circle corresponding to the event. The radial lines overcome some of ChronoView’s ambiguity.

Exploiting three types of view, as shown in the use case, we can observe the differences in event occurrence times for each day of the week and for weekdays and holidays. We can understand details of the events by seeing the original tweet text. In this way, we can analyze the different features of daily events to give an overview and breakdown of the details from various views in the ChronoView interface.

It is possible to understand distinct features of daily events in different periods by switching between the three view types and using the radial line representation. Therefore,

we found that we could show different features of the same event in ChronoView by comparing multiple data from different periods.

VI. CONCLUSION

In this paper, we proposed ChronoView, a technique to visualize data containing temporal information. One of the most important features of ChronoView is that it represents a set of time-stamps as a position on a two-dimensional plane. This can be used to visualize many events with one or more time-stamps. Some attributes of the events can also be represented visually. Using this technique, we could grasp some features about the periodicity of the events' occurrence. Moreover, we were able to observe the distribution of an event's time-stamps by adding radial lines and switching between the three view types.

We performed a use case, which focused on tweet data from Twitter, to demonstrate the usefulness of ChronoView. This showed the usefulness and effectiveness of ChronoView by analyzing tweet data. ChronoView can be applied not only to tweet data, but also to other data with temporal information, e.g., access logs of Web pages and the sales history of products. In future work, we will display the versatility of ChronoView by performing several case studies.

REFERENCES

- [1] Sônia Fernandes Silva, Tiziana Catarci: Visualization of Linear Time-Oriented Data: A Survey, *Proceedings of the First International Conference on Web Information Systems Engineering (WISE'00)*, Vol. 1, pp. 310–319 (2000).
- [2] Geoffrey M. Draper, Yarden Livnat, Richard F. Riesenfeld: A Survey of Radial Methods for Information Visualization, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 15 Issue 5, pp. 759–776 (2009).
- [3] Aleks Aris, Ben Shneiderman, Catherine Plaisant, Galit Shmueli, Wolfgang Jank: Representing Unevenly-Spaced Time Series Data for Visualization and Interactive Exploration, *HumanComputer Interaction INTERACT 2005*, pp. 835–846 (2005).
- [4] Paolo Buono, Aleks Aris, Catherine Plaisant, Amir Khella, Ben Shneiderman: Interactive Pattern Search in Time Series, *Proceedings of Conference on Visualization and Data Analysis (VDA 2005)*, pp. 175–186 (2005).
- [5] Miloš Krstajić, Enrico Bertini, Daniel A. Keim: CloudLines: Compact Display of Event Episodes in Multiple Time-Series, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 17, pp. 2432–2439 (2011).
- [6] Dinh Quyen Nguyen, Christian Tominski, Heidrun Schumann, Tuan Anh Ta: Visualizing Tags with Spatiotemporal References, *15th International Conference Information Visualisation (IV2011, July 12-15, 2011, London, UK)*, pp. 32–39 (2011).
- [7] Kazuo Misue: Overview of Networks Information by Using Anchored Maps, *11th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES2007)*, pp. 1269–1276 (2007).
- [8] John V. Carlis, Joseph A. Konstan: Interactive Visualization of Serial Periodic Data, *Proceedings of the 11th annual ACM symposium on User interface software and technology (UIST'98)*, pp. 29–38 (1998).
- [9] Christian Tominski, Heidrun Schumann: Enhanced Interactive Spiral Display, *Proceedings of the Annual SIGRAD Conference 2008*, pp. 53–56 (2008).
- [10] Pierre Dragicevic, Stephane Huot: SpiraClock: A Continuous and Non-Intrusive Display for Upcoming Events, *CHI'02 Extended Abstracts on Human Factors in Computing systems (CHI EA'02)*, pp. 604–605 (2002).
- [11] J. Bertin: *Semiology of Graphics: Diagrams, Networks, Maps*, Univ. of Wisconsin Press (1983).