Dear Paper Chairs,

We greatly appreciated this opportunity to update our paper submission, entitled “WordStream: Interactive Visualization for Topic Evolution”. We have addressed the reviewers concerns, and believe that it is now a worthy short paper submission for the EuroVis conference. Here, we provide a summary of all of the main changes we have made to the paper:

- We included a supplementary video that shows the main functionality of and user interaction with the WordStream visualization (R3). The video is now available on our project website: https://idatavisualizationlab.github.io/WordStream/video.html
- We added an argument into our Introduction section. Even though the combination of the word cloud and stream graph has been presented before (R1, R2), there are still a lot of rooms for optimization, especially when the topic streams highly fluctuate. In fact, TIARA only shows a few word clouds into the stream boxes where space is sufficient. Therefore, maximizing the space usage within the stream layers was not the priority of this technique. Similarly, TextFlow and EvoRiver focus on the topic evolution and their critical events such as birth, death, split, merge but not space usage. In particular, a word cloud at a particular timestamp can be only displayed on request. This is also the main reason why we did not directly compare WordStream with earlier attempts (R1, R2).
- We have removed the G2 and re-labeled other goals since we agree with the argument from (R2,R3) that G2 has not been reached. WordStream was designed to communicate the global trends and highlight important topics rather than focusing on the evolution of individual terms (G2). G2 can be partially satisfied through user interactions as demonstrated in the supplementary video. In this regard, we also added a new figure (Figure 7) to show examples of term selection.
- We have provided more details of the "sudden attention technique" and a citation for the entity recognition technique that we used to classify the texts into people names, locations, and organizations: https://explosion.ai/demos/displacy-ent
- The revised paper has been thoroughly checked for grammatical errors and typos (R2,R3). Many sections and paragraphs have been rewritten/retouched to make the paper flows better.
- Finally, the paper now includes a reference to the WordStream project home page, located at https://idatavisualizationlab.github.io/WordStream/, which includes all source code and other examples of our WordStream visualization;

We also provide in-line commentary that responds in further detail to each of the individual reviewers’ comments (we use blue text to indicate our reply to the main issues). The detailed report is available on our Github website:
https://github.com/iDataVisualizationLab/WordStream/blob/master/EuroVis2019Reviews.pdf

Again, we thank the reviewers for their attention to detail and their very helpful comments. We hope to see you in Porto, Portugal!

Sincerely,

Tommy Dang, Huyen Nguyen, and Vung Pham
****

coordinator review
score **3**/5

Summary Review Text

This is a borderline paper with the reviewers leaning more towards “weak accept” than “weak reject”. Should the paper be conditionally accepted, the issues raised under “Weaknesses” should be addressed in a revision.

Strengths:
- A promising tool that combines word clouds and stream graphs (R1,R2)
- The paper is generally well written and has a good review of related work (R1)

Weaknesses:
- The combination of word cloud and stream graph has been presented before (see intro), and the paper does not explicitly describe a novel better layout algorithm, it is unclear what exactly the contribution is (R1, R2)
- No comparison with earlier attempts to combine word clouds and stream graphs (R1, R2)
- It seems that design goal G2 has not been reached: The method displays the most important terms at the times when they are most prominent, but does not show the temporal evolution of these terms. Neither it is easily possible to observe whether terms disappeared or appeared over time and when. (R2,R3)
- Missing implementation details, either the novel algorithm or the exact stating and reference to existing algorithms (R1, R2)
- Missing any explanation of interactions with the system; a video showing the system in action will be very helpful (R3)
- Provide definitions, justification, and explanation of use for "sudden attention technique" and TF/IDF (R2,R3)
- Some grammatical errors and typos (R2,R3)

Summary Rating

Borderline: the strengths and weaknesses balance for this paper.

Paper Type

System

Rating

Borderline: the strengths and weaknesses balance for this paper.

Expertise

Knowledgeable

The Review

This paper describes a wordcloud visualization with an additional time component. It is a combination of wordle and streamgraph as acknolwedged by the authors.

Further this combination itself is also not new -- quoting from the paper "The combination of word cloud and stacked graph has been seen in several studies."

Still, the system seems useful and slick. I would argue for acceptance if the authors provide more technical information (as detailed below) and carefully revise the final version with a native speaker at hand.

Major issues:

It is unclear how the set of topics is extracted. For the VIS literature it seems that the "topics" are given (papers in SciVis, VAST, etc.) but is the partition/clustering into topics assumed to be part of the input? This would make the system more difficult to apply to text that is not already pre-clustered.

In Section 4 (first paragraph), we have cited the entity recognition technique that we used to classify the texts into people names, locations, and organizations: https://explosion.ai/demos/displacy-ent

It is also unclear whether there's a control over the repetition of words/phrases. Is there a way to control this? In Fig1 there's a noticeable absence of "Obama" -- is this an effect of the "sudden attention technique"?

Yes, this is the effect of the "sudden attention technique". We also added more details on this term filtering technique

Explain what the "sudden attention technique" does. Why is this chosen over TF or TF/IDF?

We have provided the details on how to compute sudden attendtion of terms in Section 4 (first paragraph).

Why is TF/IDF used as a quality measure?

We have removed this measure. Compactness is a more suitable quality measure in this context.

How is the TF/IDF quality measure adapted to topics that change over time?

We have removed this measure.

The "Informal User Study" section does not add much; it would be much better to provide more details about how the system actually works.

We incorporated this suggestion. In particular, we reduced the text in this Section and add a new Figure 7 to show how interaction works

Grammatical errors and typos (using examples mostly from the first page):

We completely revise the Introduction section as well as the entire paper.

