Visualization of Time & Time-Oriented Data, A Survey

Shashank Reddy P
500658817
Ryerson University
Toronto, Canada
Email: shashank.pandillapal@ryerson.ca

***Abstract*—Time is an important dimension with distinct characters and it is common across many application domains like medical, business, science, history etc. The hierarchical structure of time and its natural cycles, re-occurrences make time-oriented data to be treated differently than other kinds of data. It requires appropriate visual and analytical methods to explore and analyze them. In this paper, we survey many visualization techniques dealing with Time-oriented data to categorize and describe tasks the user seek to accomplish using visualization methods. Time & Time-oriented data visualization is an important concern in Visual Analytics, with this survey we try to help users and researchers identify right tasks and techniques to pick for their data in Visual Analytics.**

***Index Terms* - Time-Oriented Data, Visualization, Analysis, User Information, Graphs, Data Mining, Clustering, Visual Analytics.**

I. Introduction

The advances in computing and storage technologies have made it possible to create, collect and store huge volumes of data in variety of data formats, languages and cultures [1]. Analysis of such data to derive insight enables analysts of the data to design successful strategies and make right decisions. Data mining and Machine learning algorithms/mathematical approaches have been developed to automatically analyze the data. These approaches have proven their usefulness in many practical applications, but still face significant challenges like algorithm scalability, increasing data dimensions and data heterogeneity [2]. If complex, interesting patterns are discovered, it is usually difficult to understand and interpret the findings in an intuitive and meaningful manner [3].

To address these challenges, visual analytics has been developed in recent years through a proper combination of automated analysis and interactive visualizations [2]. The analysis of data should require humans to involve and interact with data representation to evaluate the data and respond in a timely manner. The first widely accepted road-map for visual analytics to meet practical requirements has been presented by Thomas and Cook in their seminal book [4], [5]. The book defines visual analytics as

> *"The Science of analytics reasoning assisted by interactive visual interfaces"*

After that the VisMaster Coordinated Action community, funded by the European Union, updated the road-map and provided a more specific definition of visual analytics:

> *"Visual Analytics combines automated analysis with interactive visualization for effective understanding, reasoning and decision making on the basis of a very large and complex dataset"* [3]

These pioneering researches [1], [5] defined the scope of field and discuss the future challenges that the field needs to face. A large number of visualization techniques have been developed subsequently after that. The rapid technological developments in this field have greatly affected the visualization area in promoting it to solve real-world problems, such as network traffic analysis, education, concepts, sports data, database analysis and biological data analysis. As a result, visual analytics has been gaining more and more attention from both industry and academia. With increasing advancement, development and popularity in this field there is an immediate need for a comprehensive survey covering the recent advances in the field.

The motivation in doing the survey is to show that the characteristics of data is vital when designing visual representations and one of the main salient characteristics being, whether data is related to time or not. Time is an outstanding dimension as reflected by Schneiderman's Task by Data Type Taxonomy[6] where temporal data are identified as one of the seven basic data types. A wide repertoire of interactive techniques for visualizing datasets with temporal dependencies is available however many visualization frameworks consider time as a common quantitative parameter rather than taking it as a special dimension. According to Thomas and Cook [5] it is in general a problem that

> *"Most visualization software is developed with incomplete information about the data and tasks. (...) New methods are needed for constructing visually based systems that simplify the development process and result in better targeted applications"*

In this paper we would like to delve into Time-series, Time and Time-oriented data, discuss about the various design aspects and their features which are important when modeling time, characterize the data based on time as a dimension and try to relate the data and time. We would then like to discuss about the visualization problems faced when dealing with time-oriented data and showcase some design examples at different levels. A survey on existing visualization techniques for

Time and Time-oriented data will be discussed, some specific to particular domain with each technique briefly describing the background, main idea and concept and application of the particular technique in a domain field. These visualization techniques are categorized based on data, time and visualization and for each criteria is further sub categorized to help user to pick the right technique for his data. We would then conclude on some results and observations based on the study, iterate some application issues, research challenges and estimate some future research in visual analytics domain needing advancement.

The paper would be organized as follows, in section II we model time based on some design aspects and their features. In section III we characterize the data component of the time-orientated data, in section IV we discuss some aspects of visualization and characterize the visualization problem by posing some practical questions and trying to answer them. Section V is the heart of this paper where we survey variety of visualization techniques and categorize them based on the discussions had in previous sections, this categorization is done to help users/researchers select the appropriate technique for their data. Section VI is where we draw out some observations based on the survey and finally in the following section with conclude with drawn out statements, application issues of techniques and areas which need further research.

II. Modeling Time

Time-series and Time-oriented is characterized by data elements being a function of time. In general, this data takes the following form:

$$D = \{(t_1, y_1), (t_2, y_2, ...(t_n, y_n)\} \quad (1)$$

with

$$y_i = f(t_i) \quad (2)$$

The data elements y_i represent different data types. Modeling time in information systems is not to perfectly imitate the physical dimension time, but to provide a model that is best suited to reflect the phenomena under consideration and support the analysis tasks at hand [7]. Time is modeled differently for different applications and there are many ways to model it in information systems. The major design aspects and their features which are important as per *Frank* and *Goralwalla et.al* [8], [9] are

1) Scale: ordinal vs discrete vs continuous
2) Scope: point-based vs interval-based
3) Arrangement: linear vs cyclic
4) Viewpoint: ordered vs branching vs multiple perspectives

A. Scale: ordinal vs discrete vs continuous

Time with scale as the design aspect is presented in 3 different ways. In *ordinal* time domain, only relative order relations are present (eg. before, after). This model would only be sufficient if qualitative temporal relationships are of interest

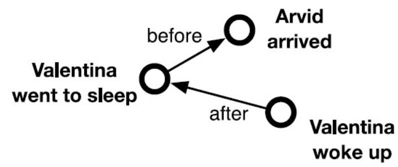

Fig. 1: Ordinal scale. Only relative order relations are present. At this level it is not possible to discern whether Valentina woke up before or after Arvid arrived.

