Mining for Surprise Events Within Text Streams

Paul Whitney, Dave Engel and Nick Cramer*

26th January 2009

Abstract

Text streams are a fundamental source of information that can be used to detect and characterize strategic intent of individuals and organizations as well as for detecting abrupt or surprising events within communities. In this paper we describe our algorithm development and analysis methodology for mining the evolving content in text streams. Text streams include news, press releases from organizations, speeches, Internet blogs, etc. Specifically, an analyst may need to know if and when the topic within a text stream changes. Much of the current text-feature methodology is focused on understanding and analyzing a single static collection of text documents. Corresponding analytic activities include summarizing the contents of the collection, grouping the documents based on similarity of content, and calculating concise summaries of the resulting groups. The approach reported here focuses on taking advantage of the temporal characteristics in a text stream to identify relevant features (such as change in content), and also on the analysis and algorithmic methodology to communicate these characteristics to a user. We present a variety of algorithms for detecting essential features within a text stream. Our approach for communicating the information back to the user is to identify feature (word/phrase) groups. These resulting algorithms form the basis of developing software tools for a user to analyze and understand the content of text streams. We present analysis results using both news information and abstracts from technical articles, and show how these algorithms provide understanding of the contents of these text streams. A critical finding is that the characteristics we used to identify features in a text stream are uncorrelated with the characteristics used to identify features in a static document collection.

1 Introduction

Massive amounts of text stream data exist and are readily available, especially over the Internet. Analyzing this text data for content and for detecting when things (such as topic or affect) change can be a daunting task. Mathematical and statistical methods in the area of data mining can be very helpful for the analyst looking for these changes. Specifically, we have implemented some of these techniques into a *surprise event detection technology* that is designed to monitor a stream of text or messages for changes within the content of that data stream.

Much of the research in information mining from text streams focuses on either the description of new events and salient features or in clustering the documents (text streams) [1, 2, 3]. Our research is focused on processing massive amounts of text streams to identify events that have just occurred or are currently occurring. You can think of this as a possible triage capability that an analyst needs to identify (surprising) events so that he can delve into the material to gain in-depth insight. However, finding these events in a timely fashion is not an easy task.

Some of the event types that one might want to detect in a text stream (which could be a sequence of news articles, a sequence of messages, or an evolving dialogue) are shown in Figure 1. In each case, time is along the x-axis. The y-axis corresponds to some measure of topic (such as the number of words or events that occur within your data). In the context of a text stream, a *point discontinuity* in topics could correspond to a document with a relatively unique content. A *jump discontinuity* could correspond to an abrupt change in the content of the text stream. A *slope discontinuity* could correspond to a ramping up or down in a topic for that text stream (e.g., emerging technologies). Typically, jump and point discontinuities are detected more readily than slope discontinuities [4].

The approach that we have taken is to monitor and evaluate the occurrence of individual terms in the text stream for changes in occurrence. As a preprocessing step, a text analysis tool is used to extract words from the text stream and give information about documents and words within the documents. With this information, algorithms are used to score each word (number of occurrences within documents). Using these statistical metrics (which we call our Surprise statistic),

*Pacific Northwest National Laboratory, Richland, Washington

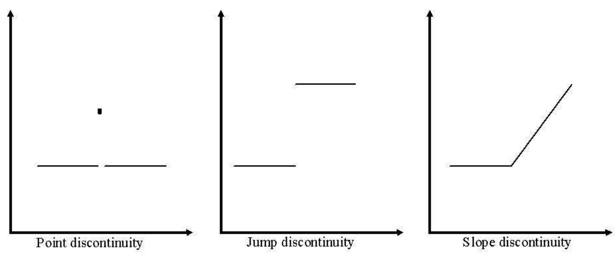

Figure 1: Typical surprising event types.

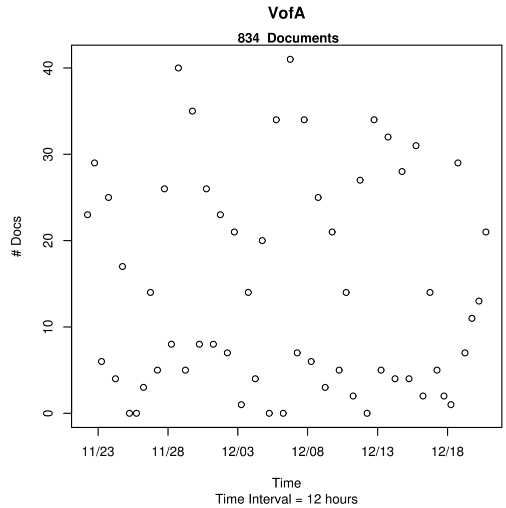

Figure 2: Number of documents for each time interval for the VofA data set.

we evaluate each word over the period represented by the text stream. When a sufficiently "surprising" word occurs, related words (based on the temporal profile) are found and are useful in explaining the broader nature of the event.

In this paper we focus the majority of our discussion on the develpment and implementation of the algorithms. A companion paper has been developed where a sensitivity analysis is performed on two different data sets with results being analyzed by a domain expert [5]. The domain expert confirmed that the results using Surprise matched historic occurrences.

Detected events and the explanatory words can be represented in a variety of ways. From our experience, graphical representations tend to be the most desirable (if not most useful) form for the analysts. The Surprise statistical algorithms along with graphical capabilities are discussed in Section 2. A comparison of our Surprise statistic with a topicality measure (which measures how interesting a word is) is shown in Section 3. Following that, we discuss the relevance of performing this analysis using two different data sets. Finally, conclusions about our data mining algorithms and analyses are discussed.

