

Time Histograms with Select and Zoom for Creating Visual Representation of Data Set for Managers

Snezana Savoska, Suzana Loskovska, Vlatko Blazeski

Abstract - Histograms are very useful tools for visualization and the most frequently used analysis techniques for managers because they show distribution of value over a data dimension. Although 2D and 3D standard histograms show the distribution of value over data dimensions, they do not consider the time as additional dimension. In this paper, we present Time Histograms with Select&Zoom technique which is an extended technique of standard histograms with possibility of selecting time step for zooming with slider in second – zoom window.

Keywords – Time Histograms, Select and Zoom, Data Visualization for Management

I. INTRODUCTION

Histograms are very useful data visualization techniques. Standard time histograms provide easy way to see data without interaction opportunity. For this reasons, the goals of our paper is to present time histograms with select and zoom possibilities and scroll bar as a tool for interaction. To achieve the goal, time dimension must be defined at the beginning. That means the time steps which will consist of the unit of selection and zooming in the process of visualization, must be specified as time slabs. For these reasons, the data in data storages has to be surveyed and analyzed for making decision which data and technique are the most appropriate for data visualization. Unfortunately, standard histograms provide just a few capabilities for customer interaction. Also, additional time dimension is seldom used.

In this paper, we introduce extended time histogram with interaction and select and zoom possibilities. This interaction and possibility to select and zoom the time zone (time slab) is enabled with slider which is located under the general static window. When the slider is moved, the time slab is changed and the time histogram is shown in the second window with zoomed data from selected time slab, located under the global time histogram and scroll bar. Also, there is some additional capability to show couple of data properties as sell region, sell sub-region, product group, etc. [3]

Data set used in this paper comes from Transactional Information Systems of the production and Trade Company for the last twelve years. The databases are extracted, loaded and transformed for the requirements of the data storage. There are 79760 data records of all transactions in the given company. The database contains all types of documented data. For us, the sell documents for the last twelve years are the most interesting at present. To make effective visual analysis, it is very useful to create Data Warehouse, extract, load and transform data from transactional databases. Also, adjustments of data and Meta data were made.

The adjustment steps are the time step specifications (time slab) for zooming and visual data rendering. These time slabs

were defined with top-down analysis in the company. The managers use visual time reports in different time steps: month, quarter, year, decade... Although the tiny time interval can be analyzed, they are not interesting for managers. With exception for time dimension, all data in data storage has couple of dimensions – attributes which can be analyzed by managers and analytical staff with 2D and 3D analysis.

The entire time space has been divided in discrete time steps (time slabs) [1], which will be the part of set for selection. The selected time slab data is shown in the additional window space, located under the time scroll bar. The customer has to select the desired time slab interactively to see the selected data in the additional zoom window. This way, there is high class display usage given, with global static view of all data in the storage (separated in different dimensional bars with different colors), scroll bar form for interactive usage and second display with data from selected time slab [8].

II. DATA VISUALIZATION GOALS

Because the standard time histograms do not provide understandable view of desired dimensions in time (attributes as regions, sub-regions, product groups), the goal of this paper is to provide data analysis with possibilities to use standard histograms as solid base and their transformation in interactive data visualization tools, which will allow time step zooming. For this reasons, our goals in designing time histograms are:

- To provide an clear and obvious way for displaying dependant time histogram for users
- To provide opportunities for displaying large scale time information in time histograms with Select&Zoom of time slabs capability.
- To save the global view of data in the time dimension and to achieve the readability of time histograms in selected time slab [1]. When the user sees the global view of all time data set in display, it is possible to select time slab and show the data in the additional histogram (window in display in the some time). In this manner, the users have the impression from the global Time Histogram, for data in the selected time slab.
- To support the capabilities of zooming with interaction with slither positioned bellow the global Time Histogram window. This interaction will provide selecting of desired dimension and time zone for analysis in the same time. The region and sub-region analysis in time slab is selected with tabs from the second – zoom window.

III. DATA SET AND RELATIONAL WORK

The data set used in this paper comes from transactional databases of the company, loaded in data storage [2]. The procedure for extracting, transformation and data loading in the data storage (Data Warehouse) depends on management needs as well as the data in transactional information system. After this procedure has finished, the data in storage has to be prepared for effective visualization. For this reason, the time slabs are defined in one month. The total numbers of time slabs are 144. Data set can be analyzed in many dimensions (six dimensions – attributes: region, sub region, town, product group, product type and time) and three measures. For different data displaying, the used Time Histogram has to answer to the managers questions.