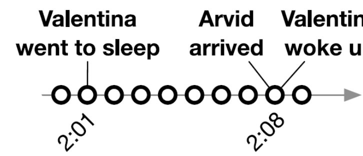

Fig. 2: Discrete scale. Smallest possible unit is minutes. Although Arvid arrived and Valentina woke up within the same minute, it is not possible to model the exact order of events.

or no quantitative information is available. This can be best understood from Figure-1 [7].

In *discrete* domains temporal distances can also be considered and time values can be mapped to a set of integers which enables quantitative modeling of time values. Discrete time domains are based on a smallest possible unit (eg. seconds or milliseconds) and they are most commonly used in information systems, check Figure-2 [7].

Continuous time models are characterized by a possible mapping to real numbers. Between any two points in time, another point in time exists example of this can be seen in Figure-3 [7].

B. Scope: point-based vs interval-based

Scope is the next basic element that form the structure of time domain. *Point-based* domains are discrete *euclidean* points in space having zero temporal extent, so no region exists between two points in time. Contrast to *Interval* based domains relate to subsections of time having a temporal extent greater than zero. Figure-4,5 [7] are examples of point and interval-based domains respectively.

C. Arrangement: linear vs cyclic

Arrangement is the third design aspect of the time domain. Based on our natural perception of time we either consider

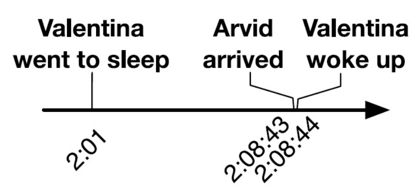

Fig. 3: Continuous scale. Between any two points in time, another point in time exists. Here, it is possible to model that Arvid arrived shortly before Valentina woke up.

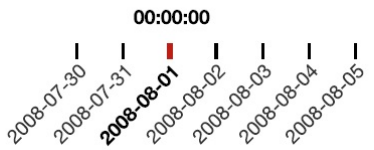

Fig. 4: Time value August 1, 2008 in a point-based domain. No information is given in between two time points.

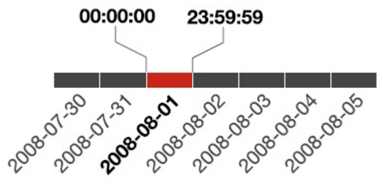

Fig. 5: Time value August 1, 2008 in an interval-based domain. Each element covers a subsection of the time domain greater than zero.

time moving in a *linear* manner i.e., from past to future (each time value has a predecessor and successor) shown in Figure-6 [8]. Or time moving in *Cyclic* manner meaning the data consists of recurring time values shown in Figure-7 [8].

D. Viewpoint: ordered vs branching vs multiple perspectives

The different views of time are modeled to be the final division of the design aspects. *Ordered* time domains consider things happening one after the other, in more detailed level we can even consider totally and partially ordered domains of time as-well. Totally ordered domain of time where only one thing can happen at a time but in contrast partially ordered domain simultaneous or overlapping events are allowed. *Branching* is the most complex form of this domain where multiple strands of time branch out and allow description and comparison of alternative scenarios. In contrast to branching where only one path through time will happen while *multiple perspectives* facilitate simultaneous views of time. An example of multiple perspectives are stochastic multi-run simulations. The Figure-8 shows an example of branching time and Figure-9 shows Multiple perspectives.

These design aspects are introduced to adequately model the time domain's scale, scope, arrangement and viewpoints before visualizing the data. Beside these general aspects, the hierarchical organization of time, characterization of data as

Fig. 6: Linear time. Time proceeds linearly from past to future.

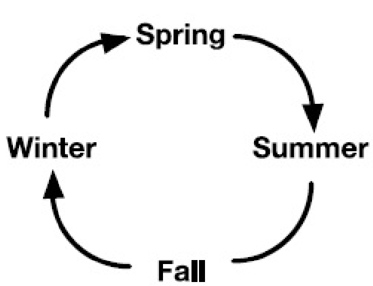

Fig. 7: Cyclic time. Set of recurring time values such as the seasons of the year.

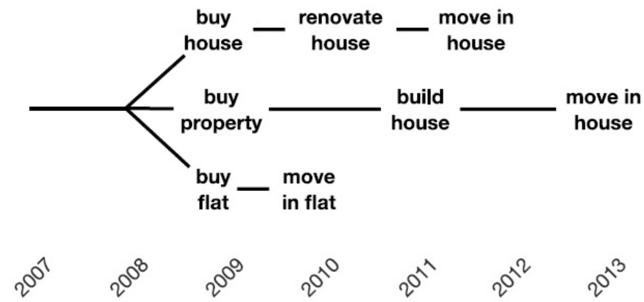

Fig. 8: Branching time. Alternative scenarios for moving into a new living space.

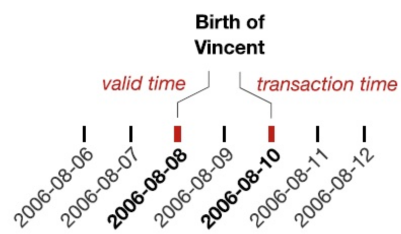

Fig. 9: Multiple perspectives. Vincent was born on August 8, 2006 (valid time) and this fact was stored in the register of residents two days later on August 10, 2006 (transaction time).

well as definition of concrete time elements used to relate data to time need to be specified [7].

III. Characterizing Data

After introducing the aspects of modeling time domain and discussing about it, we now consider characterizing time-oriented data. As we have mentioned data which has time primitives or in the format (1) are considered *time-oriented data* and the modeling approaches range from continuous to discrete data models [10].

One useful concept for modeling time-oriented data along cognitive principles is the *pyramid framework* [11] shown in Figure-10. This model is based on the three perspectives location(where is it?), time(when is it?) and theme(what is it made of?) at the level of data [11].

