Watching What and How Politicians Discuss Various Topics - A Large-Scale Video Analytics UI

Emily Song, Joseph G. Ellis, Hongzhi Li and Shih-Fu Chang
Columbia University
{emily.song, jge2105, hongzhi.li, sc250}@columbia.edu

ABSTRACT

Accurately gauging the political atmosphere is especially difficult in this day and age, as individuals have access to a constantly growing collection of written and audiovisual news sources. This is especially true with regards to the U.S. presidential election, as there are numerous candidates, countless stories, and opinion articles discussing the merits of each particular candidate. It is therefore challenging for people to make an accurate assessment of what each candidate represents and how they would act if they were elected into office. To address this problem, we present a large-scale dataset comprised of videos of politicians speaking organized by the topics they are speaking about, and a user interface for exploring this interesting dataset. Our interface links people and events to relevant pieces of audiovisual media, and presents the desired information in a meaningful and intuitive manner. Our approach is unique by direct linking to actual speaking by politicians about specific topics, rather than links to textual quotes only. We describe the larger underlying infrastructure, a novel automated system that crawls thousands of internet news sources and 100 television news channels daily, and automatically discovers entities and indexes the content into events and topics. We examine how our user interface provides helpful and unique insights to its users, and give an example of the type of large scale trend analysis that can be performed with our system. Our online demo can be accessed at:
http://www.ee.columbia.edu/dvmm/PoliticialSpeakerDemo

Keywords

Multimedia Application, Topic Linking, User Interface, Visualization Platform, Person Naming, Event Detection

1. INTRODUCTION

In the world today, we are exposed to politics through a wide variety of sources. With evening newscasts, talk shows, newspaper articles, and other forms of digital media all around us, we have greater access to news on this subject matter than ever before. We are subjected to a constant stream of stories and opinions, skewed in one way or another, to the point where we are unsure of how to make sense of it all. In this environment it is difficult to make well-founded judgments about important issues, and where candidates stand on them.

Political awareness is crucial to understanding the most pressing matters of today, but it is difficult to achieve, considering the abundance of news available. This is especially true with regards to the political election, where extensive media coverage has resulted in an overflow of information. Given that the presidential election is arguably the most important political event in the nation, there is an urgent need for a system that organizes this wealth of material in a coherent manner.

With so many candidates, and so many stories, people are overwhelmed and struggle to see the complete picture. How can individuals keep track of what different candidates are saying, what stories are trending, and how their statements on these topics changed over time or vary in comparison to one another? To address this problem, we built a user interface that allows a user to search and display the actual

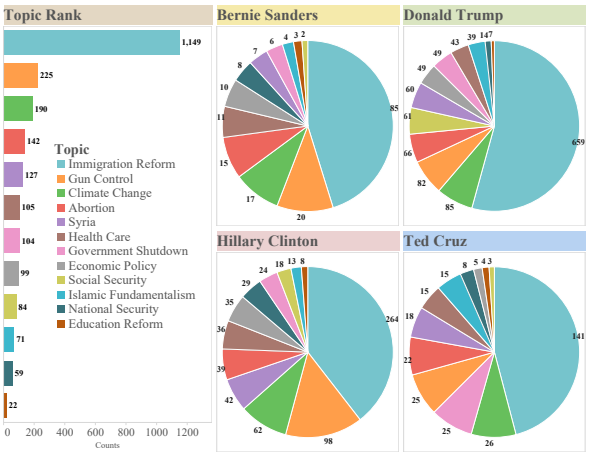

Figure 1: Breakdown of the political topics covered by different politicians and the aggregate statistics of the whole dataset broken down by topic.

ICMR '16, June 06−09, 2016, New York, New York USA

ISBN 978-1-4503-4359-6/16/06…$15.00

DOI: http://dx.doi.org/10.1145/2911996.2912025

presidential candidates talking about different political topics in videos, rather than text quotes only. We automatically discover from raw television news streams when a politician appears and link what they are talking about to important political topics that take place in this election. Our system allows a user to search our video database by political candidates and topics, and to view the desired information on a well-organized timeline.

'A simple search with this tool leads to valuable and unique insights. For instance, users will quickly find that some candidates speak more on certain topics than others, and that clusters of speaker segments indicate trending topics. In addition, users can take advantage of the scalable timeline feature in order to analyze how individuals have changed their opinions over time, or how individuals have experienced surges in popularity at certain moments in time.

