

ChronoView: A Method to Overview Many Temporal Patterns

Kazuo Misue, Member, IEEE

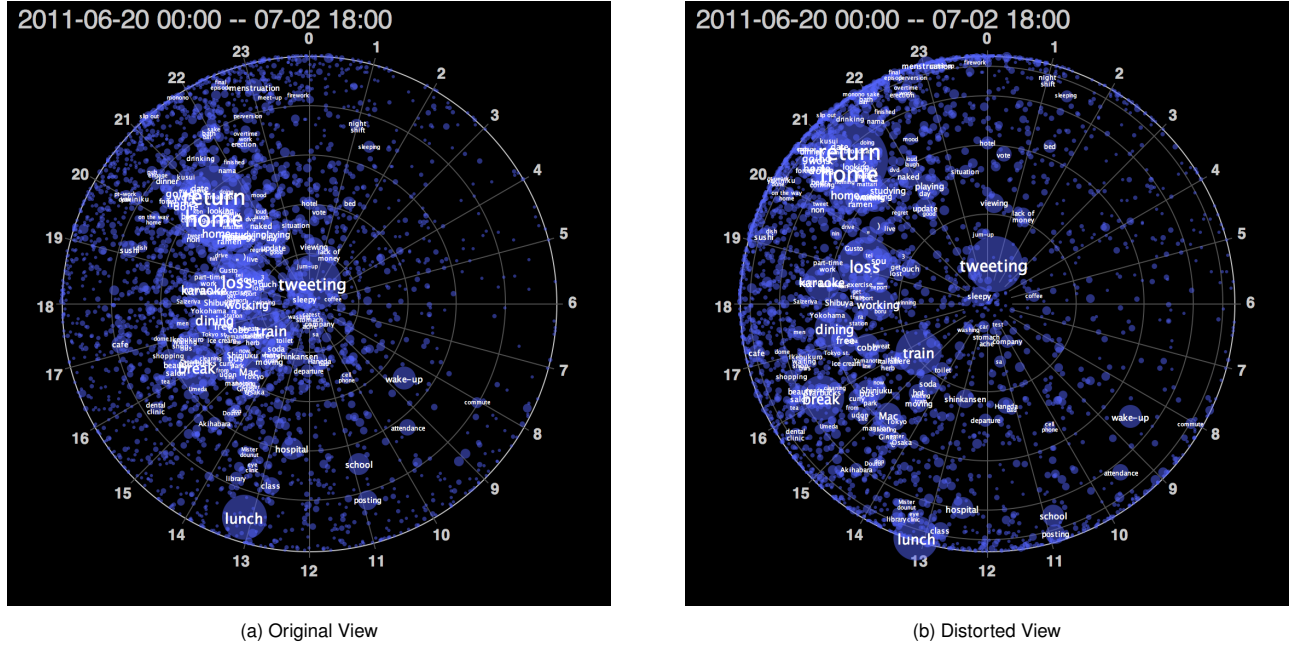


Fig. 1. ChronoView representation of temporal tweet patterns (see Section 6 for more information).

Abstract—A temporal pattern is a set of timestamps that accompanies events. Temporal patterns contain diverse valuable information and suggest tendencies of human activities, occurrences of accidents, etc. Therefore, it is important to identify distinguishing features of temporal patterns in event data. Many methods exist for representing time-oriented data. However, it is difficult to observe many temporal patterns at a glance because most naive methods require a certain area to visualize a temporal pattern. This paper presents a space-efficient technique for visualizing temporal patterns that facilitates simultaneous visualization of many event groups. Moreover, the limitation of this technique is experimentally studied and a modification is proposed to mitigate it.

Index Terms—Information visualization, Time-oriented data, Temporal patterns, Overview, ChronoView.

1 INTRODUCTION

Numerous events occur daily: web pages are visited, tweets are sent, purchases are made, photos are taken, and hospitals are visited. Many of these events are recorded as log data. A set of such events is associated with a set of timestamps. This paper focuses on temporal patterns that are features of timestamps. Fig. 2 shows examples of temporal patterns. In the figure, each orange circle represents a timestamp. There are three patterns: A, B, and C. These patterns have the following features: Pattern A is independent of the time of day, B is concentrated in the morning, and C is extremely concentrated around 8 p.m.