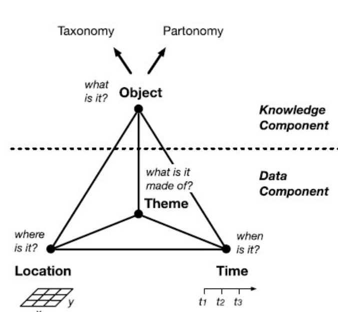

Fig. 10: Pyramid framework. Data are conceptualized along the three perspectives of location, time, and theme. Derived interpretations form objects on the cognitively higher level of knowledge. Source: Adapted from Mennis et al. (2000).

Based on the purpose of the analysis, different points of view can be taken, one example would be considering distinct conceptual entities that are related to time (objects) vs. the observation of a continuous phenomenon, like temperature over time (values). There cannot be a single model that is ideal for all kinds of applications. So, in order to characterize time-oriented data we focus on data component i.e., the lower part fo the pyramid framework as depicted in the Figure-10.

1) Scale: quantitative vs qualitative
2) Frame of reference: abstract vs spatial
3) Kind of data: events vs states
4) Number of variables: univariate vs multivariate

A. Scale: quantitative vs qualitative

Quantitative variables are based on a metric(discrete or continuous) range that allows numeric comparisons. In contrast, the scale of qualitative variables includes an unordered or ordered set of data values. Data scale is fundamental characteristic which needs to considered to design appropriate visual representations.

B. Frame of reference: abstract vs spatial

Abstract data means a data model that does not include the where aspect with regard to the pyramid framework, i.e., abstract data are not connected per se to some spatial location [11]. In contrast to this, spatial data contain an inherent spatial layout, i.e., the underlying data model includes the where aspect. The distinction between abstract and spatial data reflects the way the time-oriented data should be visualized [7]. For spatial data, the inherent spatial information can be exploited to find a suitable mapping of data to screen. The when aspect has to be incorporated into that mapping, where it is not always easy to achieve an emphasis on the time domain.

C. Kind of data: events vs states

Events can be seen as markers of state changes, example departure of plane. States can be characterized as phases of continuity between events, example plane is in the air. States and events are two sides of the same coin. However both should be specifically mentioned before visualizing.

D. Number of variables: univariate vs multivariate

This criteria clearly shows the number of time-dependent variables. Representing data where each time primitive is associated with only one single data value is univariate and if it associated with multiple data values it is multivariate.

Summarizing, in this section we structured and specified the characteristics of time and time-oriented data. So far we have modeled time based on complexity into different categories and then we characterized data mainly focusing on temporal relations of data variables. We can see that how data variables are associated with time primitives using distinction of internal and external time. All these aspects need to be considered when visualizing and analyzing data variables over time. In the following section we will discuss about the visualization problems and some design examples.

IV. Aspects of Visualization

Data related to time comes in many variants like meteorological data, financial data, census data, medical data, simulation data, news articles, photo collections or projection plans are just a few examples which contain temporal information. All these types of data contain *time* as a constraint so these data should be representable with one and the same visualization approach but in practice the data exhibits different characteristics, hence each of the above given example requires a dedicated visualization. Example visualization techniques pertaining to each data sample will be shown in the following section but for now in this section we try to characterize the visualization problem raised. A systematic view of visualization of time-oriented data is required [12].

A. Characterization of the Visualization Problem

To organize our systematic view we need a structure, the structure is geared to answer three practical questions that are sufficiently specific for researchers and at the same time easy to understand for practitioners [12], [7]:

1) What is presented? - *Time & data*
2) Why is it presented? - *User tasks*
3) How is it presented? - *Visual representation*

The first question addresses the structure of time and the data that have been collected over time. The motivation for generating the visualization is addressed by the second question and finally how the data is represented is covered by the third question.

1) What is presented? - Time & data: Temporal dimension is crucial aspect that any visualization approach for representing time and time-oriented data has to be considered. It is impossible to design effective visual representation without the knowledge of the characteristics of given data and time domain. The first characteristics of time and data have already been fleshed and discussed in detail out in section II and III.

2) Why is it presented? User tasks: User tasks are usually given at a lower level in the form of informal verbal lists only. An accepted low-level task description specifically addressing the temporal domain has been introduced by McEachren [13]. The tasks are defined by a set of important questions that users might seek to answer with the help of visual representations [7]:

- Existence of data element
- Temporal location
- Time interval
- Temporal pattern
- Rate of change
- Sequence
- Synchronization

The list covers two basic cases. First, having at hand one or more data values, the user is searching for time primitives that exhibit these values, and second, having at hand one or more time primitives, the user seeks to discern the data values associated with them [7]. This reflects the well-established distinction between identification (i.e., looking for data values) and localization (i.e., looking for when and where in time and space) [13].

3) How is it presented? - Visual representation: This question sums up the answers and results of previous questions in a visual manner. The next section i.e., section V will show a large variety of visual approaches which provide very different answers to this question. The fundamental criteria of this question can be divided into:

- Mapping of time
- Dimensionality of the presentation space

Any data variable that needs to be visualized, the dimension of time has to pass the mapping step of the visualization pipeline [7]. Mapping can be done in two ways: mapping of time to space and the mapping of time to time. Mapping of time to space is where data and time are represented in a single coherent visual representation which are *static* while mapping of time to time is based on the dependency of the data which is *dynamic*. Both static and dynamic representations will be further discussed in the section V below. Dimensionality of presentation space is divided into 2D and 3D representation of data which will also be covered in section V.

In summary to solve the visualization problem primarily requires answering the three questions: (1) What is visualized? (2) Why is it Visualized? and (3) How is it Visualized? The first two questions if answered will answer the third question by implication. A variety of techniques for handling and accounting these key characteristics will be seen in the next section.

V. Survey of Visualization Techniques

In this section we survey some existing visualization techniques for time and time-oriented data. The complexity of the visualization problem, which arises from multitude of aspects having an impact on the visual representation which already suggests that there must be variety of techniques to deal with the multitude of characteristics of data behavior. We couldn't do a exhaustive survey as visualization of time-oriented data is a hot research area which is constantly yielding new techniques. But covered most of the important ones which were application dependent and satisfied majority of the problems posed and helped user to analyze data efficiently and very recent one were chosen.

