Distance Aware Tag Clouds

Iris Adä, Kilian Thiel, Michael R. Berthold
Nycomed Chair for Bioinformatics and Information Mining
University Konstanz
78457 Konstanz
{Iris.Adae, Kilian.Thiel, Michael.Berthold}@Uni-Konstanz.de

***Abstract*—Distance aware tag clouds add visualization of relations between terms to standard tag clouds. In addition to term importance (which is usually depicted through font size) the placement of terms represents the relation between words in the corpus. These relations are modeled as similarities between words and are visualized via the distance between the corresponding tags in the tag cloud. In this paper a modified multi-dimensional scaling (MDS) approach for tag positioning is presented. Applying standard MDS results in unsatisfying and unusable representations due to two problems. The first problem stems from word overlap which is orthogonal to the second problem: excessive empty space. Both of these shortcomings are addressed by introducing methods for overlap removal and empty space reduction. We show that these two modifications only moderately increase the resulting MDS stress value of the new positioning while they remove most of the overlaps and reduce the amount of white space.**

***Index Terms*—distance aware tag cloud, multidimensional scaling, term visualization, overlap removal.**

I. Introduction

The basic purpose of a tag cloud is the presentation of a visual overview of a text collection. Regular tag clouds, which can be seen on various websites, represent the frequency or importance of a word or term related to a text collection via the size of the term. In most cases the terms are positioned as an alphabetical list or in a compact way to reduce unused white space in order to save space. Usually the positioning of the terms is not used to display additional information, such as the relations between terms. By taking these relations into account concerning the positioning of terms the expressiveness of a tag cloud visualization can be increased. Furthermore the visibility of terms for the viewer can be increased by certain positionings. As Bate et al. [1] evaluated, terms displayed in the first or in the last line of a tag cloud are often overlooked, especially if those terms are displayed with a smaller font size since they are less important or less frequent. Still less important terms can be meaningful and significant in combination with other terms e.g. to dissolve ambiguities, clarify meanings, or point out relations. Additionally the users attention can be drawn to certain terms.

In this paper we introduce a positioning method for terms in tag clouds that takes the relations between terms into account. These terms, extracted from a document set, are weighted based on their frequencies in the documents and positioned according to their relatedness. Relatedness is based on the co-occurrences of terms but, of course, other types of similarities or distance measures can be used as well.

The aim of this visualization is split into four objectives:

1) displaying the importance of a term,
2) visual representation of relatedness (similarity),
3) compactness and
4) readability.

The importance of a term is commonly visualized by its font size in standard tag clouds. Therefore a measure of importance or weight is needed for each term, which is mapped to a range of font sizes. Mostly the frequency of a term in a certain document or corpus is used as weight. We propose the visualization of term relations (retrieved from similarities in the original term space) via the Euclidean distances in the two-dimensional projection plane of the tag cloud.

Finally the last two objectives are usually provided by standard tag clouds but will become more important here. They are important to increase the usability and aesthetics of the tag cloud. The third requirement relates to the compactness of the tag cloud. The amount of white space should be reduced to a reasonable proportion. This requirement is indirectly involved in the last objective. The selected and displayed terms should be clearly readable, whereas an overlapping of terms needs to be avoided or at least reduced to a minimum.

The paper is organized as follows. In the next section an overview of related work concerning tag cloud visualization techniques is given. In Section III the used data, preprocessing, term extraction as well as the term distance computation is described. In Section IV first a short overview of the MDS method used is given and secondly the applied improvements of term positioning are explained in order to fulfill the constraints mentioned above. A visual evaluation is presented alongside the visualization section and a quantitative of our main example in the evaluation section. The paper is concluded with an overview.

II. Related Work

According to Viégas et al. [2], one of the first examples of tag clouds may have been created by Milgram et al. in 1976 [3]. In an experiment people were asked to name landmarks in Paris. A collective map was created based on these landmarks, using font size to visualize how often each place was mentioned.

In 2002 the website Flickr used tag clouds to show the popularity of various tags [2]. For positioning, the tags or terms are sorted alphabetically and displayed line by line. The more popular or important a tag is, the bigger the font size.