Determining the temporal features of events is very important. Consider the following questions: Given a sales log, can we determine when the products were sold? Given an incident log, can we determine when the site was attacked? These are not simply questions about time, but also about features of timestamps, i.e., temporal patterns. To determine temporal features, we need to overview temporal patterns.

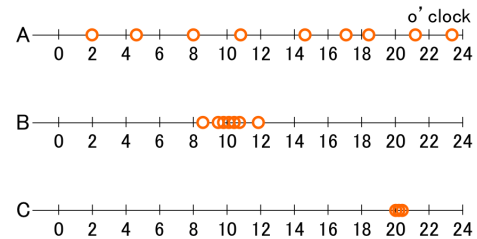


Fig. 2. Examples of temporal patterns

Shneiderman [6] proposed an information-seeking mantra in which an overview is first performed followed by zooming and filtering, and then details-on-demand are extracted. In the early stage of visual tasks, we need a suitable overview of the data.

However, two kinds of problems are associated with providing an overview of many temporal patterns. One problem is that a temporal pattern occupies space. To illustrate a temporal pattern, strip charts (Fig. 2), histograms, and line charts may be used. The size of these visual representations often makes it difficult to encode many visual patterns in a limited space. The other problem is the difficulty in finding features. One method to observe many temporal patterns at a glance

• Kazuo Misue is with the University of Tsukuba.
E-mail: misue@cs.tsukuba.ac.jp.

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org.
Digital Object Identifier: xx.xxx/TVCG.201x.xxxxxxx

is to lay out all possible histograms on a display. However, this does not help us discover similar patterns and outliers because it illustrates features of individual patterns.

In this paper, we present ChronoView, a method to visualize many temporal patterns simultaneously. The defining feature of ChronoView is that each event group (i.e., set of events) is represented by a position on a two-dimensional plane. This enables high space efficiency when many patterns are involved; however, there exists ambiguity in representation. We performed some experiments to characterize the ambiguity and considered measures to reduce it.

2 RELATED WORK

Aris et al. aimed to treat unevenly spaced time series data and developed several visualization methods [1], while Buono et al. developed a visualization tool for the interactive analysis of time series data [2]. These tools enabled us to know event occurrences and find the co-occurrence of several events by visualizing individual timestamps. Krstajić et al. developed CloudLines to deal with time-based representations of large and dynamic event data sets in limited space [7]. Their representation adapts to the data by employing a decay function to let items fade away according to their relevance. Nguyen et al. developed visualization techniques for temporally referenced tags that exploit their text size and brightness [8].

Dragicevic et al. developed SpiraClock, which displays events in a manner analogous to that of an analog clock but with a spiral representation [4]. Carlis et al. developed a spiral visualization to represent both the continuity of time and the periodicity of a week, month, or year [3]. 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]. These techniques help in easily analyzing multivariate data containing some periodicity.

To obtain an overview of many temporal patterns, we need more space-efficient methods.

3 CHRONOVIEW

We developed “ChronoView” [5], a method that facilitates the visualization of many temporal patterns in a limited area.

3.1 Basic Idea

The layout rule used by ChronoView is very simple. Each event group with a set of timestamps is plotted according to the following procedure:

1. Prepare a plane similar to an analog clock with a suitable period. In Fig. 3, the period is 24 h ($c = 24h$).
2. Represent timestamps as positions on the circumference of the clock face circle.
3. Place each event group at the barycenter of all timestamps associated with the group.

For example, assume that there are three event groups: $A = \{e_{1:00}\}$, $B = \{e_{4:00}, e_{8:00}\}$, and $C = \{e_{12:00}, e_{18:00}, e_{20:00}\}$, where $e_{1:00}$ denotes an event with timestamp 1 a.m. As shown in Fig. 3, Group A is placed at the 1 a.m. position because A is a singleton; Group B is placed at the barycenter of 4 a.m. and 8 a.m.; Group C is placed at the barycenter of 12 p.m., 6 p.m., and 8 p.m.