2 Surprise Statistics

Different algorithms for calculating "surprise" have been researched and five of these algorithms have been implemented into a toolkit. For each algorithm, the unit of calculation is a single word. Each of our algorithms requires a preprocessing of the time-sequenced documents. The processing that we use includes breaking the overall time period into intervals (we use hourly, daily, or weekly intervals, depending on the temporal granularity of the data being examined) and the number of documents in which the word occurred. The Surprise statistic is calculated in each time interval by comparing that occurrence with a moving window of the past occurrences.

To illustrate the use and also the differences between these algorithms, we used a data set from Voice of America (VofA) [6]. VofA is the official external radio and television broadcasting service of the U.S. federal government. Its oversight entity is the Broadcasting Board of Governors. Data streams are available through the Internet in both streaming media and downloadable formats.

The temporal profile of the number of documents (per time interval) within this data set (VofA) is shown in Figure 2. For this analysis, we chose 12 hours as the width of each time interval, which led to 58 time intervals.

The first Surprise statistic algorithm, within our toolkit, is based on a chi-square statistic (Pearson method) constructed from the following 2x2 table [7]:

x_t	$N_t - x_t$
$m\bar{x}_{-t}$	$N_{-t} - m\bar{x}_{-t}$

For this table, x_t is the count (number of documents containing this word) at time t, N_t is the total number of documents at that time, $\bar{x}_{-t}$ is the average of the counts of the word in all of the cells prior to time t, m is the number of cells prior to time t (so that $\bar{x}_{-t}$ is an average of the m counts), and N_{-t} is the total number of documents prior to time t. Note that the counts for $\bar{x}_{-t}$ and N_{-t} were restricted to a window of time steps prior to t. The amount of time (both the width of a time interval and the number of time windows, m) is a user-selected parameter of the procedure. A value sufficiently large for a chi-square statistic was one way to flag a surprising event/word. This statistic looks for deviations in the relative occurrence of x_t in N_t compared with the relative occurrence in the preceding time window (that is, in N_{-t}).

The formula used for the χ^2 statistic is

$$\chi^2 = \frac{n_{..}(|n_{11}n_{22} - n_{12}n_{21}| - Y\frac{n_{..}}{2})^2}{n_{1.}n_{2.}n_{.1}n_{.2}} \tag{2.1}$$

where the previous 2x2 table is rewritten as

n_{11}	n_{12}
n_{21}	n_{22}

and

$$\begin{aligned} n_{1.} &= n_{11} + n_{12} \\ n_{2.} &= n_{21} + n_{22} \\ n_{.1} &= n_{11} + n_{21} \\ n_{.2} &= n_{12} + n_{22} \\ n_{..} &= n_{11} + n_{12} + n_{21} + n_{22} \end{aligned}$$

Also, Y in Equation 2.1 is either 0 or 1. If Y is 1, the Yates continuity correction is applied for the low sample size in which the count in at least one cell is $\leq$5 [8].

The results of this algorithm applied to the VofA data set are shown in Figures 3 and 4. In Figure 3, the Surprise statistic calculated for every word of our text stream is plotted. The y-axis represents the sorted Surprise statistic, with the top 100 surprising words (largest Surprise score) being labeled. The temporal profiles, shown in Figure 4, represent the number of documents per time interval for the top 30 surprising words. Each profile is normalized by the maximum number of occurences (for a specific word). The (red) circle in each profile represents the time interval in which the Surprise statistic was the greatest.

The second algorithm for calculating our Surprise statistic is another form of the chi-square statistic known as the likelihood ratio. The likelihood ratio (for a hypothesis) is the ratio, of the maximum value of the likelihood function over the subspace represented by the hypothesis, to the maximum value of the likelihood function over the entire parameter space [9]. This statistic is calculated using the same 2x2 table as above and is shown in the following equation:

$$\chi^2 = \frac{n_{11}log\frac{n_{11}}{m_{11}} + n_{12}log\frac{n_{12}}{m_{12}} + n_{21}log\frac{n_{21}}{m_{21}} + n_{22}log\frac{n_{22}}{m_{22}}}{2} \tag{2.2}$$

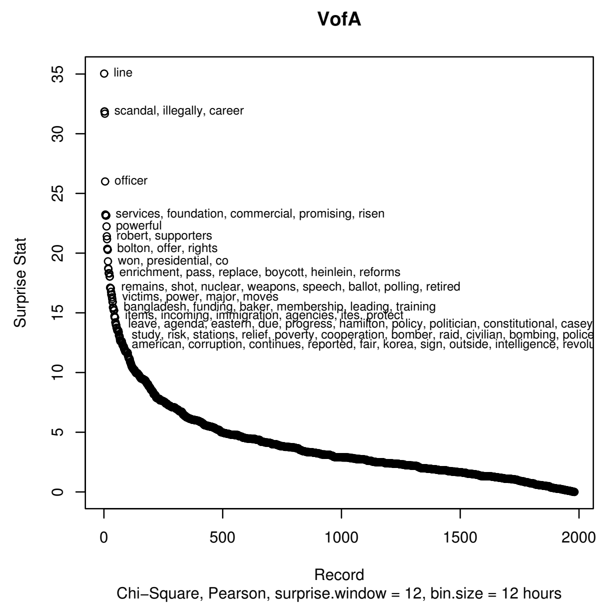

Figure 3: Sorted Surprise scores for the VofA data set, top 100 most surprising words labeled.

where

$$\begin{aligned} m_{11} &= (n_{11} + n_{12})(n_{11} + n_{21}) \\ m_{12} &= (n_{11} + n_{12})(n_{12} + n_{22}) \\ m_{21} &= (n_{11} + n_{21})(n_{21} + n_{22}) \\ m_{22} &= (n_{12} + n_{22})(n_{21} + n_{22}) \end{aligned}$$

Another of our algorithms for calculating the Surprise statistic is a Gaussian statistic. The Gaussian statistic is based on comparing the observed value x_t with the average over the previous values ($\bar{x}_{-t}$), normalized by the standard deviation of values in $\bar{x}_{-t}$ (s). We put a floor of 1.0 on the standard deviation used (because we are dealing with count data). This statistic is

$$G = \frac{x_t - \bar{x}_{-t}}{s \cdot (1 + \frac{1}{n})} \tag{2.3}$$

where n is the number of time intervals in the window and s is the standard deviation.