To keep the categorization at a manageable level we do not use the complete scale classification introduced but focus on three key criteria as follows:

- data
  - frame of reference - abstract vs spatial
  - variables - univariate vs multivariate
- time
  - arrangement - linear vs cyclic
  - time primitives - instant vs interval
- vis
  - mapping - static vs dynamic
  - dimensionality - 2D vs 3D

wherever possible further distinction will be given. However, this is not possible for every technique discussed as more general and flexible visualization approaches are designed by majority of the researchers as their core work. The techniques list follows the top most criteria to bottom since if we started to sort techniques based on name or year published would scatter them all over the place. Example we start with data criteria with abstract then more to spatial and then select univariate then move to multivariate and so on.

A. SparkClouds

This is a technique where tag clouds visualize a set of keywords weighted by their importance. By varying font size, color, or other visual variables important keywords are emphasized over less-important keywords. Classic tag clouds, however, are incapable of representing the evolution of keywords. Lee et al. [14] integrate sparklines into tag clouds in order to visualize temporal trends in the development of keywords. The idea being visually combine a keyword (a tag) and its temporal evolution. The importance of keyword is encoded with font size used to render the text, where size can correspond either to the overall importance of the keyword for entire time series or to the importance at a particular point in time. A sparkline is attached to the keyword where it represents the keyword's trend. A color gradient is shown in the background of each keyword-sparkline pair to make this design perceivable as a visual unit. SparkClouds has gone through a user studies and was confirmed that it is useful and have advantages over alternative standard methods for visualizing text and temporal information. Figure - 11 shows the visual representation of the technique.

The criterion this technique falls under is:

- data
  - frame of reference - abstract
  - variables - univariate
- time
  - arrangement - linear
  - time primitives - instant
- vis
  - mapping - static
  - dimensionality - 2D

B. Horizon Graph

This representation of data by Horizon Graphs [15] as a visualization technique for comparing large number of time-dependent variables. Horizon graphs are based on the two tone pseudo coloring technique by [16]. The left part of the Figure-12 shows the construction of horizon graphs. Starting from a common line plot, the value range is divided into equally sized bands that are discriminated by increasing color intensity towards the maximum and minimum values while using different hues for positive and negative values [15]. Then, negative values are mirrored horizontally at the zero line. Finally, the bands are layered on top of each other. This

Fig. 11: Display of the 25 most important keywords in a series of twelve measurements. The bigger the font size is, the more important is a keyword. Keywords that are not among the current top 25 important keywords but have been among them at an earlier point in time are attenuated by using dimmed color and smaller font size [14].

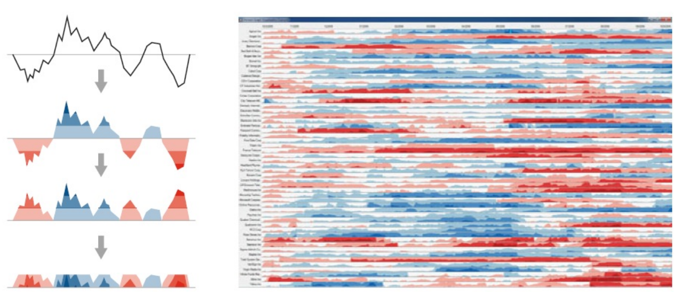

Fig. 12: The construction of a horizon graph from a line chart is illustrated on the left. Because horizon graphs require only little screen space they are very useful for comparing multiple time-dependent variables as shown to the right for stock market data [17].

way, less vertical space is used, which means data density is increased while the resolution is preserved.

The criterion this visual representation falls under is:

- data
  - frame of reference - abstract
  - variables - univariate
- time
  - arrangement - linear
  - time primitives - instant
- vis
  - mapping - static
  - dimensionality - 2D

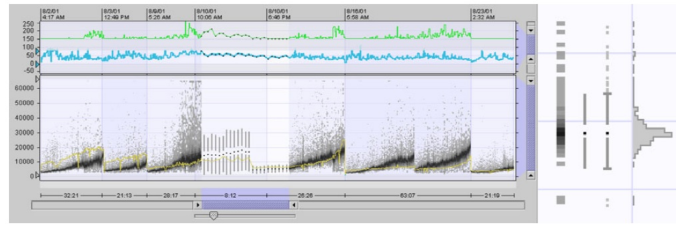

Fig. 13: The main panel shows the raw data (drug discovery data) and the top panel depicts derived statistical values. Depending on the available screen space, four different levels of visual abstraction are used: density distributions, thin box plots, box plots plus outliers, and bar histograms (as illustrated to the right) [15].

C. *TrendDisplay*

TrendDisplay is a technique by Brodbeck and Girardin [17] which allows the analysis of trends in larger time-series. This technique was used for a real world application for drug discovery process and in quality control. The display consists of two panels, main panel on the bottom showing measured (raw) data and top panel depicts derived statistical values shown in Figure-13. Density distributions, thin box plots, box plots plus outliers and bar histograms are the four different details used in this method to keep up with large number of time points. Brushing & linking as well as smooth transitions complete the highly interactive interface.

The Criteria being:

- data
  - frame of reference - abstract
  - variables - univariate
- time
  - arrangement - linear
  - time primitives - instant
- vis
  - mapping - static
  - dimensionality - 2D

D. *TimeNets*

TimeNets approach by Kim et al. [18] aims to visualize genealogical data with bth family structures of interest and also temporal relationships. It represents persons as individual bands that extend horizontally along a time axis from left to right. Each band shows a label of the persons name and different colors are used to encode sex: red is reserved for females, and males are shown in blue [18]. Marriage of persons is visualized by converging the corresponding bands, while divorce is indicated by diverging bands.When a child is born, a new band is added to the display. A so-called drop line connects the band of the child to the parents bands to convey the parent-child relationship [18], [7]. This technique has a DOI (degree of interest) algorithm which helps users to focus on relevant parts of the data. Users are also able to select multiple persons to focus on an each change of the focus,