978-1-4244-6587-3/10/$25.00 É2010 IEEE

This simple layout technique, is often applied to tag clouds. The alphabetically sorted tables and lists are so-called "tag clouds of the first generation". The key benefits of these tags clouds are their low computational complexity. Sorting terms alphabetically as well as displaying them in a table is time-efficient. The drawback of these first generation approaches is that the possibility to apply positions to the terms of the tag cloud is not used. By using positioning methods to increase the expressiveness of the tag cloud, one can also add new information, such as the relation of the displayed tags, with regard to semantic similarity, to the visualization. Figure 1 depicts a tag cloud, showing important terms of the inaugurals of the last three presidents of the USA, visualized as an alphabetically sorted list.

Fig. 1. A first generation tag cloud visualization of important terms of inaugurals of the last three presidents of the USA.

In second generation tag clouds, the positioning of terms is much more important. In Seifert et al. [4] words are placed in a polygon. The biggest ones are placed in the middle and the others are placed around the middle, with the aim to minimize the empty space and to avoid overlapping. John Feinbergs' Wordle[1] is considered to be a second generation tag cloud as well. These tag clouds overcome on the one hand the drawback of alphabetical ordering [5], but on the other hand are still static and relations between words are not visualized. Figure 2 shows a Wordle of the most important terms of the inaugurals. It can be seen that the most frequent terms are placed in the middle and thereby attract the attention of the viewer faster than in Figure 1.

To visualize the semantic relations of terms, usually based on their co-occurrence in documents, multi-dimensional scaling (MDS) methods have been applied to position terms in

[1]available at `http://www.wordle.net`

Fig. 2. A Wordle visualization of important terms of inaugurals of the last three presidents of the USA.

tag clouds, e.g. by Martín-Merino et al. [6]. However when using MDS, problems like 0-distances leading to an overlap of terms or the *curse of dimensionality* can affect the visualization negatively. Semi-supervised approaches ([7], [8]) have been applied to overcome the latter and to improve the discriminant power of tag clouds by taking advantage of a priori categorized documents.

A non static approach that visualizes terms over time can be seen in an online project [2]. The most frequent terms, after removing common words, of the speeches of the presidents of the United States are visualized by an alphabetically sorted table. The more frequent a term, the bigger its font size. This approach enables one to notice topic changes over time and the different presidents. However, it is cognitively exhausting to follow particular terms, since they may appear at different positions.

Just recently, a new word cloud visualization was proposed by Cui et al. [9], combining the visualization of a corpus evolving over time and the visualization of term relations via MDS positioning. The temporal evolution of content in the corpus is represented as the evolution of relations between important terms, thereby changing the positions of terms in the tag cloud.

III. Data and preprocessing

In this section, first, data used for the visualizations is described briefly. Afterwards the used document preprocessing is introduced, and terms selection criterias and distance measurement are described.

A. Data

Each data set consists of a set of document D, from which the required terms T are selected and the distances between each pair is calculated. For testing our visualization approach we used three different corpora. At first articles from the *Times* document collection[3], which contains 425 articles from the time magazine in the year 1963. As a second approach the *PubMed*[4] database was queried for the terms *breast*, *mice* and *cancer*. And finally we compared the inaugurals of the last three presidents of the United States[5].

B. Preprocessing

After parsing and creating a bag-of-words representation for the selected documents, we applied typical text mining preprocessing routines. All numbers, stop words, very short words and non-nouns were filtered and the remaining terms were stemmed.

Since a tag cloud typically contains between 30 and 150 terms the number of remaining terms still needs to be reduced. Too many terms can confuse the observer, are cognitively exhausting, and make it difficult to have an easy and quick overview of the important terms. To further reduce the total

[2]available at `http://chir.ag/projects/preztags/`
[3]available at `http://www.cs.utk.edu/~lsi/corpa.html`
[4]available at `http://www.ncbi.nlm.nih.gov/pubmed`
[5]available at `http://www.bartleby.com/124/`

number of terms, a weight for each term was determined, representing its importance. Based on this weight, the most important terms are visualized in the tag cloud.

