

Investigating the Direct Manipulation of Ranking Tables for Time Navigation

Romain Vuillemot, Charles Perin

▶ **To cite this version:**

Romain Vuillemot, Charles Perin. Investigating the Direct Manipulation of Ranking Tables for Time Navigation. CHI '15 - 33rd Annual ACM Conference on Human Factors in Computing Systems , Apr 2015, Seoul, South Korea. pp.2703-2706, 10.1145/2702123.2702237 . hal-01138447

HAL Id: hal-01138447
https://inria.hal.science/hal-01138447

Submitted on 2 Apr 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Investigating the Direct Manipulation of Ranking Tables for Time Navigation

Romain Vuillemot
Harvard University
Cambridge, MA, USA
romain_vuillemot@hks.harvard.edu

Charles Perin
INRIA, Univ. Paris-Sud & CNRS-LIMSI
Bat. 660, Univ. Paris-Sud, 91405 Orsay, France
charles.perin@inria.fr



Figure 1. The direct manipulation of tables, our new technique, combines a standard table (far left and far right) as background, with a line chart that plots the table's values as foreground, during the table's update for overview and guidance. In this example, the user explores Japanese incomes between 1990 and 2010 by dragging the cell corresponding to the country towards the lower right along the slope of its line chart during this time period.

ABSTRACT

We introduce a novel time navigation technique to update ranking tables by direct manipulation. The technique allows users to drag a table's cells to change the time period, while a line chart overlays on top of the table to provide an overview of the changes. The line chart is also a visual hint to control the pace at which data are updated. We explore the design and usability of this technique for table variations in size, time spans and data variability. We report the results of a usability study, using academic citation rankings and economic complexity datasets, and discuss design implications coming with real-world scenarios such as missing data and affordance.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

Tables display data as text using rows and columns. Such representation is probably the most faithful to raw data and provides a direct access to their editing and sharing. Tables are implemented in popular tools such as spreadsheets (e. g., Microsoft Excel), offering a clutter-free and configurable grid-like layout with simple, yet powerful features such as rows ranking and columns re-organization for comparison [7]. Such interactions scale up to large and multidimensional datasets, relying on vertical or horizontal scrolling to reach non-visible areas. Tables are often updated over time, but tracking and controlling them remains difficult using standard widgets [6].

In this paper, we introduce a novel technique to better control updates in ranking tables while getting an overview of those updates (Figure 1). The technique is activated by a cell drag (a) which displays the line chart of the column's values over time (b). By dragging the mouse along the line chart, the table behind is updated accordingly, rows moving vertically to their updated position (c). When the drag is over, the line chart disappears and the table is updated. The line chart provides an overview of the changes over time, and its brevity does not alter any of the table's aforementioned assets. This technique originates from the well-established paradigm of direct manipulation [8], which, among many gains for the user, reduces his distance to the manipulated data. We hypothesize the following benefits of the technique for the user:

H1 *It preserves the user's focus* as the data in dragged cells match their representation on the line chart.
H2 *The layout consistency between the line chart and the table* enables comparison between rows.
H3 *The original table remains visible* as a background, providing direct access to raw data.
H4 *An overview is provided* by the line chart over multiple time points.

To evaluate the effectiveness of this technique, we identified a series of frequent temporal tasks related to tables from [1, 2, 6]. We focus on the following ones:

T1 *Retrieving values* for a given time point, which is a standard direct low-level benchmark task [1]. Several studies show that tables are more efficient than line charts for direct tasks (e. g., [3], see [5] for a review), and standard widgets are efficient for this kind of task in the time domain [6].

Published in CHI 2015, April 18 - 23 2015, Seoul, Republic of Korea
This is the authors version of the work.