Finally, combining the previous algorithms (chi-square and Gaussian) forms the final two algorithms within our toolkit for the Surprise statistic. Each combined statistic is accomplished by taking the square root of the chi-square statistic plus the absolute value of the Gaussian statistic, as follows:

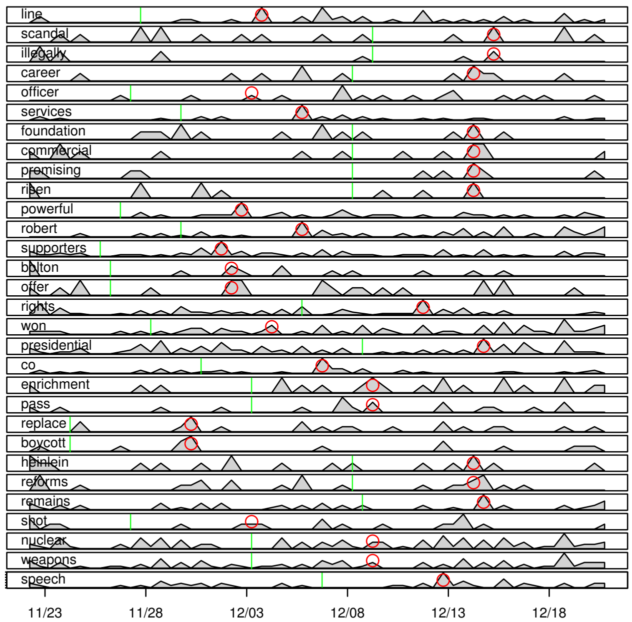

Figure 4: Temporal profiles for the top 30 surprising terms (VofA data set).

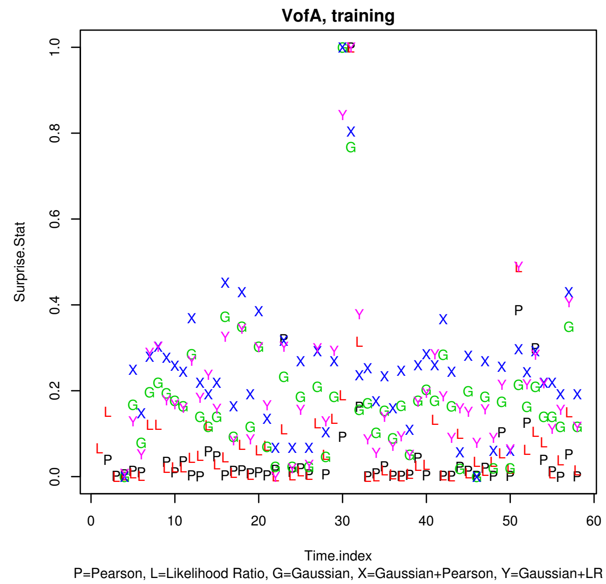

Figure 5: Comparison of actual Surprise statistic for each algorithm and each time interval for the word "training" (VofA data set).

$$C_{surprise} = \sqrt{\chi^2} + |G| \tag{2.4}$$

The results of applying all five of our algorithms on the VofA data set are shown in Figures 5 and 6. In Figure 5, the actual Surprise statistic for the word ("training"), for each time interval, and for all algorithms (including the combined algorithm shown in Equation 2.4) is shown. Note, that for this word, the time interval at which the Surprise statistic is maximum is different for the chi-square methods (represented by a black P for the Pearson method and a red L for the likelihood ratio) than for the Gaussian (green G) and the combined methods (blue X and pink Y). What is also interesting is that the two chi-square algorithms selected a surprising event of a negative slope discontinuity. These two points are illustrated in the temporal profiles shown in Figure 6, where the circles represent the time interval for the maximum Surprise score for all five algorithms.

2.1 Related Words. Similarities between words within a given set can give an analyst more information than just a single word can provide. We assess similarity based on the distances between vectors of the temporal occurrence of each word (i.e., temporal profiles as shown in Figure 4). There is a large number of candidate algorithms for calculating distances between the temporal profiles. Our preferred implementation is based on the

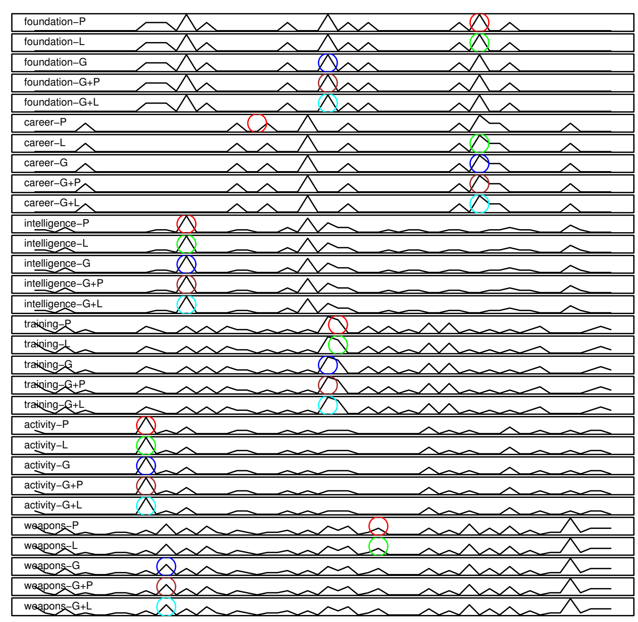

Figure 6: Comparison of Surprise statistic temporal location for each algorithm (VofA data set).