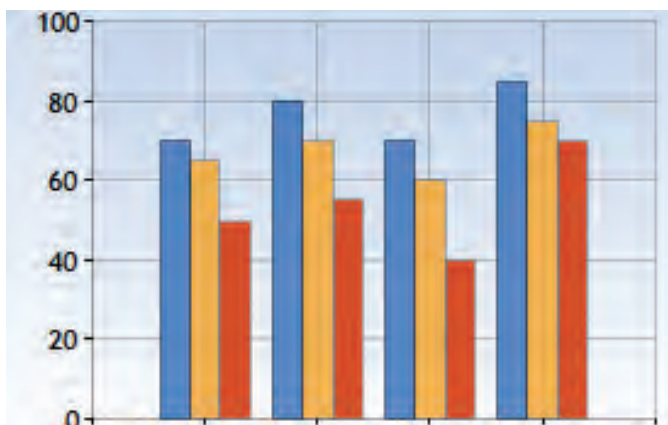


Fig. 1. Time histogram for product distribution in value (Denars) in regions

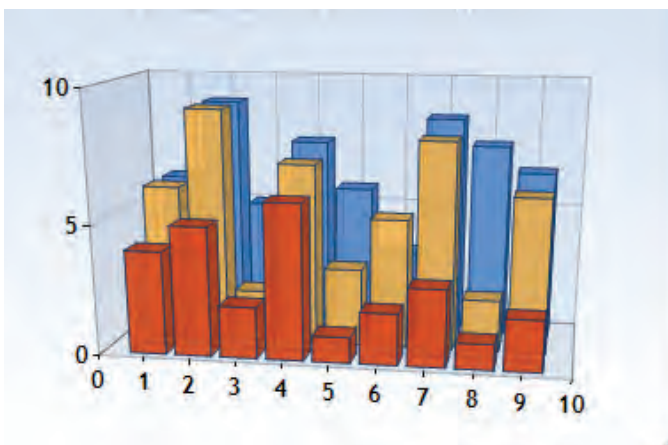


Fig. 2 3D Time histograms for products distribution (Millions) in time step

When different dimensions (attributes) are analyzed, various histogram techniques may be used. If the regions distribution analyses are the focus of manager interest, the parallel histogram is the most suitable technique. Also, for product type distribution, the time histograms presented in different time steps, with different colors and opacity are more appropriate (Fig.1). History bars have good user feedback when used in process visualization. Also, they provide some inspection of trends when they are used in the marketing data set. When monthly sell data in years has to be compared, 3D histograms are very convenient. The concept of the 3D time

histograms is to draw the histogram for each time step as arrows of cuboids rather than rectangles, and arrange them one behind the other along the time dimension (Fig 2).

In this case, the user can always see the simple global time histogram and in the same time, selected time slab with desirable data and information for managers.

IV. TIME HISTOGRAMS WITH SELECT AND ZOOM

One of the goals in the time histograms was to retain the simplicity and easy understanding of the standard histogram as much as possible. When we discussed 2D histograms, the use of colors for the best visualization process is desirable. But, when we discuss about 3D histograms, colors are used, as well as the other visualization primitive attributes, such as shadow, texture and lighting. In this manner, some additional attributes are described. When the 3D histograms are used, it is more conventional to pursuit 3D cuboids if there are Select&Zoom possibilities. But, although 3D histograms can be easily navigated with rotation possibility and zooming, this type of view has the restriction in two axis problems. The augmentation and reduction, left hand and right hand rotation and zooming of the rows in 3D cuboids are the most useful capabilities. Also, the wrap properties and global high control of the bars in histogram can be selected as a property. But, the goals of 2D and 3D histograms are not to show the exact values of the attributes, but to show the trends. But, the ToolTip [8] is possibility to give the correct value of displayed measures – the value of the presented bar in the graph. The ToolTip property can be set, if the managers want to see the correct value of the dimensions. For this purpose, the interaction of the users with enabled selection of time slab and zooming of the selected data in the additional window is the best solution for data analysis.

Some of the interactions are always possible, depending on the desired view. When the slider is used for interaction, the time slab can be changed and data can be displayed in additional zoom window. The ToolTip property can also be set and the customer can see the value of the bar in the time histogram. The values of the bars present the summarized amount of money in each time slab – monthly money amount. In the first window, the total money amount is shown, and in the second – a zoom window, the amount of money for each region and sub-region are being shown. Some interaction techniques provide possibility to change the upper and lower limit of display. This technique with select and zoom to the selected time slab can be compared with the perspective wall distortion method, but in this distortion technique one region is enlarged and the other region is squashed. In our technique, the enlarged region can be seen in the current time in additional window of display.