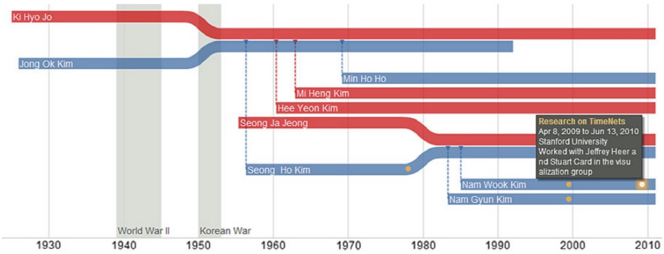

Fig. 14: TimeNets visualize temporal and structural aspects of genealogical data. Bands that extend along the horizontal time axis visualize individuals. Marriage and divorce are indicated by converging and diverging bands, respectively. Children are connected to their parents via drop lines. Labels are shown for the person's names as well as for historical and personal events [18].

the visualization shows a smooth transition of the display to keep users oriented. Figure-14 shows TimeNets visualizing temporal and structural aspects of genealogical data.

The technique criteria fall as follows:

- data
  - frame of reference - abstract
  - variables - univariate
- time
  - arrangement - linear
  - time primitives - interval
- vis
  - mapping - static
  - dimensionality - 2D

E. SpiraClock

SpiraClock was introduced by Dragicevic and Huot [19] using clock as a metaphor. The visual representation resembles the clock having two hands indicating hour and minute and its interior shows a spiral that extends from clocks circumference towards its center. This is shown in the Figure-15 where time intervals are represented as thick segments along the shape and these segments show when intervals start and end too. Users can also see if certain appointments are in conflict if they overlap. As time advances, the spiral is constantly updated and future intervals gradually move outward until they are current [19]. Past intervals gradually fade out, in this sense, the SpiraClock enhances classic clocks with a preview of the near future and a brief view to the past. Users can also drag the clock handles to visit different points in time, and intervals of interest can be highlighted and corresponding textual annotations can be displayed.

The criteria it falls under is:

- data
  - frame of reference - abstract
  - variables - univariate
- time
  - arrangement - cyclic

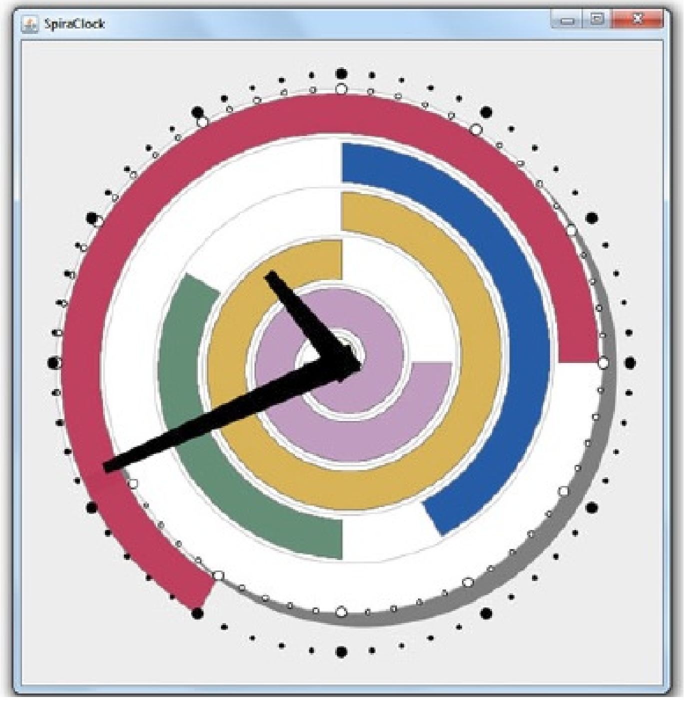

Fig. 15: The SpiraClock builds upon the clock display. The minute hand currently points to a meeting that has already started. Future appointments are aligned along a spiral on the clock face [19].

  - time primitives - interval
- vis
  - mapping - dynamic
  - dimensionality - 2D

F. Spiral Display

This is an interactive display developed by Carlis and Konstan [20] using Archimedean spirals to represent the time domain. Data values at different time points are visualized as filled circular elements whose area is proportional to the data value shown in Figure-16(top left). When it comes to interval-based data, filled bars are aligned with the spiral shape to indicate start and end of intervals [20] in Figure-16(center). If multivariate data are given at time points, the spiral is tilted and data values are visualized as differently colored spikes, where spike color indicates variable affiliation and spike height encodes the corresponding data value shown in Figure-16(bottom-left) [20]. Users are even use the z-axis to separate the display of multiple variables Figure-16(right). The cycle lengths of spirals can be adjusted interactively and can also be animated automatically for discovering periodic patterns.

This technique falls under:

- data
  - frame of reference - abstract
  - variables - univariate, multivariate

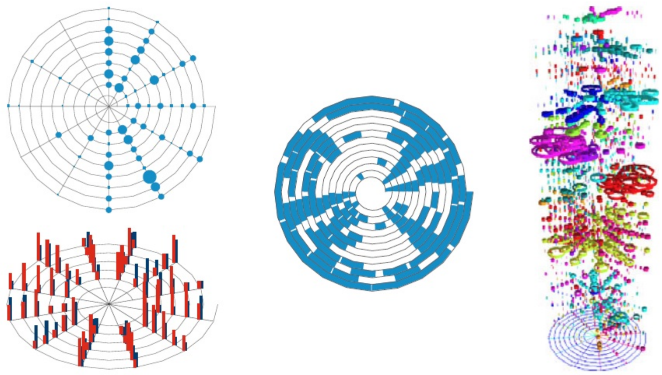

Fig. 16: Data can be visualized along a spiral in different ways: by the area of circular elements (top-left), by the sizes of multiple spikes (bottom-left), as bars marking start and end of intervals (center), or by the volume of hollow cans aligned at different layers along the vertical axis (right) [20].