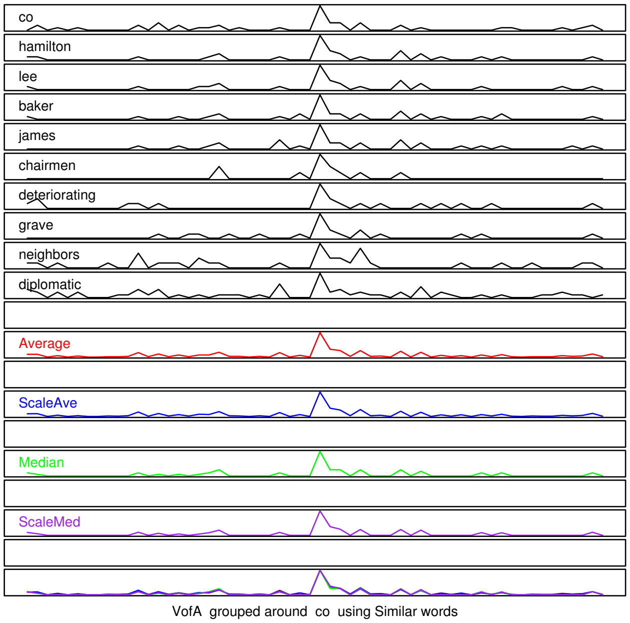

Figure 7: Top related terms (to "co") based on correlation distance (VofA data set).

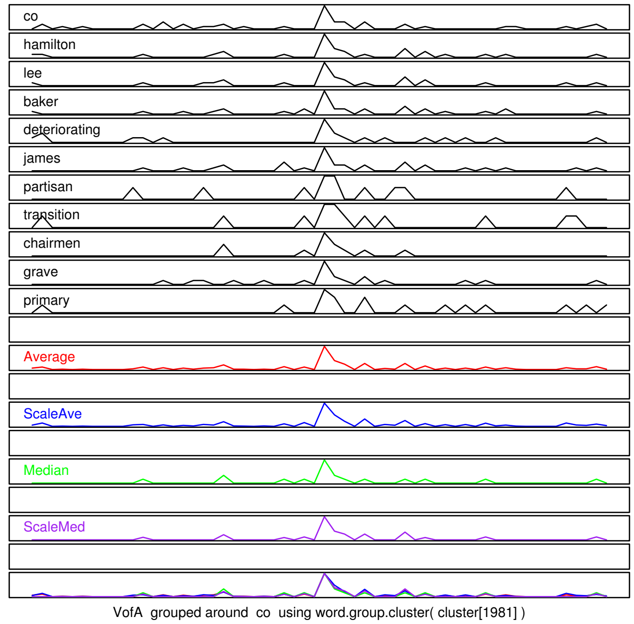

Figure 8: Top related terms (to "co") based on cluster grouping method (VofA data set).

correlation function between the vectors and is, for two such vectors (x, y), equal to $1 - |corr(x, y)|$. This distance often results in interpretable word groupings. Using combined related word profiles, one can gain more detailed information about the events. The combining (grouping) of the related words is described in the following section.

As an example of related words, Figure 7 displays the top 9 words (temporal profiles) related to the word "co". These related words give the analyst more information than what can be gained from just the single word. For example, we could get some entity information (Lee Hamilton and James Baker) and a tone of the text (grave and deteriorating) and some information about the actual topic (diplomatic and chairmen). When looking at the actual text streams for which this information was gathered, we see that the word "co" actually goes with the word "chairmen" (co-chairmen), referring to Lee Hamilton and James Baker who were co-chairmen of the Iraq study group, reporting to the President and Congress, back in 2006. These text streams were talking about the (war) strategy in Iraq (grave and deteriorating).

2.2 Groupings. In the previous two sections, we discussed words that are similar to each other (related) and also methods for selecting which words are of key interest (Surprise words). Typically, a large set of words exist in the text streams that we analyze. Grouping similar words can help reduce the dimensionality of the data set. Two grouping methods are presented in this section. First, the agglomerative hierarchical clustering algorithm is used to cluster the words around a selected word, which we call our clustering grouping method. Results of clustering the VofA data set for the word "co" are shown in Figure 8. For this analysis, complete linkage and a correlation distance algorithm was used [10]. From the clustering analysis, 10 words were selected to cluster (group) with "co". The number of members within the cluster was determined systematically within the algorithm. Also shown in Figure 8 are four aggregation methods, including a straight average at each time, an average at each time of scaled traces (scaled by the maximum of all included profiles), median value selected at each time, and the median at each time of the scaled profiles (same scaling as before). All four of these aggregation methods are plotted together in the last (bottom) profile in Figure 8.