The contributions of our work can be summarized as follows:

- Using our automated video processing pipeline, we have generated a large collection of video political speech segments from the U.S. presidential candidates in the 2016 presidential election.
- Our system automatically detects and links video political speech to specific news events, and broad political topics.
- We demonstrate an interesting data analysis study that can be done with the massive collection of speech segments.
- Our video collection is fully searchable, and we provide a novel user interface for analyzing politicians' views over time and compared to one another.

1.1 Related Work

Political analysis using multimedia information is an interesting research topic. Many works from interdisciplinary researchers focus on this area. The automatic tools developed in the data mining, machine learning, and computer vision community play an important role in political science. Stewart [4] analyzed and provided a set of natural language processing tools for political scientists. Information extraction from "big data" [1], event detection, linking and coreference techniques discussed in [9] and [11] are useful components in building a political analysis system. In this paper, we discuss how we leverage state-of-the-art techniques to build a fully automatic data collection, analysis and representation system that explores information regarding political candidates and topics.

2. SYSTEM OVERVIEW

In this section we detail our video processing and recording pipeline that has enabled the creation of our archival political demo. The system architecture is shown in Figure 2.

2.1 Recording Infrastructure

Our system has access to over 100 US news channels. Our infrastructure can record and process 12 channels simultaneously in real-time. We record both broadcast and cable television news programs, examples of the programs that are collected are: *CNN Newsroom*, *The Today Show*, *ABC World News*, *NBC Nightly News*, *News 4 New York at 5*, and *The Situation Room*. Our system automatically tunes a digital TV tuner to record any news programs that are available at given points in time, based on the information provided by an Electronic Program Guide. Each of the chosen programs are recorded in their full digital transport stream (.ts), and then transcoded down to a more compact format (.mp4) to conserve space on our storage servers and allow for display in all browsers, which support this format.

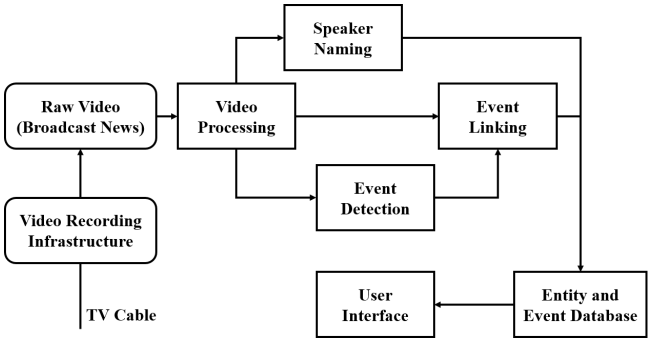

Figure 2: System Architecture

Once the videos are fully recorded we then begin processing each video by performing automatic speech recognition on the audio track within the video using the CMU sphinx library [7]. We extract the closed caption text that accompanies each video program, and then time-align the closed caption transcipt to the generated speech transcript from ASR. The closed caption text tends to be more accurate, but is shifted from what is actually being said on screen. By combining the two sources we are able to obtain a time-aligned accurate transcript of the speech on screen. Across one news program many different topics are covered, we want to separate each story, so that we can accurately search and determine what each politician is speaking about. To accomplish this we segment the news program into distinct "story segments", which are portions of the news program in which only one topic is covered. To accomplish this we utilize a variety of multimodal cues, such as closed caption special characters, black screen detection, shot boundary detection, consistent visual logos, and more to detect commercials in the news program and change of segments across the program. For more information on the algorithms used in this work to perform story segmentation and commercial detection, please refer to [5], [6], and comskip [1] .

2.2 Event Detection and Person Detection

Once we have separated the full news program into separate news stories we then attempt to detect when breaking news or important news events have been generated. To perform this operation we have developed a data mining algorithm for finding breaking news events. The details of the algorithm can be found in [6] and [8].

A unique aspect of our system is the ability to detect and find speakers from our collected and recorded news stories. Politicians are of particular interest in news, and our system is able to automatically detect who is talking on-screen with a precision of 86%. We employ two algorithms for speaker detection, which are documented and described here [6] and

[1]comskip: http://www.kaashoek.com/comskip/

[3]. Due to space considerations, we have omitted the details of our speaker detection algorithms from this paper. We encourage the reader to explore the algorithms in the cited works.

2.3 Political Topic Linking

To allow a user to browse through the broader political topics that are shaping the 2016 political election we assign each political speaker segment to its appropriate political topics. First, we have developed a crawler to automatically monitor the topics and content that are aggregated by Google News. We download new articles and Google news topics every 5-10 minutes. To date, we have found 45,000 distinct news topics and 25 million articles that are linked to those topics.