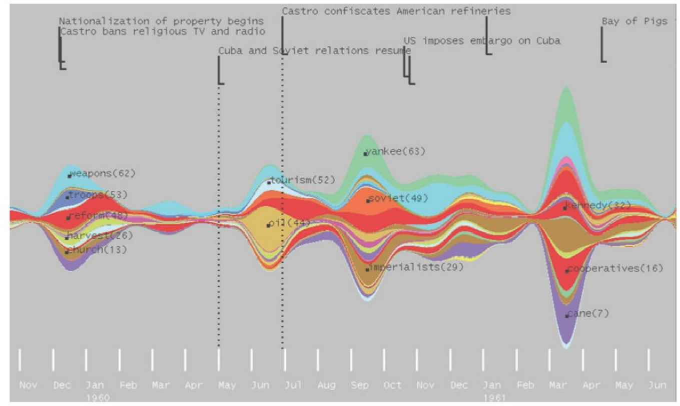

Fig. 17: The ThemeRiver representation uses the metaphor of a river that flows through time. Colored currents within the river reflect thematic changes in a document collection, where the width of a current represents the relevance of its associated theme. [21].

- time
  - arrangement - cyclic
  - time primitives - instant, interval
- vis
  - mapping - static
  - dimensionality - 2D, 3D

G. ThemeRiver

ThemeRiver technique developed by Havre et al. [21] representing changes of new topics in the media. Each topic displayed as a colored current whose width varies continuously as it flows through time. The main point of ThemeRiver is that it provides an overview of the topics that were important at certain points in time. Hence, the main focus is directed towards establishing a picture of an easy to follow evolution over time using interpolation and approximation [21]. It was mainly invented to visualize thematic changes in document collections, it is also suited to represent other multivariate, quantitative data. Because perception of data differs depending on where in the river individual variables are shown, it is important to provide interaction techniques to allow users to rearrange the horizontal position of variables [21]. It is visually represented in Figure-17.

This technique falls under:

- data
  - frame of reference - abstract
  - variables - multivariate
- time
  - arrangement - linear
  - time primitives - instant
- vis
  - mapping - static
  - dimensionality - 2D

H. TimeWheel

TimeWheel was developed by Tominski [22] where it visualizes multiple time-dependent variables. It has single time-axes and multiple data axes for data variables. The time axis is in the center of display and data axes are associated with individual colors, arranged circularly around the time-axis. TO visualize data each data axes are emanated from time axis to establish a visual connection between points in time and associated data values. The visual patterns formed by these axes allow users to identify positive or negative correlation with the time-axis, trends and outliers. Users can rotate the TimeWheel to bring data axes of interest into focus.Interactive exploration, including navigation in time, is supported through different types of interactive axes [22]. A visual representation of technique is shown in Figure-18.

The criteria of this technique is:

- data
  - frame of reference - abstract
  - variables - multivariate
- time
  - arrangement - linear
  - time primitives - instant
- vis
  - mapping - static
  - dimensionality - 2D

I. Kiviat Tube

The Kiviat tube by Tominski [23] visualizes multiple time-dependent variables. The construction of Kiviat tube is done just stacking multiple Kiviat graphs [24] along a shared time axis. Each Kiviat graph in the tube represent multiple variables for a specific point of time. On a 3D surface all these graphs are combined to represent the dataset as a whole. The peaks or valleys in data can be easily identified by the users over time.

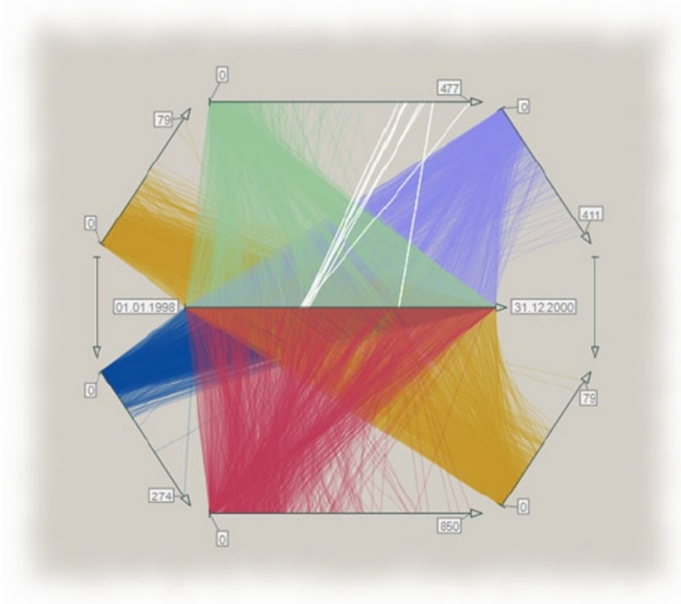

Fig. 18: The TimeWheels central axis represents time. The axes in the periphery represent time-dependent variables; here we see the number of cases for eight diagnoses. Days with particularly high numbers of influenza cases are highlighted [22].

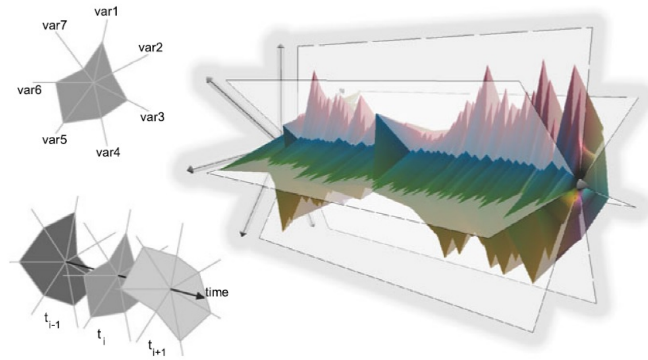

Fig. 19: Construction of a three-dimensional Kiviat tube representing seven time-dependent variables. Peaks and valleys indicate ups and downs in the evolution of the data over time.Wings assist in associating features of the Kiviat tube to particular variables in the data [23].

Additional semitransparent wings assist in relating identified patterns to particular variables. Common interaction methods can be used for zooming and rotation around arbitrary axes. Rotation specifically around the time axis enables users to quickly access variables on all sides of the Kiviat tube. Interactive axes allow users to navigate back and forth in time to visit different intervals of a possibly large time-series [23]. A visual representation of this is shown in Figure-19.