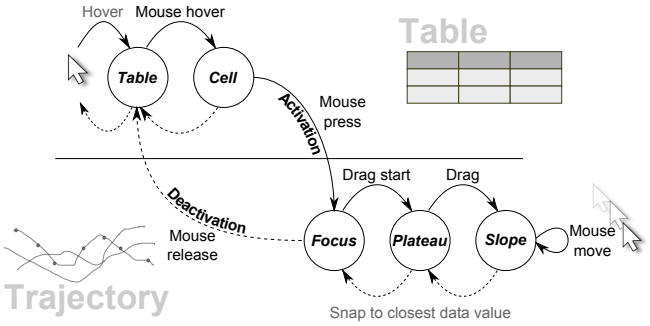


Figure 2. Graphical representation of the states and transitions from the table to the line chart, and vice versa.

T2 *Comparing values* over time, such as trend estimation and trend comparison [2]. Line charts have been found to be more efficient than tables for synoptic tasks, e. g., analyzing trends, comparing patterns, and comparing values [5, 6].

T3 *Performing tasks in the values domain,* such as estimating the time point or time interval for which a dimension of the table has a given value or trend [2]. Such tasks require alternative representations and interaction techniques to navigate in the values domain instead of the time domain [6].

This paper also explores several design variations for the technique that suit representative datasets of real scenarios with table variations in size, time spans and data variability.

DIRECT MANIPULATION OF RANKING TABLES

The novelty of the technique is to overlay two different representations from the same dataset on top of each others. First, as background a table that displays raw data as text. Then, as foreground, a line chart that plots the values for a column over several time points. The overlay is interactive, as the line chart is only visible once the user interacts with the table, i. e. drags a cell he is interested in. Figure 2 explains the series of interactions during which the line chart is visible and can be summarized as follows:

1. The user *starts* by dragging a cell on the table; the line chart appears.
2. The user *keeps* dragging; the line chart is horizontally translated while the table remains fixed. This ensures that the dragged cell always matches the line chart (**H1**) and the intersection remains the focus point (Figure 3).
3. The user *stops* dragging; the line chart disappears.

Although the line chart may have different sizes and configurations (e. g., scale, axis, interpolation, etc.), it will remain centered around the focus point. The line chart may eventually be decorated with additional visual marks (e. g., by showing other data points as circles and encoding those with other data). We will refer later to this particular mapping as *additional* dimension representation.

IMPLEMENTATION

We implemented the technique in JavaScript and SVG, using D3 for both the table and the line chart. SVG provided us with flexibility for custom transitions between the two representations. For instance, it lets us smoothly transform the rectangle of a cell into the circle of a focus point on the line chart. The prototype follows figure 2 in terms of interaction steps and we implemented several design variations illustrated in figure 3 for the line chart. The prototype is available online (http://romsson.github.io/dragit/), as well as the source code (https://github.com/romsson/dragit) with a permissive license.

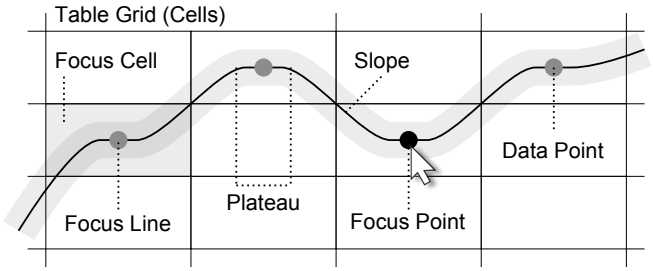


Figure 3. The line chart is the visual connection of data points, spatially ordered by time (from left to right) which match with the table's cells.

During the prototype's development, we evaluated the technique's usability with constant feedback from users. It informed us of the many design decisions to make, such as the type of scale, line interpolation and transitions between the two representations. Since the line chart design can be very different depending on the task or dataset, we provided the user with a way to change those parameters from pre-defined configurations.

CASE STUDIES

We now report results from an exploratory study we ran by interviewing two experts using our implementation. We succinctly presented them the technique, and then they spent between 35 minutes and 1 hour freely exploring a dataset popular in their research community we pre-loaded in our tool.