Word cloud is designed to give an engaging visualization of texts based on font

sizes and colors while stacked graph is a popular method for visualizing topic evolution.
-->
Word clouds are designed to give an engaging visualization of text based on font sizes and colors, while stacked graphs are a popular method for visualizing topic evolution.
Fixed. Thanks,

The words are applied with spatial orientation
-->
In our visualization words have spatial orientation
This sentence has been removed.

forming the actual stream layer and increasing coverage over the stream
-- this is doesn't make any sense
This sentence has been removed.

The illustration of topic evolutions has existed for a long time.
-- reword
This sentence has been reworded

In recent researches, there are several effective techniques for the demonstration of the progression of topics.
-- reword
This sentence has been shorten and combined to the next one.

the changes in the widths of each stacked layer along this timeline depicting the amount of evolution of the topic
-- this doesn't make sense
This sentence has been replaced

On the other hand, the words themselves demand to be visualized in such a way that they are memorable and users can quickly capture their meaning as a whole.
-- reword!
-- words don't demand anything...
-- the meaning of words "as a whole": what do you mean?
This sentence has been removed.

Wordle [VWF09] is a useful tool for emphasizing essential words by utilizing the randomized greedy algorithm to place words within two-dimensional space
-- this is a poor introduction of wordle; what are "essential words"? what is "the randomized greedy algorithm"?
-- reword

This sentence has been removed.

"The benefit from these word cloud models for learning and memory can lead to considerable potential for embedding in a larger model."
-- this sentence is just a bag of words, without structure or meaning!
-- what "embedding"? what "larger model"?

The Introduction section has been rewritten.

"The second challenge is how to demonstrate the comparison between topics and their corresponding terms."
-- what does that mean?

This sentence has been removed.

"G1. Display the evolution of stream topics"
-- in the teaser figure you have the same 4 topics through the entire time...

"G3. Emphasis important terms in their corresponding topic layers at their corresponding time steps."
-- emphasize

Fixed. Thanks

Goals 3 and 4 can be conflicting

We disagree with this. We can rescale the font sizes to fit more terms (G4) without de-emphasize the important terms (based on their relative font sizes)

"Arrange texts according to the orientation of its stream."
-- all words are horizontal; what is "orientation" here?

We have clarified this sentence.

"Less level of freedom can present more words."
-- reword

We have removed this sentence

The entire paper needs to be carefully proof-read by a native speaker.

---

committee member review
score **3**/5
Paper Type
Algorithm / Technique
Rating
Borderline: the strengths and weaknesses balance for this paper.
Expertise
Passing Knowledge

The Review

The paper presents Wordstream - a visualization combining word cloud and stream graph. The presented approach has several layout variants. The paper is very well written and the results look good. Related work is cited. As the combination of word cloud and stream graph has been presented before (see intro), and the paper does not explicitely describe a novel better layout algorithm, it is unclear what exactly the contribution is.

Introduction states that there are "technical challenges" to the current approaches. It is unclear which exactly? How does the presented approach overcome them? How does evaluation show this improvement?

The approach section describes how the visualization is constructed. Unfortunately the implementation details, either the novel algorithm or the exact stating and referencing to existing algorithms is missing. In case there is no space, this could be included in annex. But it is necessary to understand and reproduce the approach. Moreover, it would enable to state clearly where the contribution in the algorithm is. For example, I like the various variants, but it is unclear what their contribution, pros and contras are.

The evaluation shows impressive set of use cases, but not the comparison to state of the art or the value added of the various layout variants.

This is a pitty as the paper is well written and the results look appealing. In case the paper is accepted, I require to include the necessary technical details. In case the paper is rejected, I would suggest revision and resubmission to a different venue.

---

committee member review

score **4**/5

Paper Type

Algorithm / Technique

Rating

Probably accept: I would argue for accepting this paper.

Expertise

Expert

The Review

The paper presents an approach to visualize topic evolution. It is based on combining a word cloud with a stream graph approach. Thus, the streamgraph allows for an intuitive understanding how the topics evolve over time (at least, when assuming a low number of topics), while the word cloud shows the most prominent terms at the time of their appearance. This is not ground-breaking work, but a

nice extension of previous techniques and, thus, can be considered suitable for a short paper.

My main concern is whether design goal G2 has been reached. The method displays the most important terms at the times when they are most prominent, but they do not exhibit the temporal evolution of these terms. Neither it is easily possible to observe whether terms disappeared or appeared over time and when. Hence, I woudl argue that design goal G2 has not been reached.

We agree with this argument. I tried to demonstrate the term evolution via interaction. Unfortunately, this has been cut off for the short paper submission. We believe this is not the focus of our technique since we try to communicate the global trends rather than individual terms. Therefore, we removed G2 in our revision.

In Section 5.3, it is mentioned that a user "chose a word and scrolled through the entire timeline to see the fluctuation of its occurrences". This type of interaction is not explained in the paper. In general, the paper talks about an interactive system, but the interaction mechanisms are never described. Also, an accompanying video would have been nice to document possible interactions (as an anonymous reviewer I am not clicking at provided links). It would actually be nice to "see the fluctuation of its occurrences", which relates to goal G2. But then, how is this done?

We have provided the video on our project website to demonstrate this feature through user interactions.

I am not surprised of the outcome of the evaluation in Section 5.1. In Figure 5, it is obvious that the "no flow, no angle variance" allows for best packing. It also allows for best readability and even for best following of the streamgraph in my opinion.

We agree. This result did not fit our initial hypothesis but make sense as we investigate in details.

Minor comments:

* The last sentence if Section 3 is not a sentence.

We itemized the two options: Flow and Angle variace. Therefore, this fixed the issue

* Terms like TD-IDF should be explained in the paper. In general, the paper assumes that the reader is familiar with topic modeling, which is not necessarily the case for all EuroVis attendees.