Another method used to group words was based on the correlated distance that was discussed in Section 2.1 (Related Words), which we call our correlation grouping method. In this method, a word is selected, and the distance (correlation) is calculated between this word and all others words. The words are then sorted, and the words with the smallest distance (statistic) are selected for the grouping. This is illustrated in Figure 9. The

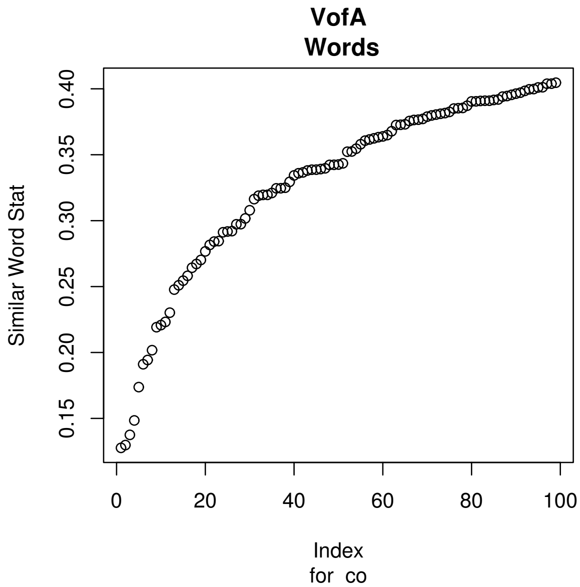

Figure 9: Distance metric (correlation) between the word "co" and the closest 100 terms.

top 9 closest words were selected (qualitatively selected from the similar/related word statistic plot) and were plotted in the temporal profiles shown earlier in Figure 7.

Utilizing these grouping methods, a new data set can be formed, in which single words are replaced with grouped words. The temporal profiles are then replaced with an aggregate profile (we used the straight average method as illustrated in Figure 8). This new data set can then be analyzed using the Surprise statistic technology. As an example, Table 1 shows a list of the top 25 groups (top Surprise scores) formed from our grouping analysis of the VofA data set. The groupings were formed using the clustering grouping method and the chi-square likelihood ratio Surprise statistic.

The analysis of this new data set produced the results shown in Figures 10 and 11. In these plots, the most surprising groups and their temporal profiles are shown. These groupings tend to provide more information about an event compared with the single word analysis as described in the previous sections.

3 Comparison with In-Spire's Topicality Index

Early, we made the claim that our characteristics used to identify features in a text stream are uncorrelated with the characteristics used to identify features in a static document collection. To illustrate, let's assume that the maximum value of the Surprise statistic (over time for a single word or term) could be thought of

Table 1: Top 25 surprising word groupings, cluster grouping method (VofA data set).

1	replace boycott
2	powerful ites
3	promising risen
4	ballot overthrew
5	specific implement
6	services nomination play vietnam colleagues
7	co hamilton lee baker deteriorating james partisan transition chairmen grave primary
8	organized square
9	heinlein bolton
10	career reforms rose
11	join saddam hussein headquarters
12	electronic equipment
13	don forum
14	roads relative
15	aziz abdul
16	agencies laws
17	bangladesh scores
18	marines anbar
19	remains closer leadership abducted
20	boy girl
21	goal sensitive
22	retired considerable
23	holy jihad lifted
24	robert gates highly fresh pentagon rumsfeld donald
25	amal nasrallah hassan

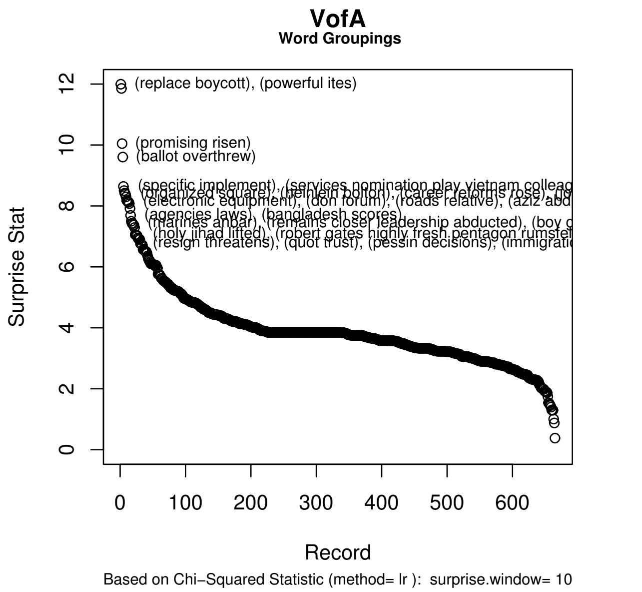

Figure 10: Sorted Surprise statistic for each word group (VofA data set).

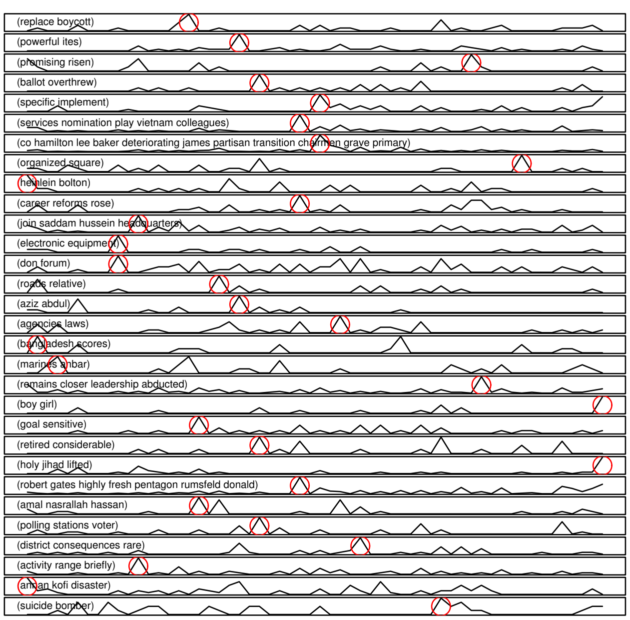

Figure 11: Temporal profiles for the top 30 word groupings (VofA data set).

