Party Tracker: Visualizing Changes in Party Affiliation over Time

Souvik Bhattacherjee
Department of Computer Science
University of Maryland, College Park
bsouvik@cs.umd.edu

Yuwei Cui
Neuroscience and Cognitive Sciences Program
University of Maryland, College Park
ywcu1990@gmail.com

Mohit Iyyer
Department of Computer Science
University of Maryland, College Park
m.iyyer@gmail.com

Alex Memory
Department of Computer Science
University of Maryland, College Park
alexmemory@gmail.com

James Parker
Department of Computer Science
University of Maryland, College Park
jprider@cs.umd.edu

ABSTRACT
Information visualization has become an increasingly important issue for political and social sciences with the accumulation of large amount of data from various sources. In this paper, we present PartyTracker, a new visualization tool that allow user to analyze the long-term party affiliation survey data from people with different demographic background across the United States. To visualize this large multivariate dataset, PartyTracker combines techniques of both geographic visualization and time series data visualization. The user can switch between a “map” view and a “time” view to explore different aspects of the data. Usability testing and expert review has been conducted to evaluate the effectiveness of this tool. The outcome of this work provides an interactive way to visualize the historical changes of party affiliation, which will inform future work of partisanship and generate insights that are interesting for general public.

Keywords
information visualization, political science, geographic visualization, themeriver, d3

1. INTRODUCTION
During the past few decades, political and social sciences have accumulated a large amount of data from various sources such as public opinion polling, demographic surveys and other empirical social science researches [3, 12]. As a result, reasoning in political science relies more and more on quantitative analysis of large dataset. This can be difficult for analysts because human capacities for interpreting large, multidimensional information are rather limited. Thus, it is important to use data visualization techniques to present such multivariate data in a meaningful way for both policy makers and the general public.

In this study, we focus on the problem of visualizing changes in political party affiliation over time. The United States’ political landscape is dominated by its two-party system, which includes the Republican Party and the Democratic Party. In every campaign cycle, political parties pay close attention to the details of every election. It has been an important topic in political science to understand how various social and demographic factors affect political party affiliations [6].

To help political analysts study multivariate data over time, we have created the interactive visualization, Party Tracker. In addition, we have performed usability tests to evaluate the effectiveness of this tool. The data used in this visualization comes from survey data provided by the Pew Research Center. This data is used in our visualization to show how demographic groups have changed their party affiliations over time.

2. RELATED WORK
The survey dataset of party affiliation has both geographical (across states) and temporal (over years) aspects, as well as several additional dimensions specify-
ing demographic variables (age, gender, income, education, state). Although we did not find many studies addressing the same problem (visualization of party affiliations), there are two related problems that are well studied in the field of information visualization, which are (1) visualization of large high-dimensional geographic datasets [7] and (2) visualization of multiple thematic variations over time [9]. We will discuss related works in these two areas in this section.

Changes of party affiliations do not develop uniformly over space. Geography is an important factor when trying to explain spatial aspects of an elections outcome and behavior of voters. In fact, electoral geography is an important subfield of political geography since the 1970s [2], where the focus is the spatial variation of voting behavior. One prevailing perspective in explaining these variations argues that social and political interests in a particular region may be very different from the national trends [1]. As a result, “people in similar socio-economic conditions may vote differently depending on where they live” [19]. This is because people lives in the same area shared same local factors. Individuals tend to join the dominant political opinion regardless of their backgrounds, which is known as the “neighborhood effects”.

Geographical data like the survey data can be naturally presented with a geographic map and several layers of information on top (known as semantic layers or thematic maps). A system designed to capture, store, analyze and present geographical data is known as the geographical information system (GIS), which represents a very well developed branch of information systems. GIS is particularly helpful to visualize the spatial location and densities/quantities of data. Geovisualization tools have been used extensively in analyzing electoral patterns [14, 4]. Such tools are great to visualize the result of a particular election. However, it is challenging to explore the change of election results or party affiliations over time with a single map. Even with animation, it can be difficult to read out the trend from successively presented maps.