Academic Citations

Our first expert was a researcher in the field of information visualization. We presented him with a ranking table containing authors of articles published at the IEEE VIS conference between its creation in 1995 and 2013 (19 years). The table's columns include the number of papers, the number of citations, and the number of cited papers for 44 authors (rows). The expert was already familiar with the data from his professional activity and his use of an existing online tool (http://www.cc.gatech.edu/gvu/ii/citevis/).

The expert started by selecting three famous authors to study (Figure 4). He sorted the table according to the number of citations and dragged one of the three cells to analyze trends. While dragging the three authors, he realized that one of them is constantly well cited since the beginning of the conference (a). Indeed, as the table was visible as a background, he observed when authors enter or exit the top 10. He also used *additional* dimensions—in this context the data points' size encodes the number of publication by the current author at the current time step—to deduce that a "best paper mention" (red outline) highly increases the number of citations (b). But the experts added that *"to have a best paper, it seems that the author needs to be cited and well-known before."*.

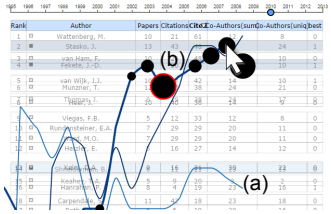


Figure 4. Comparing three authors' rankings according to their number of citations. Black circles encode the number of papers and the red outline indicates a best paper mention.

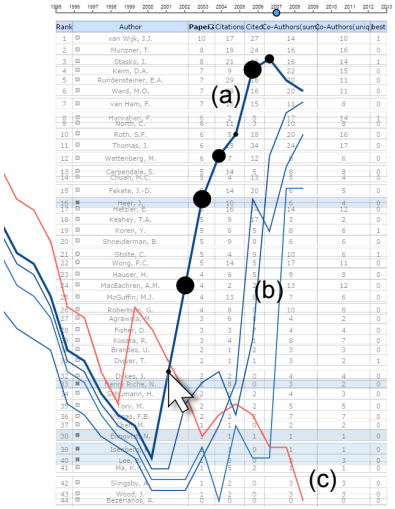


Figure 5. Comparing four authors' rankings according to their number of published papers.

His second selection focused on the younger generation of authors in the field (Figure 5). He commented that one of the 5 selected authors had an *"incredible slope"* starting in 2004 (a). Then he realized that the three selected females were reaching the same level almost simultaneously (b), and that there was a delay in their steep phase (starting in 2006). He was also surprised to realize that one of the 5 selected authors was unexpectedly going down (c). The expert found the technique very useful to visualize both the animated table—to follow entries, observe values, and compare dimensions—and the line charts at the same time *"to identify trends and make comparisons, find mountains and valleys. And especially because I can filter which line chart to show."*. He used the feedback provided on the time slider at the top of the table to match his time-stamped memories. He did not often change the design parameters: he sporadically used step-before interpolation to express the discrete aspect of data, such as the number of papers, and he did not activate all entries but used the multiple selections to filter the data. After a while, he asked to explore other examples available with the technique, such as algorithm sorting executions available in our repository. He found that the technique made it possible to analyze the behavior of sorting algorithms, emphasizing the pedagogical potential of the technique that he would like to use during his classes. Finally, the expert had several suggestions for improvement: increasing the size of the cursor, displaying a cursor on every line chart, and showing a vertical line as a landmark following the cursor's horizontal position to make comparisons easier.