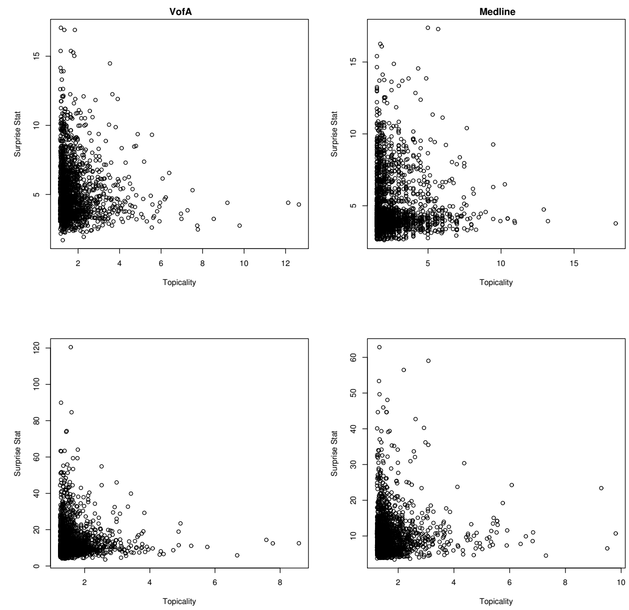

Figure 12: Comparison of Surprise statistic (y-axis) to the In-Spire topicality measure (x-axis) for four different data sets.

as a measure of interestingness of a word in a time sequence of documents. We compared the value of topicality with our measure of temporal interestingness to ascertain whether new information from our statistic is being gained regarding the document set. Note that the "topicality" is a measure used by In-Spire^{TM} and is a measure that defines discriminating terms within a set of documents. In-Spire is a text analysis and visualization tool developed by the Pacific Northwest National Laboratory [11]. This comparison was done using four different corpora (data sets). Figure 12 shows the top few thousand topical words for each of the corpora and how that topicality value compares with the Surprise statistic. The fundamental observation is that the metrics are uncorrelated, at least for these four different corpora, because no correlation is seen in any of these plots, which suggests that our temporal interestingness (Surprise) is new information, not currently extracted via topicality (from In-Spire).

4 Discussion

So far, we have discussed the algorithms that encompass our Surprise statistic. We have also shown an example of the algorithms as applied to a single data set (VofA). In this section, we continue the discussion about the results from the VofA data set (specifically about grouping results) and present some results from a data set that

Table 2: Word groupings of the VofA data set using the correlation distance method.

1	co hamilton lee baker james chairmen deteriorating grave neighbors
2	services nomination highly seize resolved risk confirmation
3	line opportunity study republican major
4	replace chiefs boycott
5	powerful ites connection warning corrupt
6	weapons china sanctions nuclear united
7	intelligence hill control agency

covers a much longer time frame.

4.1 Grouping Results. In Table 1, we showed the groupings (of the VofA data set) as produced using the clustering grouping method. As a comparison, Table 2 shows groupings as produced using the correlation grouping method. Remember, this method is more subjective, in that the analyst chooses the number of related words for each group.

These groupings (Table 2) tend to be somewhat different from the groupings produced using the clustering method. When looking at the actual text streams that produced these results (from VofA), we can see a good explanation for most of the groupings, including:

1. Lee Hamilton and James Baker were co-chairmen of the Iraq Study Group. The text discussed strategy in Iraq (war) not working (grave and deteriorating), along with the relationship with neighboring countries.

2. Robert Gates' nomination as the new U.S. Defense Secretary.

3. Group is loosely related, several topics involved.

4. Group is also loosely related.

5. Powerful Iraqi Shi'ite clerics and their connection with corruption.

6. Discussions (Chinese) about sanctions for nuclear weapons in North Korea.

7. Assistant Secretary of State Christopher Hill discusses nuclear weapons and North Korea.

4.2 Time Intervals. So far, we have presented several algorithms for detecting surprising events within a text stream. We have shown results based on a single corpora (Voice of America, 11/22/2006 – 12/20/2006). The time span of this data is relatively short (about a month). What happens if we have data over a long period of time? For instance, PubMed is a service of the U.S. National Library of Medicine that includes over 18 million citations from MEDLINE and other life science journals for biomedical articles dating back to the 1950s [12]. PubMed includes links to full-text articles and other related resources. Data from MEDLINE are presented in Figures 13 through 15 and discussed here. In Figure 13, we see the number of documents that are available for each time interval, divided into 3-month intervals. The results of applying our Surprise technology for this data set are shown in Figure 14, with the profiles of the top 30 words shown in Figure 15.

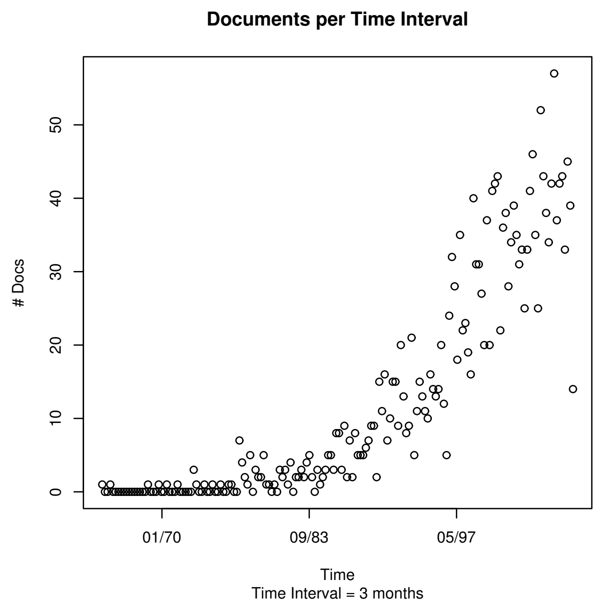

Figure 13: Number of documents per time interval for the MEDLINE data set.