Another prevailing view in electoral studies argues that “similar people vote similarly, independent of their location” [16]. This is known as the compositional view. According to this view, voters should be studied as a national unit, not necessarily regionally [19]. For this purpose, it would be suitable to visualize the electoral results or public opinion polling data with temporal data visualization techniques. Line chart and scatter plot are two commonly used methods for temporal data visualization. But they are not very space efficient and it is hard to visualize the trend when many time series are simultaneously displayed. Horizon graph is a recently proposed time series visualization techniques with high data density and better graphical perception than line chart [11]. In a horizon graph, a line chart is divided into bands and overlaid to reduce the space. ThemeRiver is another popular method for temporal data visualization, where the strength of multiple themes over time is shown as “rivers” with changing width [9]. More sophisticated methods transform the time series into a symbolic representation, such as trees [15] or glyphs [8]. There are also several techniques dealing with the hierarchical-temporal data. For example, Thern et al. [17] proposed to use a tree-ring like layout and Lammarsch et al [13] developed a pixel-based visualization that overlay several time granularities to utilize the inherent hierarchical structure of time.

The existing works in visualizing electoral results or partisanship affiliations either use geographical visualization techniques (NYTimes 29 Nov 2012[18]) or time series visualization methods (e.g., NYTimes 7 Nov 2012 [5]). However, to have an complete understanding of the dataset, it would be necessary to visualize the data both geographically and temporally. Therefore, we developed PartyTracker with the need to accomplish these tasks in mind.

3. DATA DESCRIPTION

The data used in the Party Tracker visualization is the result of 442,262 telephone surveys conducted by the Pew Research Center from 1990 to 2013. Each data point contains demographic information including the respondent’s age, gender, income level, race, education, and geographical location. In addition, the data points include the date that the survey was conducted, as well as the survey’s post stratification weight.

The post stratification weight is needed to ensure that the survey data is a representative sample of the overall population [10]. When computing statistics about the data, we take this weighting into account by counting each data point by its post stratification weight. Since we used survey data, we were also limited in the results we could present to users. For a result to be significant and accurate, it needs to be supported by at least 100 weighted data points. When filtering the data in the visualization, it is possible for this condition not to be true. As a result, the visualization indicates to the user that there is insufficient data whenever this happens.

4. DESIGN AND IMPLEMENTATION

We decided to implement two different views, the map view and the time view, in our visualization in order to make it easy for our users to investigate both nationwide variations and temporal changes in parti-
Figure 1: User interface for MapView. The map view contains a geographical map of the United States. Each state is colored on a red/blue spectrum depending on its ratio of Republicans to Democrats. The user can specify the time range with the slider above the map and select a demographic group of interest with the panel on the right.
Figure 2: User interface for TimeView. The time view presents the trend of partisanship change for a selected demographic group (specified by the filter).

Partisanship affiliation. Both views share the same basic layout, which consists of a main panel for the view, a filter panel that allows users to select demographics of interest, and a legend panel. Since Pew wants to post our visualization on their website, we designed it with the client-side experience in mind. Early on it became apparent that our dataset was too large to send to the client, and so we created a web service that queries data on the server-side and sends very small data files to the client every time a filter is updated.

The map view, which is by default the first view the user sees, displays a map of the United States where each state is colored on a red / blue spectrum depending on its ratio of Republicans to Democrats (Figure 1). While early versions of our visualization changed the color domain every time the filters were changed, we found that the map is more intuitive if we use a fixed color domain. Our current version associates the extreme red (Republican) value with a Republican / Democrat ratio of 5 and the extreme blue value with a Democrat / Republican ratio of 5. Users are able to click on states and zoom in, which is useful for smaller states such as those in the Northeast. In addition, mousing over a state activates a hover window that shows a detailed break-down of the party affiliations within that state. We implemented the map using a combination of jQuery and d3.js.

Above the map is a range slider that facilitates moving forward and backward in time over the domain 1990-2013. In addition, the slider features two adjustable handles that control the number of years over which to perform aggregation (the window size). We initially created the visualization with a fixed five-year window, but our mentor suggested that we give users the ability to increase and decrease it. At smaller window sizes, comparisons between individual years can be made. However, the finer granularity comes at a cost: more states have statistically insignificant data at these smaller window sizes. Similarly, while more states have significant data at higher window sizes, users are only able to compare fairly large periods of time. We implemented the range slider using jQuery - while earlier versions of our slider were coded entirely in d3, we found it difficult to implement the adjustable ranges while retaining slider dragging, so we switched to a modified version of an existing jQuery slider.