To discover broader political categories we have manually defined 12 distinct topics that are prevalent in coverage of the 2016 presidential election. Examples of these topics are *Immigration Reform*, *Gun Control*, and *Abortion*, and a complete list can be seen in Table 2. To link each speech segment to a political topic we extracted the closed caption from a given segment, and used n-gram similarity to find the closest political topic. A more detailed look at our topic linking algorithm can be seen in [8] and [6].

3. STATISTICAL ANALYSIS

Using our video political speech dataset one can extract insightful observations, examples of which are summarized in the included Tables 1 and 2 and Figures 1 and 3. Upon examining the number of videos per speaker and the number of videos per topic, we find that media coverage of the presidential election is far from balanced. This finding strongly suggests the presence of media bias, which in earlier studies has been found to greatly influence voter opinion and election outcome [2]. For example, Donald Trump and Hillary Clinton have so far dominated the media coverage in their respective parties, and not surprisingly at the time of publication they are both the odds-on favorites for the primary nominations. Interestingly, Ted Cruz at the time of writing has performed second to Donald Trump in an aggregation of national polls[2], and we have found that Ted Cruz is in the top three in total speech segments for republicans in our dataset. Conversely, Jeb Bush appears the second most often in the dataset, but is currently more than 10% behind Ted Cruz[2]. It is obvious that their is a correlation between performance in national polls and media coverage, whether it is a causal relationship is unclear. We hope that further analysis with our video dataset could be enlightening as to the causal nature of this relationship.

Furthermore, upon examining the breakdown of the topics discussed by each speaker, we discover noticeable fluctuations of the topics covered over time in Figure 3. What candidates are talking about is altered by when particular events happen, and the overall interests of the media shift and change.

Our comprehensive presidential candidate speaking set is useful in that it allows us to highlight disparities in media coverage and to pinpoint trends in political discussion, thereby allowing us to present a more complete picture of the speakers and topics. For example, if we look at Figure 1 we see that Immigration Reform is popular, but Donald Trump

[2]fivethirtyeight.com

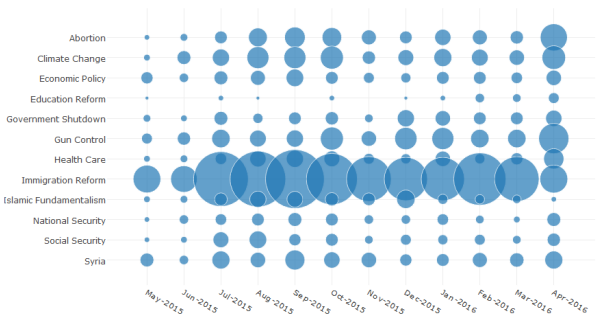

Figure 3: The amount of video story segments for each topic over time.

Table 1: Breakdown by Speaker

Speaker	Number of Segments
Jeb Bush (R)	817
Ben Carson (R)	67
Chris Christie (R)	109
Ted Cruz (R)	520
Carly Fiorina (R)	44
John Kasich (R)	111
Rand Paul (R)	143
Marco Rubio (R)	224
Rick Santorum (R)	76
Donald Trump (R)	1950
Hillary Clinton (D)	1143
Bernie Sanders (D)	321

speaks about immigration reform proportionally more than any of the other leading candidates. This information may give us insight into the demographics and geographic portions of America that are more likely to vote for him. In the world of biased news reporting and a constantly changing political climate, this large scale data analysis can be very enlightening to expose biases or a candidate's tendencies. The patterns discovered here speak volumes about the issues most important to the candidates and their party affiliations [10].

4. USER INTERFACE

Our user interface exhibits a scalable timeline to represent the desired political candidate and topic the viewer is interested in. The visualization demo initially prompts the user to select multiple speakers and one topic from two separate drop down menus, and this search query then determines what speaker segments appear on the generated timeline.

The timeline has several notable features: (1) The timeline is scalable, allowing for easy comparison of statements over time. (2) Each speaker segment is an item on the timeline, flagged with a photo of the speaker and a title for the video. (3) Clicking on an item on the timeline generates the video, along with its date, title, and speech transcript.