Select & Zoom in time histogram is a software tool created for a specific use [5]. With selection of dimensions, the attribute of focus is shown in the global window with large scale of measuring, depending on the selected dimension and focus of the selected time slab. For this purpose, there are few possibilities to select what dimension to see in the second zoom window. To select what will be shown in the window the system provides the radio buttons or tabs pages. In this

The screenshot shows the Microsoft SQL Server Enterprise Manager interface. At the top, the title bar reads "SQL Server Enterprise Manager [Design Time 'Polars' in release on 'LOCAL']". The main window displays a query plan for a query. The query plan consists of four tables (T1, T2, T3, T4) and a summary table (SUMMARY). The tables are connected by lines representing joins. The summary table is connected to the other four tables. The query plan shows a Hash Join operation. Below the query plan, the SQL text is displayed. The SQL text is as follows:

```
SELECT  dbo.MG.MG_SUM(dbo.Recent1.kolona) AS Recent1, SUM(dbo.Recent1.kolona) AS Recent2, SUM(dbo.Recent1.kolona) AS Recent3, SUM(dbo.Recent1.kolona)
        AS Recent4
FROM    dbo.MG LEFT OUTER JOIN
        dbo.Recent ON (dbo.MG.MG = dbo.Recent1.MG LEFT OUTER JOIN
        dbo.Recent2 ON (dbo.MG.MG = dbo.Recent2.MG LEFT OUTER JOIN
        dbo.Recent3 ON (dbo.MG.MG = dbo.Recent3.MG LEFT OUTER JOIN
        dbo.Recent4 ON (dbo.MG.MG = dbo.Recent4.MG
GROUP BY  dbo.MG.MG
```

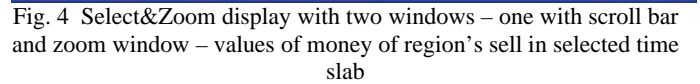


Fig. 5 Select & Zoom display – sell amount (in denars) in selected time slab sub-regions

When the data analysis is made in 3D histograms, different data dimensions can be analyzed, especially when the size of each bar in 3D histogram requires being analyzed. If the ToolTip property is set, the exact value of the bars shown in the ToolTip window. The slicing properties in 3D plane and their 2D plane can be very useful. But, the arrangement in 3D space is not arranged well enough in the stacked cuboids. The 2D histograms are much clearer (Fig.6).

The screenshot shows a Java Swing window titled "Farm1" with a standard Mac OS X title bar (red, yellow, and green buttons). The window contains a bar chart. The x-axis represents time, with labels for 3/2005, 1/2006, 6/2005, 1/2006, 4/2007, and 9/2007. The y-axis represents a numerical value, ranging from 0 to 8,000,000 in increments of 2,000,000. The chart displays a series of blue bars, which are identified in the legend as "Scene1". The bars show a fluctuating pattern of values over time, with a significant peak around 6,000,000 in mid-2005. A legend in the top right corner of the window area shows a blue square next to the text "Scene1".

The data presentations with these visualization tools have many advantages and disadvantages, and the fact is that they are not clear enough. The proposed technique of Select&Zoom in Time Histograms has many advantages for the mentioned purposes.

VI. CONCLUSION

Time Histograms with Select&Zoom property are extended technique of the well-knowns time-dependent data histograms. They are useful tools which provide effective data visualization, particularly for large and complex data sets. The presented technique provides a global view of data in the storage which can be an interactive “brush”. With the slider, located under this graph, and the additional zoom window with data for the selected time slab, separated depending of selected dimension can be seen. Also, this technique provides additional information as bar’s value (ToolTip property) and color for dimension values distinguishing.

Time Histograms with Select&Zoom were discussed in this paper with company data storage. In future work, the Time Histograms technique will be included in 3D, with possibility of row and columns selection for effective visual data rendering of knowledge-based data discovery algorithms.

ACKNOWLEDGMENT

We would like to thank the managers of the selected data company for their engagement and help in Top-down analysis, as well as the insightful and detailed remarks.

REFERENCES

- [1] Kosara R., Bendix F., Hauser H., “TimeHistograms for Large, Time-Dependant Data”, VrVis Research Center, Vienna,Austrija, IEEE TCVG Symposium on Visualization (2004)
- [2] Bonifati A.,Cattaneo F., Ceri S., Fugetts A., Paraboschi S., “Designing Data Marts for Data Warehouse, ACM No.4, 10.2001, Pages 452-483
- [3] Hansen C., Johnson C.,”The Visualization handbook”, USA-2005, ISBN 0-12-387582-X,
- [4] Ware C. Information Visualization,
- [5] Dundas tutorial and Dundas help, Dundas.com
- [6] Few S., Visualization Multidimensional Data through Time, Perspeptual Edge
- [7] Card S., Grudin J., Nielsen J., Information vialization – Percepction for design – Second edition, 2004
- [8] -, C# tutorial, Microsoft.com
- [9] Nguylu Q.V., Huang M.L., EncCon: an approach to constructing interactive visualization of large hierarchical data, Information visualization (2005) 4,1-21
- [10] Chen C., Yu Y., Empirical studies of information visualization: a meta-analysis, Interactive J.Human-Computer studies (2000) 53, 851-860
- [11] North C., Conklin N., Indukiri K., Saini V., Visualization schema and a web-based Architecture for cusom multiple-View visualization of multiple table databases, Center for Human-Computer interaction, 2001