Many important questions can be answered with the map view alone, especially those focusing on regional trends. However, temporal trends aren’t immediately obvious from this view; the user has to drag the slider around to reveal them. In order to make this process
easier, we implemented the time view, which presents all of the data for a specified set of filters in one easy-to-understand visualization (Figure 2). Users first select a filter category to display timelines by an attribute of interest (e.g., Race or Gender). For each of the options in the selected category, we display a separate timeline. For example, if Gender was selected, we display one timeline for males and one for females. We used the ThemeRiver to visualize the temporal trend [9]. The x-axis of each timeline represents the time in years, while the y-axis represents three bands stacked on top of each other, which represents Independents, Democrats, and Republicans respectively. The width of each band represents the fraction of people with that partisanship affiliation. While the time view doesn’t offer any regional insights, questions about the relative proportion of parties in the selected demographic can be answered quickly, which makes it valuable enough that we made it a separate view. The time view was implemented using only d3.

5. USABILITY TESTING

Subjects for our usability tests are users who read news via the internet. The subjects were asked to fill in a survey which included personal information and their domain knowledge which in this case is US political history. Before starting the usability test with the subject we spent some time introducing our visualization software and maintained a checklist with us to make sure we have covered all the relevant features of our software that the user is supposed to know. Our visualization software consists of two major views: Map view and Time view. Under each view we make sure the user understands the following:

Map View:
1. Coloring of the states based on Republican (red) or Democrat (blue) leaning
2. Hovering over each state would provide the relevant information for state
3. Zoom-in and zoom-out feature of the map

Range Slider in the Map View:
1. The selection of the years as a range
2. The labels in the slider with the President names (the different positions of the labels depict the year in which they assumed office)

Time View:
1. Coloring of the time view based on Republican or Democrat leaning
2. The different timelines based on Gender, Race, Age and Education

In addition we have a common filter for both the views and we made sure that the user understands how to operate the filters to narrow down to a particular group of people.

Figure 3: Usability Testing Result. (A) Correct rate for the 11 questions we asked the user to answer. The first three questions are related to MapView, while the remaining questions are best answered with the TimeView (B) Evaluation score on the four aspects of the tool.
We start the usability testing after this brief training session and inform the user that he/she can quit the testing process at any point in time. Also the user might ask us about any clarifications that might be required while the testing is on. The users were provided with the following set of tasks to accomplish by using the PartyTracker visualization software. We asked the users 11 questions related to identifying changes of partisanship affiliation over time. The questions are listed as follows:
1. Which states switched partisanship from 2005 to 2006?
2. In what three-year window was California the most Republican?
3. Who was/were the Presidents during this window?
4. How has the partisanship of women changed over time?
5. How has the partisanship of college graduates and post-graduates changed over time?
6. Do Hispanics become more democratic or more republican over time?
7. In which age group do Independents currently make up the highest percentage?
8. Are people with at least some college education trending towards being more Republican or less Republican?
9. Are there clearer trends for minorities within age groups?
10. Which gender tends to have a higher percentage of Democrats?
11. For people with less than high school education, is there any trend towards a particular party in recent years?

Some questions involves usage of the map view (Q1-3), while others can be better addressed with the time view (Q4-11). The users are able to correctly answer most of the problems, with an average score of $8.5 \pm 1.3$. However, we found some questions are easier to answer the others. For example, only 4 out of 10 subjects correctly answered the first question, which involves identified switch of partisanship. In the subsequent survey section (see below), three of the subjects complained that it is hard to tell switch of colors with a graded color scheme. We have carefully observed the subjects behavior during the usability testing, and documented the difficulties encountered by subjects. After the usability testing is done, we asked the users to fill up a survey form to rate their experience of using PartyTracker to accomplish the tasks. On a scale of 10, with 10 being excellent and 1 being poor, we ask the following questions:
1. How intuitive was the tool and its controls?
2. How easily could you identify changes in partisanship over time?
3. What is overall impression of the map view?
4. What is your overall impression of the time view?

The users generally felt the PartyTracker software is intuitive to use (8.1/10) and helps to identify changes in partisanship over time (8.2/10). The users had a good overall impression with the map view in particular (8.5/10), while some users had difficulty to use the time view (7.9/10). We also asked the user about any feedback/suggestions they might have for improving both map/time view. A shared concern for the map view is that We have carefully considered these suggestions and fixed several features that many users have difficult to use.

6. DISCUSSION

7. ACKNOWLEDGEMENTS

We would like to thank our mentor Paul Hitlin from Pew Research Center and Ben Shneiderman from University of Maryland for helpful suggestions during the development of this study. We would also like to thank Robert Suls for help with data analysis.

8. ADDITIONAL AUTHORS

9. REFERENCES