The weight or importance $\mathrm{i}(t) = \sum_{d \in D} w_d(t)$ of a term t in a document corpus D is determined by cumulating its tfidf[6] values $w_d(t)$ for each document d of the corpus. The weight $\mathrm{i}(t)$ is used in the visualization to determine the font size.

After term extraction and the computation of term weights, distances between all pairs of terms are computed. Each term is represented as a term vector in document space $\mathbf{t} \in \mathbb{R}^{|D|}$, using the tfidf values as vector values at the corresponding indices $(\mathbf{t})_d = w_d(t)$. Based on these term vectors, the cosine distance $\mathrm{d}(t_i, t_j) = 1 - \cos(\mathbf{t_i}, \mathbf{t_j})$ between two terms t_i and t_j is computed. The cosine measure results in small distances for terms occurring in the same proportion in all considered documents. The set of the selected terms is described with T.

The complete preprocessing of the document collection was done with KNIME [10]. KNIME is a free data mining tool, which can be used to analyze and visualize huge amounts of data. For our approach we used the text-processing plugin[7], which already provides quite a lot of text mining routines.

IV. Visualization

In this section we introduce the positioning of tags in a tag cloud. The well-known MDS method to map the term distances onto the two-dimensional drawing pane is introduced followed by our improvements of the tag positioning. However, first the determination of font sizes is briefly discussed.

A. Font Size

As already mentioned, the font sizes represent the importance $\mathrm{i}(t)$ of a term. Using a logarithmic adaptation yields good results, as the existence of few very large values is smoothed and the differences between smaller values become more visible [4]. In this work, the view is always determined by a preconfigured minimum fs_{min} and maximum fs_{max} font size. In our experiments we always choose a range from a minimal font size of 10pt to a maximal font size of 50pt for the most important terms. The actual font size $\mathrm{fs}(t)$ used to display the term t is determined by

$$\mathrm{fs}(t) = \mathrm{fs}_{min} + \frac{\ln(\mathrm{i}(t)) - \ln(\mathrm{i}_{min})}{\ln(\mathrm{i}_{max}) - \ln(\mathrm{i}_{min})} * (\mathrm{fs}_{max} - \mathrm{fs}_{min}) \ ,$$

where $\mathrm{i}_{min} = \min\{\mathrm{i}(t) | t \in T\}$ and $\mathrm{i}_{max} = \max\{\mathrm{i}(t) | t \in T\}$ are the minimum and maximum importance values of all displayed terms, respectively.

B. Term Placement

After calculating the font sizes, the high-dimensional term vectors are mapped via MDS onto two dimensions. For each term, a two-dimensional point is determined, which is used as the center of the rectangle in which the corresponding term is drawn. Note that the size of the rectangle is determined by the number of characters of the related term and its font size.

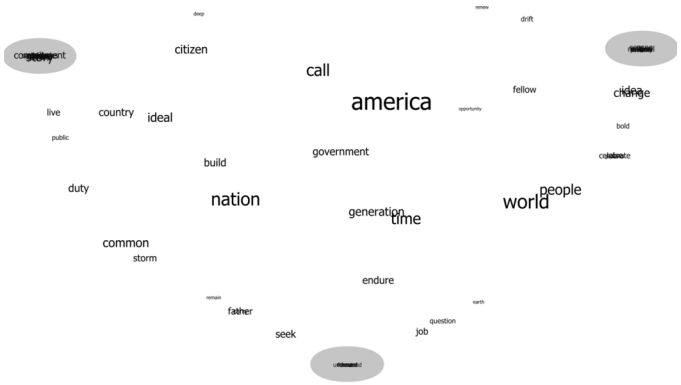

Fig. 3. Pure MDS tag cloud without removal of overlaps. The tag cloud compares the 30 most important terms of each inaugural of the last three American presidents. Terms of the three dense clusters (highlighted with gray circles) only occur in one of the speeches. Due to the huge amount of overlap many terms of this tag cloud are not readable.