The design of the timeline allows for accurate snap judgments, but also for careful exploration of the timeline content. An accurate snap judgment may occur when a user sees a clustering of videos around a certain date - this immediately suggests a trending event. An instance where

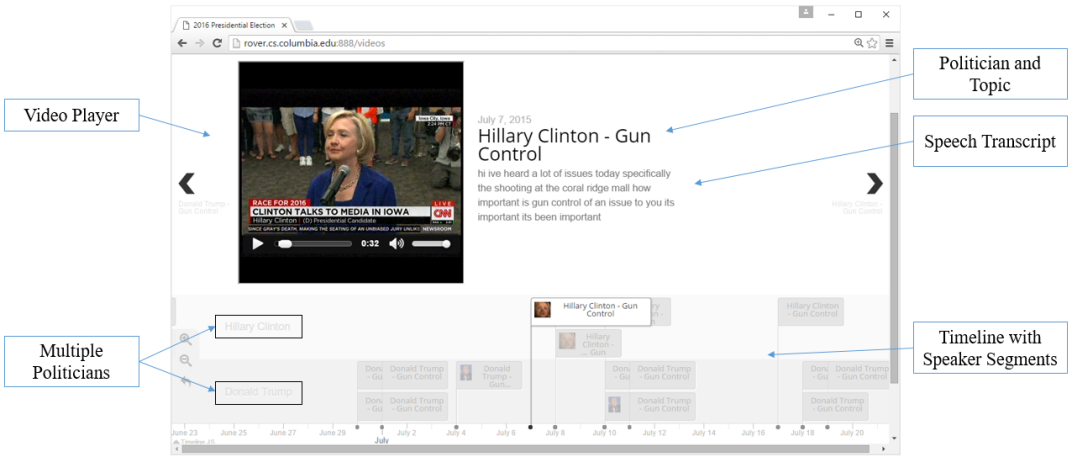

Figure 4: Topic linking user interface: Scalable timeline allows a user to easily navigate through the videos temporally. Videos can be viewed in time sequence, allowing for direct comparison of two speakers discussing the same event.

Table 2: Breakdown by Topic

Topic	Number of Segments
Islamic Fundamentalism	124
Syria	195
Health Care	153
Climate Change	268
Government Shutdown	138
Immigration Reform	1729
Gun Control	315
Social Security	113
National Security	100
Education Reform	28
Economic Policy	142
Abortion	221

careful exploration of the timeline content may be helpful is when someone wants to directly compare several candidates - systematically going through videos of different candidates speaking on the same days allows the user to find differences between their responses to the same event.

5. CONCLUSION

We have presented an automated video processing system and user interface that allows a user to search and watch videos of political candidates by political topics. This visualization demo aims to illuminate political attitudes in the presidential election with an intuitive timeline populated with relevant speaker segments. Our platform is notable for the following reasons: (1) Using our automated video processing pipeline, we have generated a large collection of video political speech segments from the U.S. presidential candidates. (2) Using this data, we have demonstrated interesting data analysis studies that can be performed. (3) Selecting people and topics with this tool can lead to meaningful realizations regarding the presidential election, enabling the voting public to be more well-informed. The demo website is live so that others can benefit from its usefulness. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-11-44155.

6. REFERENCES

[1] K. Arceneaux, M. Johnson, R. Lindstädt, and R. J. Wielen. The influence of news media on political elites: Investigating strategic responsiveness in congress. *American Journal of Political Science*, 2016.

[2] D. D'Alessio and M. Allen. Media bias in presidential elections: a meta-analysis. *Journ. of Comm.*, 2000.

[3] J. G. Ellis, B. Jou, and S.-F. Chang. Why we watch the news: A dataset for exploring sentiment in broadcast video news. In *ICMI*, 2014.

[4] J. Grimmer and B. M. Stewart. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 2013.

[5] W. Hsu, L. Kennedy, C.-W. Huang, S.-F. Chang, C.-Y. Lin, and G. Iyengar. News video story segmentation using fusion of multi-level multi-modal features in trecvid 2003. In *ICASSP*, 2004.

[6] B. Jou, H. Li, J. G. Ellis, D. Morozoff-Abegauz, and S.-F. Chang. Structured exploration of who, what, when, and where in heterogeneous multimedia news sources. In *ACM MM*, 2013.

[7] P. Lamere, P. Kwok, E. Gouvea, B. Raj, R. Singh, W. Walker, M. Warmuth, and P. Wolf. The cmu sphinx-4 speech recognition system. In *ICASSP*, 2003.

[8] H. Li, B. Jou, J. G. Ellis, D. Morozoff, and S.-F. Chang. News rover: Exploring topical structures and serendipity in heterogeneous multimedia news. In *ACM MM*, 2013.

[9] R. Li, K. H. Lei, R. Khadiwala, and K. C.-C. Chang. Tedas: A twitter-based event detection and analysis system. In *IEEE ICDE*, 2012.

[10] B. Yu, S. Kaufmann, and D. Diermeier. Classifying party affiliation from political speech. *Journal of Information Technology & Politics*, 2008.

[11] T. Zhang, H. Li, H. Ji, and S.-F. Chang. Cross-document event coreference resolution based on cross-media features. In *EMNLP*, 2015.