3.2 Formalization

In a sales log, let us assume that a teenage boy buys a beverage at time t . This is considered an event and is accompanied by timestamp t . Events that satisfy certain conditions are considered to be a part of an event group. For example, all events of buying bottles of orange juice constitute one event group while all events in which teenagers make purchases constitute another group. A temporal pattern is a set of timestamps accompanying an event group.

Let us now formally describe ChronoView. Function f_0 determines the position of a timestamp on the circumference. Function f takes

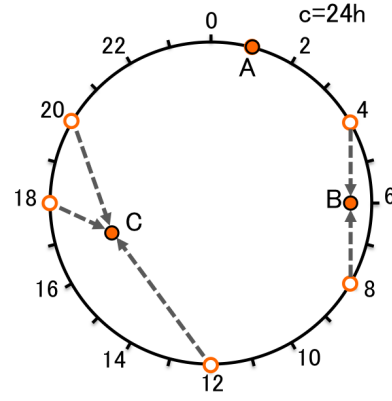


Fig. 3. ChronoView layout rule

an event group and associates it with a position. This position is the barycenter of the group’s event timestamp positions.

Function f_0 places all timestamps on the circumference, while moving clockwise according to the time elapsed from t_0 . If we consider U to be the set of all timestamps, the position of a timestamp is expressed by the function $f_0 : U \rightarrow R^2$, given by

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

where r denotes the radius of the clock face circle. When the period is c ,

$$\theta_t = \frac{\pi}{2} - 2\pi \frac{t - t_0}{c}. \quad (2)$$

Suppose E is an event group (a finite set of events), i.e., $E = \{e_1, e_2, \dots, e_n\}$ and $t(e)$ denotes the timestamp of event e . We assign the barycenter of the positions of $t(e_i)$ ($i = 1, \dots, n$) to the set E . The position of an event group can be represented by the function

$$f(E) = \frac{1}{|E|} \sum_{e \in E} f_0(t(e)). \quad (3)$$

3.3 Potential Features and Limitations

One of the most important features of ChronoView is that a set of timestamps is represented by a position on a two-dimensional (2D) plane. This provides several advantages:

- Easy to read features of temporal patterns: Position is known to be excellent for representation of values [10].
- High space efficiency for many patterns: Each set is expressed by a point; an area parameter is essentially not used.
- Transformation to a density plot. A density plot, for example, a heat map, is available for many patterns.

Unfortunately, ChronoView also has limitations in that a set is represented by a position on a 2D plane — a kind of dimension reduction — causing ambiguity in representation. It is acceptable to place two similar temporal patterns at positions close to each other. However, a problem that needs rectification is that the two patterns that it places close to each other are not always similar. One way of doing this is to add additional marks to show detailed patterns. Our tool [12] provides a function to show the details of timestamps.

4 DISTRIBUTION ON THE CLOCK FACE CIRCLE

Many temporal patterns seem to occupy positions close to the center of the clock face circle. We exhaustively generated many temporal patterns under some conditions and plotted them on the circle. Fig. 4 illustrates the results of plotting the following temporal patterns.

T5L6: At each of the five timestamps {00:00, 04:48, 09:36, 14:24, 19:12}, we assume 6 levels of frequencies {0, 1, 2, 3, 4, 5}. The number of temporal patterns (i.e., dots) is 7,775 ($= 6^5 - 1$).

T12L4: At each of the 12 timestamps {00:00, 02:00, 04:00, ..., 22:00}, we assume 4 levels of frequencies {0, 1, 2, 3}. The number of temporal patterns is 16,777,215 ($= 4^{12} - 1$).

T24L2: At each of the 24 timestamps {00:00, 01:00, 02:00, ..., 23:00}, we assume 2 levels of frequencies {0, 1}. The number of temporal patterns is 16,777,215 ($= 2^{24} - 1$).

Fig. 5 shows the distribution of temporal patterns on the clock face circle. In each histogram, the horizontal axis represents the distance from the center, and the height of a bar represents the number of temporal patterns placed in the area. These histograms show that many temporal patterns tend to occupy the central region.