Multi-dimensional scaling is used to map objects of a high-dimensional space onto a lower-dimensional space by trying to preserve the pairwise distances between objects ([11], [12]). The high-dimensional space here is the document space in which the cosine distances between the term vectors is determined. The lower-dimensional space is the two-dimensional drawing area in which the Euclidean distance metric is used.

The aim of multi-dimensional scaling is to find a positioning for the high-dimensional points in the lower-dimensional space in a way such that distances, here cosine, are still preserved for each pair of points. To achieve this dimension reduction, a stress minimization is applied. The method used here was suggested by Sammons [13] and tries to find a minimum for the stress function E

$$\mathrm{E} = \frac{1}{\sum_{t_i \neq t_j} \mathrm{d}(t_i, t_j)} \sum_{t_i \neq t_j} \frac{1}{\mathrm{d}(t_i, t_j)} \left(\mathrm{d}(t_i, t_j) - \delta(t_i, t_j)\right)^2 \ ,$$

where $\delta(t_i, t_j)$ is the euclidean distance between the terms t_i and t_j, i.e. the distance between the centers of their bounding boxes. After using a random initialization for the points in every iteration, the stress is decreased by gradient descent. A detailed description of this method can be found in [14].

Figure 3 shows the positioning of terms of the three inaugurals determined by the MDS result. In the visualization, the distance of the selected terms of the three inaugurals is visualized. Terms placed next to each other have been used frequently in the same speeches. Unfortunately, a lot of terms overlap and are thereby hardly or not at all readable. Most notably, terms being considered important in only one of the speeches overlap more than others. This results from a 0-distance of these terms in document space and hence they are mapped via MDS to the same center point and are displayed as an unreadable batch.

The visualization problems, based on the mapping of term distances in document space onto a two-dimensional projection

[6]Term frequency - inverse document frequency is a weighting of terms in a document corpus, calculated by multiplying the relative term frequency with the inverse document frequency.

[7]available at `http://labs.knime.org/textprocessing`

plane, can be broken down to two main issues. These issues are identified considering the aim of getting a compact but still readable tag cloud view. The first problem yields from terms with 0-distances, resulting from an exclusive co-occurrence in all documents they occur, leading to a cosine similarity of 1. The perfect positioning would be to display them exactly on top of each other, which results in an unreadable visualization. The second problem is the amount of white space in the tag cloud. The main reason for this effect is the distribution of distance values. A few large distances in comparison to many small distances leads to a distortion of the tag cloud and a huge amount of white space, since the MDS is trying to preserve these large distances. However, the exact representation of these large distances is not necessary. Instead it is important to give the viewer an impression of which terms are strongly or weakly related.

In the following we apply two modifications, resulting in a more readable and aesthetic view. The amount of white space in proportion to the total size and the amount of overlapping areas is reduced. Additionally we suggest the usage of an automatically-determined scale factor to fit the selected font size range and to minimize the overlap. The overall stress of the positioning is thereby increased, but only by a marginal amount as you can see in the evaluation section, where we compare the stress, overlapping and white space before and after our distance harmonization approach.

C. Distance harmonization

The aim of distance harmonization is to adjust the Euclidean distances in the tag cloud obtained from standard MDS, in order to improve the readability and aesthetics of the visualization, with respect to the objectives listed in the introduction of this work.

The size $\mathrm{size}(t_i)$ of a term rectangle is determined by the font size, calculated by its weight and the characters of the corresponding term $t_i \in T$. The pre-fixed size of the rectangles leads to a distortion of the two-dimensional mapping, which can be avoided by appropriate distance scaling. Therefore the minimal space required to display all term rectangles is determined by cumulating the area of all rectangles. This is the absolute minimum size needed to display all terms.

We assume that scaling the mean of all distances by the square root of the minimal required space is sufficient to deskew the view. Hence, the scaling factor is defined by

$$\mathrm{mds}^d_{\text{scaling}} = \sqrt{\left(\sum_{t \in T} \mathrm{size}(t)\right)} \Big/ \frac{\sum_{t_i \neq t_j} \mathrm{d}(t_i, t_j)}{m \cdot (m-1) \cdot \frac{1}{2}}$$

with respect to the distance function $\mathrm{d}(\cdot, \cdot)$.