For more information about these words and events, grouping analysis was performed. A list of the top 25 groups (based on the Surprise statistic) is shown in Table 3.

Now, what if we change the interval resolution, say from 3 months (365.25/4) to monthly (362.25/12)? The temporal profiles for this new analysis for the top 30 (surprising) words are shown in Figure 16. These words were then grouped and are shown in Table 4.

Comparing the top 30 surprising words from Figure 15 (3-month resolution) to the top 30 surprising words from Figure 16 (1-month resolution), we observe only a few overlapping words (encoding, iga, vaccination/vaccinated, expressing, vector). Also, the groupings from the two analyses (Tables 3 and 4) are totally different. The results of this comparison show that the

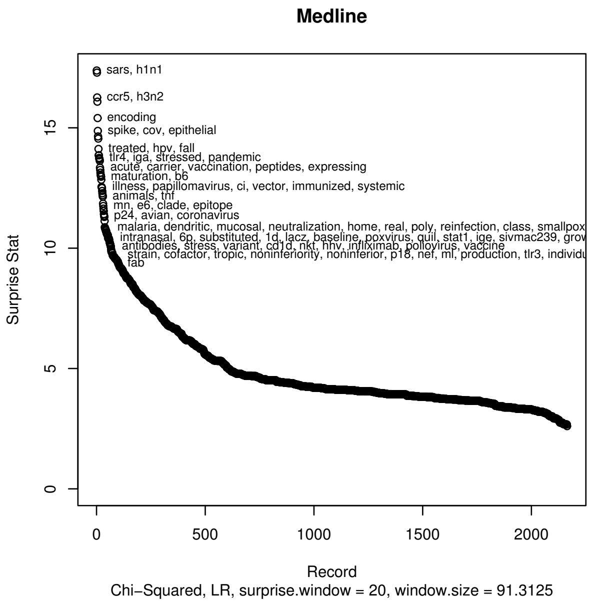

Figure 14: Sorted Surprise statistics, top 50 surprising terms labeled (MEDLINE data set).

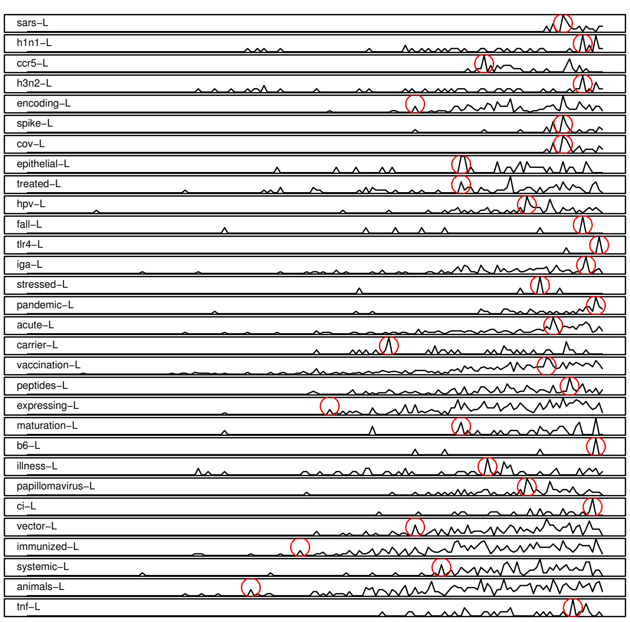

Figure 15: Temporal profiles for the top 30 surprising words (MEDLINE data set), 3-month time resolution.

Table 3: Top 25 surprising word groupings using the clustering method (MEDLINE data set), 3-month time resolution.

1	h1n1 h3n2 season
2	ccr5 genotype delta32
3	stressed corticosterone restraint transmammary
4	fall noninferiority noninferior hbsab
5	tlr4 tlr3 livers
6	tnf alpha regression
7	b6 hsp65
8	substituted 1d insoluble
9	gmc b16f10 mac phosphatase unmanipulated
10	malaria falciparum
11	hpv papillomavirus e6 e7 cervical
12	sars spike cov coronavirus syndrome poxviruses m2e copv trimera
13	vagina oncoproteins
14	subtype subtypes
15	illness gastroenteritis
16	pigs guinea
17	p24 gz
18	sivmac239 postchallenge
19	clade neurocognitive
20	baseline cd3 golden
21	sign ebola
22	ci impairs
23	evetpirn aa6 agms hrsv pseudovirus
24	poly prototypical
25	h5 asthma h7 ae

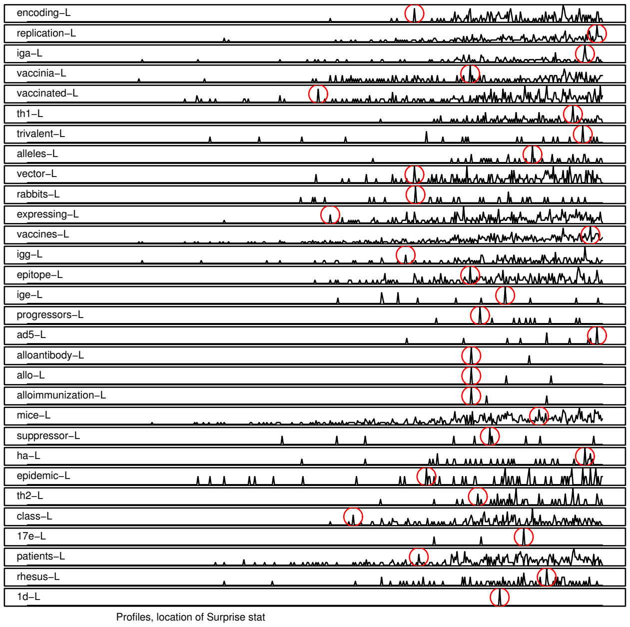

Figure 16: Temporal profiles for the top 30 surprising terms (MEDLINE data set), monthly time resolution.