The criteria it falls under is:

- data
  - frame of reference - abstract
  - variables - multivariate
- time
  - arrangement - linear
  - time primitives - instant
- vis
  - mapping - static
  - dimensionality - 3D

J. Helix Icons

Helix Icons by Tominski et al. [25] are useful for emphasizing the cyclic character of spatio-temporal data. This visualization is derived from model presented in [26] in the form of a space-time cube, which maps the spatial context to the x-axis and the y-axis, and the dimension of time to the z-axis of a virtual 3D cube. The helix ribbon is constructed to unroll the time domain along the z-axis. each segment of the helix ribbon visualizes a specific instant in time by means of color-coding. Multiple time-dependent variables can be visualized by subdividing the helix ribbon into narrower sub-ribbons, each of which represents a different variable. Using unique hues for each sub-ribbon helps the user distinguish variables [25]. The inherent flaws in this technique in 3D representation i.e., information displayed on helix back faces or inter-icon occlusion are dealt with by offering 3D navigation through the space-time cube and rotation of helix icons. A visual representation is shown in Figure-20.

- data
  - frame of reference - spatial
  - variables - multivariate
- time
  - arrangement - cyclic
  - time primitives - instant, interval
- vis
  - mapping - static
  - dimensionality - 3D

In summary in this section we have reviewed 10 existing visualization techniques for time and time-oriented data and categorized them accordingly as per our distinction mentioned in previous sections yet there might be some imbalances and further classification to be done for finer results yet it is sufficient enough to help researchers and beginner practitioners to get an idea which technique would be more suitable for them.

VI. Observation

From the survey we can draw out some conclusions based on our categorization. Some possible facts worth noting are *Data - Variables*, the number of techniques for univariate and multivariate data are almost balanced. while many classic techniques consider univariate data while modern approaches take on the challenge of dealing with multiple variables.

Time - Arrangement, Most techniques in survey are supporting linear time; the approaches with cyclic time are

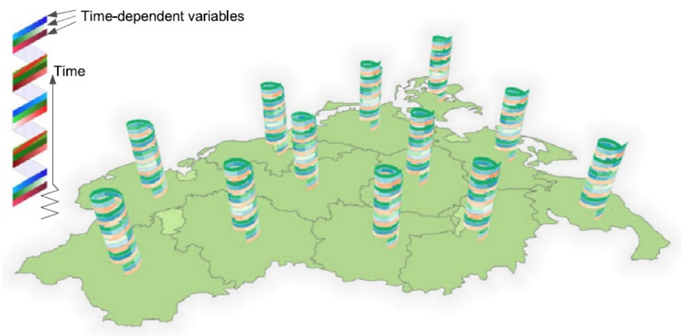

Fig. 20: Helix icons use color-coding to visualize multivariate spatio-temporal data along helix ribbons, which emphasize the datas cyclic temporal character. The spatial aspect of the data is illustrated by embedding helix icons in a space-time cube [25].

significantly outnumbered. Reasons might be users are only interested in trends evolving from past, to present, to future rather than in finding cycles in the data. But effective data analysis requires cyclic representations.

Vis - Mapping, Static pages are better suited for showing static techniques. The survey mostly contained static approaches which isn't a fair one. Dynamic animation is equally important and often the first solution offered when time-oriented data have to be visualized.

Vis - Dimensionality, 2D visual representations are often preferred over 3D ones as they are more abstract and thus easier to understand. However modern designers and advancement in visualization techniques try to implement 3D and users to navigate in 3D. Helix Icons [25] is one such technique attempting to do it. 3D is more useful when spatial references have to be visualized.

One more observation we can conclude is that most approaches address the model of *ordered* time domain while only few of them explicitly consider visualization of *branching* alternate strings of time. Therefore, branching time and time with multiple perspectives deserve more research attention in future [7]. We can discuss the closing conclusions, application issues and areas of further research in the next section.

VII. Conclusion

In this paper we have modeled time based on various design aspects like scale, scope, arrangement and viewpoint which are further divided into sub categories for further classification. we have characterized the data component of the time-oriented data as data is distinct and behaves in multitude of ways so that there cannot be a single mode that is ideal for all kinds of applications. This characterization was done based on scale, frame, kind of data and number of variables which is then further divided into their own sub categories.

We then characterized the visualization problem using some simple practical questions like *What is presented?*, *Why is it presented?* and *How is it presented?* and these are further explored with more simple questions trying to answer them and in turn getting all the answers. Finally we surveyed variety of techniques, categorized them based on the distinctions discussed and drew out some interesting conclusions from the survey and we can conclude that the survey has given us enough evidence that time is an important dimension that deserves special treatment in visual, interactive, and analytical methods. However, tools or systems that provide the broad functionality demanded specifically for time-oriented data are not available. Moreover, there are several open issues to be addressed in the future.

VIII. Future Research Challenges

Designing appropriate visual representations and tools for time-oriented data as well as making them applicable in real world problem solving scenarios requires further scientific investigation. In the field of software engineering it is generally acknowledged that the first step in developing tools and user interfaces should be a sound analysis of the given problem domain [27]. The same applies for designing visual representations. Choosing adequate visualizations is one point, but providing a set of suitable techniques that cover all the different aspects of time is another concern. Although a large diversity of powerful visualization techniques for time-oriented data have been developed, most of them support only certain parts of the introduced time and data categorization [7].

Visualization techniques show data themselves rather than information on data quality or data provenance, taking uncertainties into account will significantly improve the expressiveness of the visual representations and also improve the user's confidence in the findings. Currently, visualization methods communicating both time and data uncertainties are not available. Therefore, new visualization strategies have to be developed. The concern of representing the quality of data has also been acknowledged as an open research topic in visualization research in general [28].