5 DISTORTION OF THE CLOCK FACE CIRCLE

Introduction of distortion methods like fisheye views may resolve visual clutter near the center. Assignment of more spaces around the center might decrease visual clutter by increasing the resolution of the central area. We consider moving the positions of temporal patterns using some non-linear mapping. Now, assume that the radius of the clock face circle is 1 for simplicity. When the original positions are expressed by (r, θ) in the polar coordinate system, we move the points to $(d(r), \theta)$ using non-linear mapping $d: [0, 1] \rightarrow [0, 1]$. To assign more space around the center, we enlarge the region near 0 by choosing concave function d . We consider the following three kinds of functions:

$$d_1(r) = \sqrt{r} \quad (4)$$

$$d_2(r) = \sin\left(\frac{\pi}{2} \cdot r\right) \quad (5)$$

$$d_3(r) = 1 - (1 - r)^2 \quad (6)$$

By using these functions, the distributions are changed to the ones shown in Fig. 6 and Fig. 7. Data sets in these figures are all T5L6. The original distribution is shown in Fig. 4(a) and Fig. 5(a). Function d_1 results in a sparse space near the center. Function d_2 may fill the center area. By observing the characteristics of these functions, we generate function d_3 . This function can distribute the dots more evenly on the circle.

Actual distributions depend on the target data, with some target data possibly requiring more space around the center while other requiring more resolution peripherally. Function d_3 is suitable for such extension. Function d_4 is a generalized form of d_3 by replacing the exponent to a .

$$d_4(r) = 1 - (1 - r)^a \quad (7)$$

When $a = 1$, the function becomes $d_4(r) = r$ and every dot is placed at the original barycenter (see Fig. 4(a) and Fig. 5(a)). When $a = 2$, the function becomes equivalent to d_3 . When $a > 2$, the function enlarges the center area to a greater extent. When $a < 1$, the function enlarges the peripheral area.

6 EXAMPLES

Fig. 1 shows two examples of ChronoView, both of which represent tweets on Twitter.

We regard tweets containing the keyword “now” in Japanese, such as “lunch, now” and “Tokyo, now,” as events. “Now” is an English word, but Japanese tweeters often use it to express their current situation. Tweets containing “now” convey casual, quotidian behavior. Therefore, analysis of such data is expected to be useful in the field of market research.

Our data includes about 30,000 tweets collected in ten days from June 20, 2011. It contains about 10,000 different words, i.e., 10,000 event groups. To generate the examples, we translated Japanese words in the tweets into English.

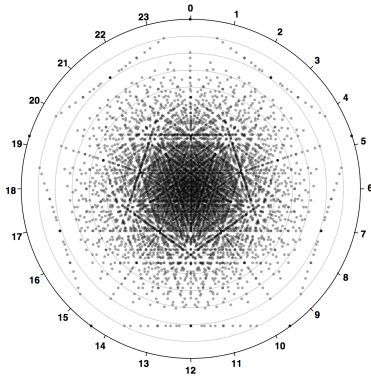
Each event group is represented by a circle, the size of which is proportional to the number of events. The examples in Fig. 1 only show event groups including two or more events. Labels are added only for event groups with 16 or more events to reduce visual clutter. Fig. 1(a) shows the original view of ChronoView, with all event groups placed at the barycenter of all included event timestamps. Fig. 1(b) shows a distorted view by using function d_4 with $a = 2$ (i.e., d_3). We can observe event groups placed near the center more clearly than in Fig. 1(a).

7 CONCLUSION

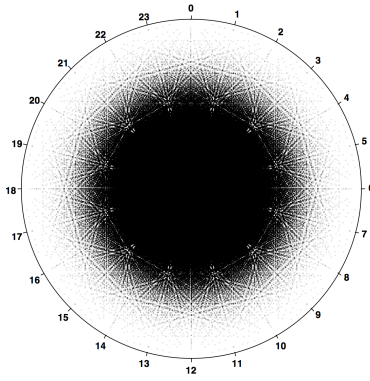
This paper presented ChronoView, which visualizes many temporal patterns. The defining feature of ChronoView is that each event group is represented as a position on a two-dimensional plane. This provides several advantages, including easy-to-read temporal pattern features, and high space efficiency for many patterns. This representation of sets by positions on a 2D plane causes ambiguity, which tends to occur near the center of the clock face circle. We exhaustively generated many temporal patterns under some conditions, plotted them on the clock face circle, and confirmed the concentration of many dots near the central area. We also tried to introduce non-linear mapping to decrease clutter near the center. Non-linear mapping induces the position of each temporal pattern to move away from the barycenter of event occurrence times. We need to survey observers to understand the effect of these movements on their impressions on a decluttered variant of ChronoView.