Table 4: Top 25 surprising word groupings for the MEDLINE data set and monthly time resolution.

iga igg
th1 th2
encoding plasmid
rsv syncytial
alloantibody allo alloimmunization alloimmune
trivalent kk10
h1n1 h3n2
replication defective
herpes simplex hsv gd
vector vectors
vaccinia expressing
ad5 icp8
noninferiority noninferior cryptic
alleles hla
agonist ampligen vietnam polyfunctional mac251
rabbits rabbit
1d copolymer
west nile
hbv hbsag hbs
adjuvant adjuvants
infliximab preemptive
protracted b57
17e deltab670 cl
sendai streptococcal
hhv sivgag

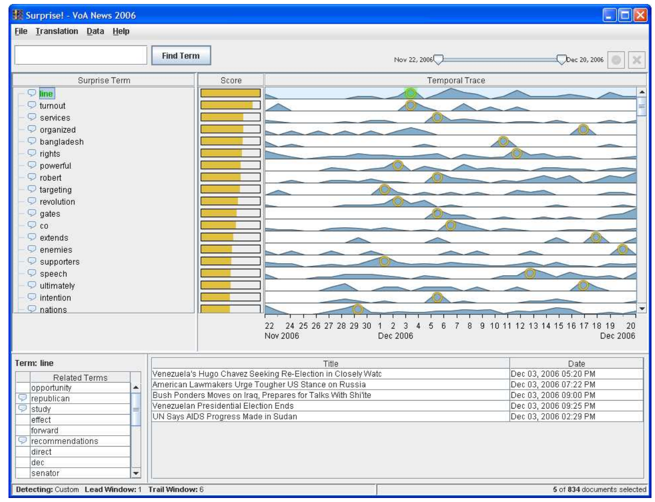

Figure 17: GUI for the Surprise software.

long time periods and differing time resolution can play a key factor in the analysis of these text streams, at least when looking for change/surprising events.

4.3 Surprise Software. In order to make the Surprise analysis viable by an analyst, a graphical user-interface (GUI) called Surprise has been developed. The GUI for Surprise is illustrated in Figure 17. Surprise allows an analyst to explore the term list and temporal profiles of the terms with the highest Surprise statistical values. The visualization and interactions are designed to help an analyst identify terms indicating an event and quickly understand the associated context.

An analyst can quickly identify events using the Surprise visualization. Interactions with the visualization allows a user to discover context about an event and select relevant documents. An analyst can sort the term list alphabetically, by surprise score or by correlation order. When a term is select, the related terms computed from the algorithm described in Section 2.1 are populated in the lower left corner of the GUI. A portion of the temporal plots can be highlighted to select documents for that time range.

Surprise is linked with In-Spire to support further analysis. When documents are selected in Surprise, a user can view those documents in In-Spire and explore their relationship in the topical space. The actual document names for which the terms occurred are shown in the bottom right window, illustrated in Figure 17.

5 Summary

In this paper, we have described our algorithmic development in the area of detecting evolving content in text streams. We have compared our results to text analysis results on a static document collection and found that our techniques produce results that are different and enhance those results. We compared our two different grouping algorithms and concluded that the correlation grouping method tended to produce groups that were more representative and informative of the data sets we analyzed.

We also looked at the effects of using our techniques on a data set that spans decades. Although words were found to be "surprising," the grouping algorithms tended to not produce information (groups) as informative as those produced using the VofA data set. We expect this is due primarily to the longer time span of the text streams and the difference in the subject matter (e.g., medical data vs. current events).

We recognize that much work is needed in this area before this technology can become a real-time analysis tool. Right now, the analyst must make too many decisions (trials) before the relevant information appears. With this in mind, this paper presents research on the first step in the development of capabilities for real-time analysis of text streams.

References

[1] Q He, K Chang, E Lim, J Zhang, *Bursty feature representation for clustering text streams*, proceedings of SIAM Data Mining Conference, 2007, http://siam.org/proceedings/datamining/2007/dm07_050he.pdf.

[2] Q Mei and C Zhai, *Discovering evolutionary theme patterns from text: an exploration of temporal text mining*, in SIGKDD, pp. 198-207, 2005.

[3] G Kumaran and J Allan, *Text classification and named entities for new event detection*, in SIGIR, pp. 297-304, 2004.

[4] RL Eubank and PD Whitney, *Convergence rates for estimation in certain partially linear models*, J. of Statistical Planning Inference, 23, 33-43, 1989.

[5] DW Engel, PD Whitney, AJ Calapristi and FJ Brockman, *Mining for Emerging technologies within text streams and documents*, proceedings of SIAM Data Mining Conference (Text Mining), 2009.

[6] Voice of America, http://VOANews.com.

[7] A Agresti, *Categorical data analysis*, 2^{nd} ed. John Wiley & Sons, New York, 2002.

[8] JL Fleiss, *Statistical methods for rates and proportions*, 2^{nd} ed. John Wiley & Sons, New York, 1981.

[9] T Dunning, *Accurate methods for the statistics of surprise and coincidence*, Assoc. for Computational Linguistics, 19:1, 1993.

[10] L Kaufman and PJ Rousseeuw, *Finding groups in data: an introduction to cluster analysis*, Wiley, New York, 1990.

[11] In-Spire, Pacific Northwest National Laboratory, http://in-spire.pnl.gov.

[12] NIH National Library of Medicine (NLM) PubMed MEDLINE database. http://nlm.nih.gov/pubs/factsheets/medline.html.