Many techniques were facing scalability issues when it comes to large amount of data. This calls for particular characteristics and functionalities of data management as well as of the visual and algorithmic design [28]. With regard to time-oriented data, we need to be able to deal with both very long time-series containing vast amounts of time primitives and large numbers of time-dependent variables in parallel [7]. These are some of the many research challenges to be worked on and improved in the future. Visual Analytics is still a budding research area with lots of improvement to be done.

Acknowledgment

The author would like to thank Professor T. McInerney for his lectures and guidance throughout the semester.

References

[1] D. A. Keim, F. Mansmann, J. Schneidewind, and H. Ziegler, "Challenges in visual data analysis," in *10th International Conference on Information Visualisation, IV 2006, 5-7 July 2006, London, UK*, pp. 9–16, 2006.

[2] G. Sun, Y. Wu, R. Liang, and S. Liu, "A survey of visual analytics techniques and applications: State-of-the-art research and future challenges," *J. Comput. Sci. Technol.*, vol. 28, no. 5, pp. 852–867, 2013.

[3] D. A. Keim, J. Kohlhammer, G. P. Ellis, and F. Mansmann, *Mastering the Information Age - Solving Problems with Visual Analytics*. Eurographics Association, 2010.

[4] J. J. Thomas and K. A. Cook, "A visual analytics agenda," *IEEE Comput. Graph. Appl.*, vol. 26, pp. 10–13, Jan. 2006.

[5] J. Thomas, *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society Press, 2005.

[6] B. Shneiderman, "The eyes have it: A task by data type taxonomy for information visualizations," in *Proceedings of the 1996 IEEE Symposium on Visual Languages, Boulder, Colorado, USA, September 3-6, 1996*, pp. 336–343, 1996.

[7] W. Aigner, S. Miksch, H. Schumann, and C. Tominski, *Visualization of Time-Oriented Data*. Human-Computer Interaction Series, Springer, 2011.

[8] A. Frank, "Spatial and temporal reasoning in geographic information systems, chapter different types of times in gis, 40–62," 1998.

[9] I. A. Goralwalla, M. T. Özsu, and D. Szafron, "An object-oriented framework for temporal data models," in *Temporal Databases, Dagstuhl*, pp. 1–35, 1997.

[10] M. Tory and T. Möller, "Rethinking visualization: A high-level taxonomy," in *10th IEEE Symposium on Information Visualization (InfoVis 2004), 10-12 October 2004, Austin, TX, USA* (M. O. Ward and T. Munzner, eds.), pp. 151–158, IEEE Computer Society, 2004.

[11] J. L. Mennis, D. J. Peuquet, and L. Qian, "A conceptual framework for incorporating cognitive principles into geographical database representation," *International Journal of Geographical Information Science*, vol. 14, no. 6, pp. 501–520, 2000.

[12] W. Aigner, S. Miksch, W. Müller, H. Schumann, and C. Tominski, "Visualizing time-oriented dataa systematic view," *Computers & Graphics*, vol. 31, no. 3, pp. 401–409, 2007.

[13] A. M. MacEachren, *How maps work: representation, visualization, and design*. Guilford Press, 1995.

[14] B. Lee, N. H. Riche, A. K. Karlson, and S. Carpendale, "Sparkclouds: Visualizing trends in tag clouds," *IEEE transactions on visualization and computer graphics*, vol. 16, no. 6, pp. 1182–1189, 2010.

[15] H. Reijner *et al.*, "The development of the horizon graph," 2008.

[16] T. Saito, H. N. Miyamura, M. Yamamoto, H. Saito, Y. Hoshiya, and T. Kaseda, "Two-tone pseudo coloring: Compact visualization for one-dimensional data," in *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005.*, pp. 173–180, IEEE, 2005.

[17] D. Brodbeck and L. Girardin, "Interactive poster: Trend analysis in large timeseries of high-throughput screening data using a distortion-oriented lens with semantic zooming," in *Poster Compendium of IEEE Symposium on Information Visualization (InfoVis03)*, pp. 74–75, Citeseer, 2003.

[18] N. W. Kim, S. K. Card, and J. Heer, "Tracing genealogical data with timenets," in *Proceedings of the International Conference on Advanced Visual Interfaces*, pp. 241–248, ACM, 2010.

[19] P. Dragicevic and S. Huot, "Spiraclock: a continuous and non-intrusive display for upcoming events," in *CHI'02 extended abstracts on Human factors in computing systems*, pp. 604–605, ACM, 2002.

[20] J. V. Carlis and J. A. Konstan, "Interactive visualization of serial periodic data," in *Proceedings of the 11th annual ACM symposium on User interface software and technology*, pp. 29–38, ACM, 1998.

[21] S. Havre, B. Hetzler, and L. Nowell, "Themeriver: Visualizing theme changes over time," in *Information Visualization, 2000. InfoVis 2000. IEEE Symposium on*, pp. 115–123, IEEE, 2000.

[22] C. Tominski, J. Abello, and H. Schumann, "Axes-based visualizations with radial layouts," in *Proceedings of the 2004 ACM symposium on Applied computing*, pp. 1242–1247, ACM, 2004.

[23] C. Tominski, J. Abello, and H. Schumann, "Interactive poster: 3d axes-based visualizations for time series data. 2005," *Poster IEEE InfoVis*, vol. 5.

[24] K. W. Kolence and P. J. Kiviat, "Software unit profiles & kiviat figures," *ACM SIGMETRICS performance evaluation review*, vol. 2, no. 3, pp. 2–12, 1973.

[25] C. Tominski, P. Schulze-Wollgast, and H. Schumann, "3d information visualization for time dependent data on maps," in *Ninth International Conference on Information Visualisation (IV'05)*, pp. 175–181, IEEE, 2005.

[26] M.-J. Kraak, "The space-time cube revisited from a geovisualization perspective," in *Proc. 21st International Cartographic Conference*, pp. 1988–1996, 2003.

[27] C. Courage and K. Baxter, *Understanding your users: a practical guide to user requirements: methods, tools, and techniques*. Gulf Professional Publishing, 2005.

[28] D. A. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann, *Mastering the information age-solving problems with visual analytics*. Florian Mansmann, 2010.