REFERENCES

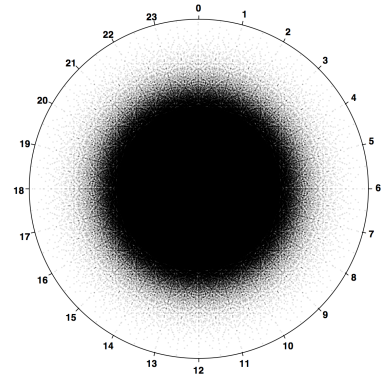
- [1] Aleks Aris, Ben Shneiderman, Catherine Plaisant, Galit Shmueli, and Wolfgang Jank: Representing Unevenly-Spaced Time Series Data for Visualization and Interactive Exploration, *In Proc. of Human Computer Interaction (INTERACT 2005)*, pp. 835–846, 2005.
- [2] Paolo Buono, Aleks Aris, Catherine Plaisant, Amir Khella, and Ben Shneiderman: Interactive Pattern Search in Time Series, *In Proc. of Conference on Visualization and Data Analysis (VDA 2005)*, pp. 175–186, 2005.
- [3] John V. Carlis and Joseph A. Konstan, Interactive Visualization of Serial Periodic Data, *In Proc. of the 11th annual ACM symposium on User interface software and technology (UIST 1998)*, pp. 29–38, 1998.
- [4] Pierre Dragicevic and Stephane Huot, SpiraClock: A Continuous and Non-Intrusive Display for Upcoming Events, *In Proc. of CHI'02 Extended Abstracts on Human Factors in Computing systems (CHI EA 2002)*, pp. 604–605, 2002.
- [5] Satoko Shiroy, Kazuo Misue, and Jiro Tanaka, ChronoView: Visualization Technique for Many Temporal Data, *In Proc. of 16th International Conference on Information Visualisation (iV2012)*, pp. 112–117, 2012.
- [6] Ben Shneiderman, The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations, *In Proc. of the IEEE Symposium on Visual Languages*, pp. 336–343, 1996.
- [7] Miloš Krstajić, Enrico Bertini, and 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.
- [8] Dinh Quyen Nguyen, Christian Tominski, Heidrun Schumann, and Tuan Anh Ta, Visualizing Tags with Spatiotemporal References, *In Proc. of 15th International Conference on Information Visualisation (iV2011)*, pp. 32–39, 2011.
- [9] Christian Tominski and Heidrun Schumann, Enhanced Interactive Spiral Display, *In Proc. of the Annual SIGRAD Conference 2008*, pp. 53–56, 2008.
- [10] Jock D. Mackinlay, Automating the Design of Graphical Presentations of Relational Information, *ACM Trans. of Graphics*, Vol. 5, No. 2, pp. 110–141, 1986.
- [11] Kazuo Misue, Yuta Kanai, and Hideaki Takagi, Visualizing the Overview of Temporal Patterns of Patients’ Activities, *In Proc. of 2013 Workshop on Visual Analytics in Healthcare (VAHC2013)*, pp. 11–14, 2013.
- [12] Kazuo Misue, ChronoView: Visualization of Many Temporal Patterns, <http://www.cs.tsukuba.ac.jp/~misue/open/software/chronoview/>