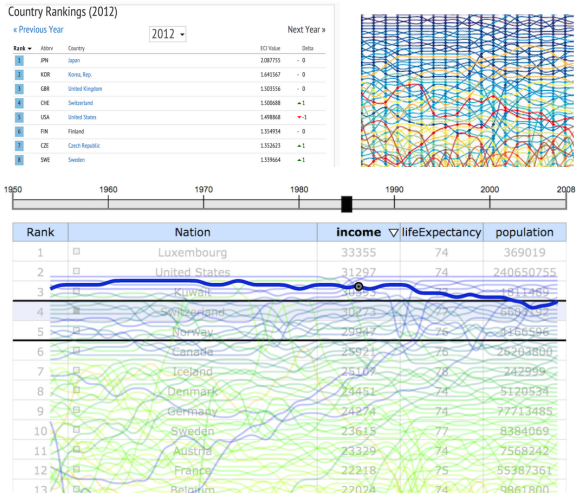


Figure 6. Top rows: table and line charts currently used separately by our expert in macroeconomics. Bottom: combined with our technique.

Economic Complexity Indicators

Our second expert is a research fellow in macroeconomics and has a continued interest in countries' rankings as a way to understand the relative impact of, among others, trade policies at a regional level. We presented him with a ranking table of countries containing World Bank Indicators and an Economic Complexity Index (ECI), a variable calculated using the number of products a country exports worldwide. Such an indicator is calculated based on 1354 product families from an open dataset on international trade data [4] from 1964 to 2012. Our expert decided to start by focusing on countries statistics he is familiar with, like population and incomes, for Switzerland. The table shows that in 1964, Switzerland is ranked first by incomes. In recent years, it has lost that ranking. To better explore this change over time, he dragged the cell containing the country and activated the line chart to immediately see a drop in the 70s and an average ranking around third place since then. Meanwhile, he observed important ranking variations for the year 1971 for all countries (Figure 6), but he knows this is normal using this dataset. The technique was very helpful to grasp the magnitude and exact time point for this change (while before he had to mentally connect a line chart and a table with the same data as shown on figure 6). He would have liked such variation not to greatly impact the navigation so he suggested to jump over those gaps when dragging. Overall, he valued the option to select multiple countries. This is because *"those indicators are not meaningful in absolute, only in relative terms as it is interesting to see the dynamics of countries over time"*. One issue was that scrolling down the page made the table's headers and years disappear. He wanted to always have access to that information. Thus, he changed the vertical scale of the line chart to make it fit the table height and found this feature very useful to provide a full overview (**H4**). He finally ended up re-creating the chart visual encoding he was familiar with (Figure 6) which greatly helped him find his feet when dealing with all the countries at once.

FINDINGS AND DESIGN IMPLICATIONS

From our study, the technique supports our three initial tasks: retrieving values (**T1**), comparing values over time (**T2**), and

performing tasks in the values domain (**T3**) with both the table and the line chart being complementary. Typically, the line chart helps to locate an extrema or a slope of interest and then the user switches back to the table to analyze the raw data across multiple dimensions. We also frequently observed experts performing higher-level interactions (e. g., moving countries where *"they should normally be"*) which remains to be fully characterized and understood.

In real-case scenarios, **missing data and high variability** happen. Missing data can be supported by skipping time intervals where data are absent, or by using a decoration (dashed lines, color) showing those gaps. For example, countries can split or merge, and sometimes they are even discarded from the rankings as their indicators are not significant enough. Regarding the variability, some sorting algorithms permute all elements during the first time steps, resulting in a sawtooth line chart that impacts the navigation with frequent ups and downs. One solution may be to use a different dragging trajectory than the current line chart, such as one that partially fits the data points to provide a smoother line.

Our second expert questioned the **affordance to activate the technique** *"How will you teach people to activate the technique? You can click on all the cells!"*. This means a mechanism should indicate where the technique can be activated or not. This is a well-known problem common with direct manipulation techniques [8]. Visual cues (e. g., icons that can be used as handles) can be added to indicate that dragging is an option. To minimize the visual overload of the table, they could only be visible when the mouse pointer hovers a cell.

LIMITS AND PERSPECTIVES

We now discuss some design decisions and limitations raised by our participants and the anonymous reviewers.