(a) T5L6: 5 timestamps, 6 levels

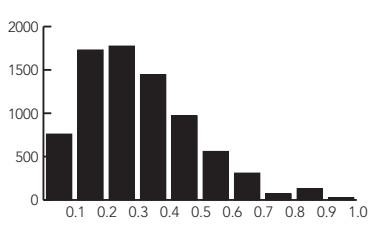


(b) T12L4: 12 timestamps, 4 levels

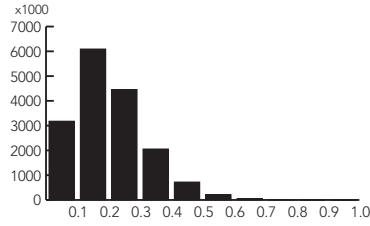


(c) T24L2: 24 timestamps, 2 levels

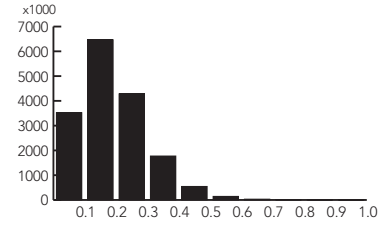
Fig. 4. Results of plotting many temporal patterns. Dataset induced variation.



(a) T5L6: 5 timestamps, 6 levels

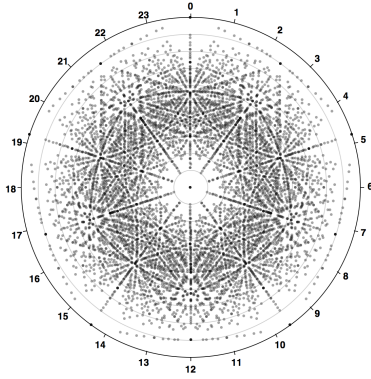


(b) T12L4: 12 timestamps, 4 levels

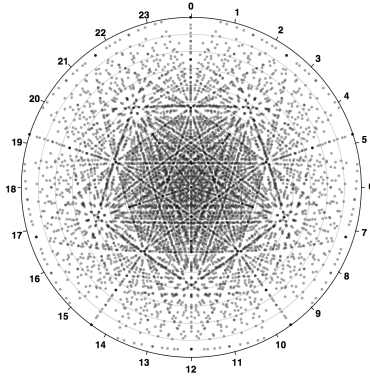


(c) T24L2: 24 timestamps, 2 levels

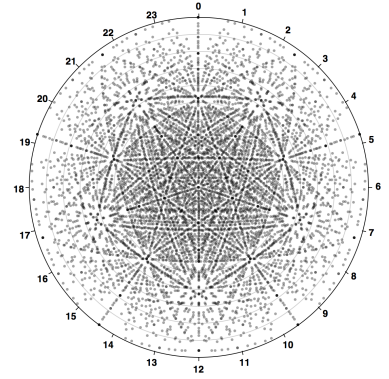
Fig. 5. Distribution of temporal patterns on the clock face circle. Dataset induced variation.



(a) $d_1(r) = \sqrt{r}$

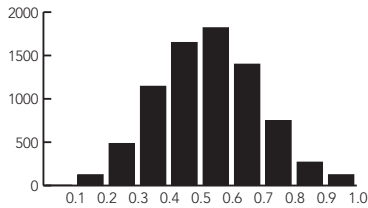


(b) $d_2(r) = \sin(\pi \cdot r / 2)$

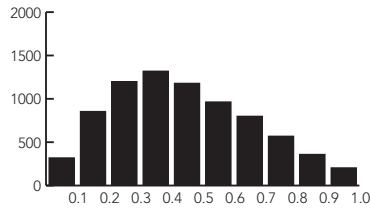


(c) $d_3(r) = 1 - (1 - r)^2$

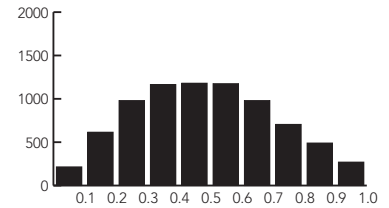
Fig. 6. Results of plotting many temporal patterns. Variations by mapping functions.



(a) $d_1(r) = \sqrt{r}$



(b) $d_2(r) = \sin(\pi \cdot r / 2)$



(c) $d_3(r) = 1 - (1 - r)^2$

Fig. 7. Distribution of temporal patterns on the clock face circle. Variations by mapping functions.