As we observed during the study, the technique is not self-discoverable. Another issue is an eventual conflict of interaction, as other direct manipulation techniques may use the same mouse move. For instance, in À Table [6] vertical dragging is used to change values and while current implementation features horizontal dragging, it is very close. Especially since we considered 2D line curvature drag (which requires following the curve exactly to change the time variable), but we found out that a horizontal drag provides excellent usability as a ranking function is injective, i. e. all values on the horizontal axis followed by the mouse match a unique value on the curve.

As the technique introduces an offset between the line chart and the current time point, the user has to look at the time slider to retrieve the current time. None of the experts raised this point as an issue and our assumption is that when interacting with the technique, as opposed to looking at snapshots, the dynamics makes it more intuitive. The animated transition between the cell and the focus point on the line chart also helps to connect the two representations. We considered adding time labels along the curve, but it created visual clutter.

The line chart's design remains yet to be fully explored. During our design and study, its representation changed to support different datasets and tasks. As we showed for academic ci-

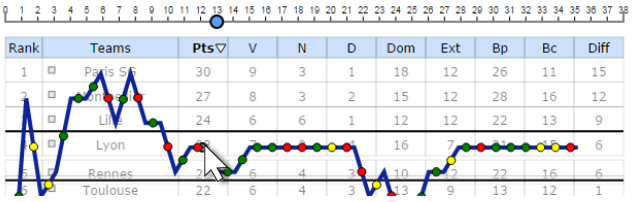


Figure 7. Additional dimension visualization for a soccer championship.

tations, just by adding **additional dimensions visualizations**, the expert managed to gain more insight without loosing usability. We also implemented several alternatives, and a notable one is ranking tables of soccer championships. In this context, the rank of a team evolves after the result of each day of the championship. Figure 7 shows how to represent this information using colored circles: green for a win, yellow for a draw, and red for a loss. However, this specific design is not applicable for countries rankings since no event, besides annually updating the values, causes change. When we asked our economics expert what would be a similar additional dimension, he said *"major economic events should be visible, such as trade agreements or crisis, but they sometimes span over multiple year and only concern a handful of countries"*.

Finally, the main perspective of our work is to extend the current technique to better support comparisons across multiple columns during the activation phase (e. g., comparing both countries' incomes and population). Similarly, also to explore the support of comparison between different time resolutions (e. g., comparing multiple soccer seasons) while dragging.

ACKNOWLEDGMENTS

We thank Stuart Russell for his help proofreading the paper and the anonymous reviewers for their relevant suggestions.

REFERENCES

1. Amar, R., Eagan, J., and Stasko, J. Low-level components of analytic activity in information visualization. In *INFOVIS 2005.*, IEEE (2005), 111–117.
2. Andrienko, N., and Andrienko, G. *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*. Springer-Verlag New York, Inc., 2005.
3. Benbasat, I., and Dexter, A. An investigation of the effectiveness of color and graphical information presentation under varying time constraints. *MIS Q. 10*, 1 (Mar. 1986), 59–83.
4. Hausmann, R., and Hidalgo, C. A. *The atlas of economic complexity: Mapping paths to prosperity*. MIT Press, 2014.
5. Jarvenpaa, S. L., and Dickson, G. W. Graphics and managerial decision making: Research-based guidelines. *Commun. ACM 31*, 6 (June 1988), 764–774.
6. Perin, C., Vuillemot, R., and Fekete, J.-D. À table!: Improving temporal navigation in soccer ranking tables. In *Proc. CHI '14*, ACM (2014), 887–896.
7. Rao, R., and Card, S. K. The table lens: Merging graphical and symbolic representations in an interactive focus + context visualization for tabular information. In *Proc. CHI '94*, ACM (New York, NY, USA, 1994), 318–322.
8. Shneiderman, B. Direct manipulation: A step beyond programming languages. *Computer 16*, 8 (1983), 57–69.