Scaling distances does not have any impact on distances equal to zero. As they are the main reason for term overlap, a further adjustment of the distances is applied. To reduce the effect of 0-distances, all distances are slightly increased by $\mathrm{mds}^d_{\text{addon}}$. Only the distance between a term and itself $\mathrm{d}(t, t))$ is not increased, to keep the metric characteristics. The minimal distance greater than 0 is used for $\mathrm{mds}^d_{\text{addon}}$ in this work.

$$\mathrm{mds}^d_{\text{addon}} = \min \{\mathrm{d}(t_i, t_j) > 0 | t_i, t_j \in T\}$$

The next distance modification aims to reduce the effect of large differences between distances. We identified two problems at the extremes when investigating distances in relation to the unused white space: first the existence of mainly small and only a few large distances and second the opposite; many large and only few small distances. For the first problem, a logarithmic function is applied to the distances, separating the small and dampening the few large distances. The opposite problem can be improved by applying an exponential function.

To identify which of these problems occur, the ratio of small and large distances is determined based on the mean Ω of all distances.

$$\Omega = \frac{1}{|T|^2} \sum_{t_i \neq t_j} \mathrm{d}(t_i, t_j)$$

Additionally, the decision criteria Ψ is calculated as the proportion between the number of values greater and number of values smaller than Ω.

$$\Psi = \frac{|\{\{t_i, t_j\} \,|\, \mathrm{d}(t_i, t_j) < \Omega, t_i, t_j \in T\}|}{|\{\{t_i, t_j\} \,|\, \mathrm{d}(t_i, t_j) \geq \Omega, t_i, t_j \in T\}|}$$

In total the distance harmonization can be described using the following equation:

$$\overline{\mathrm{d}(t_i, t_j)} = \begin{cases} \ln\left(\mathrm{d}(t_i, t_j) + \mathrm{mds}^d_{\text{addon}} + 1\right) & \text{, if } \Psi > \frac{1}{\tau} \\ \exp\left(\mathrm{d}(t_i, t_j) + \mathrm{mds}^d_{\text{addon}}\right) & \text{, if } \Psi < \tau \\ \mathrm{d}(t_i, t_j) + \mathrm{mds}^d_{\text{addon}} & \text{, else} \end{cases}$$

Results have shown that values between 0.7 and 0.9 for τ yield good visualizations. Finally the positions of the term centers are calculated using the same multi-dimensional scaling methods as described in Section IV-B on the modified distances.

Figure 4 depicts two tag clouds created from two different corpora (left, right column) using three different approaches of distance harmonization. An MDS with scaled distances is applied on the top. It is obvious that this visualization is not helpful due to the amount of overlapping terms. In the middle only offset $\mathrm{mds}^d_{\text{addon}}$ and scaling is used, which already increases the usability of the view. At the bottom the logarithmic (right) and the exponential (left) adjustment is additionally used.

D. Minimizing remaining overlapping

The distance harmonization already yields proper results for the document sets it was tested on. But sometimes overlapping of terms still poses problems. To reduce the remaining overlap, the force transfer algorithm (FTA) [15], which was constructed to remove node-node-overlapping in graph visualizations, is applied. The problem of overlapping between terms in a tag cloud is basically a problem of overlapping between rectangles.

Briefly the FTA works as follows: In each iteration, clusters are determined, where a cluster is described by a set of rectangles overlapping each other. The overlap in this cluster is than reduced starting from a core point. Huang et al. [16]

Fig. 4. MDS tag clouds of two document sets (left and right column) with different steps of distance harmonization. *top*: MDS with scaling, *center*: additionally with $\text{mds}^{d}_{\text{addon}}$, *bottom*: complete harmonization. On the left hand side, data obtained from the Times Corpus is shown with a logarithmic adjustment applied. Data obtained from Pubmed with an exponential adjustment applied is shown on the right hand side.

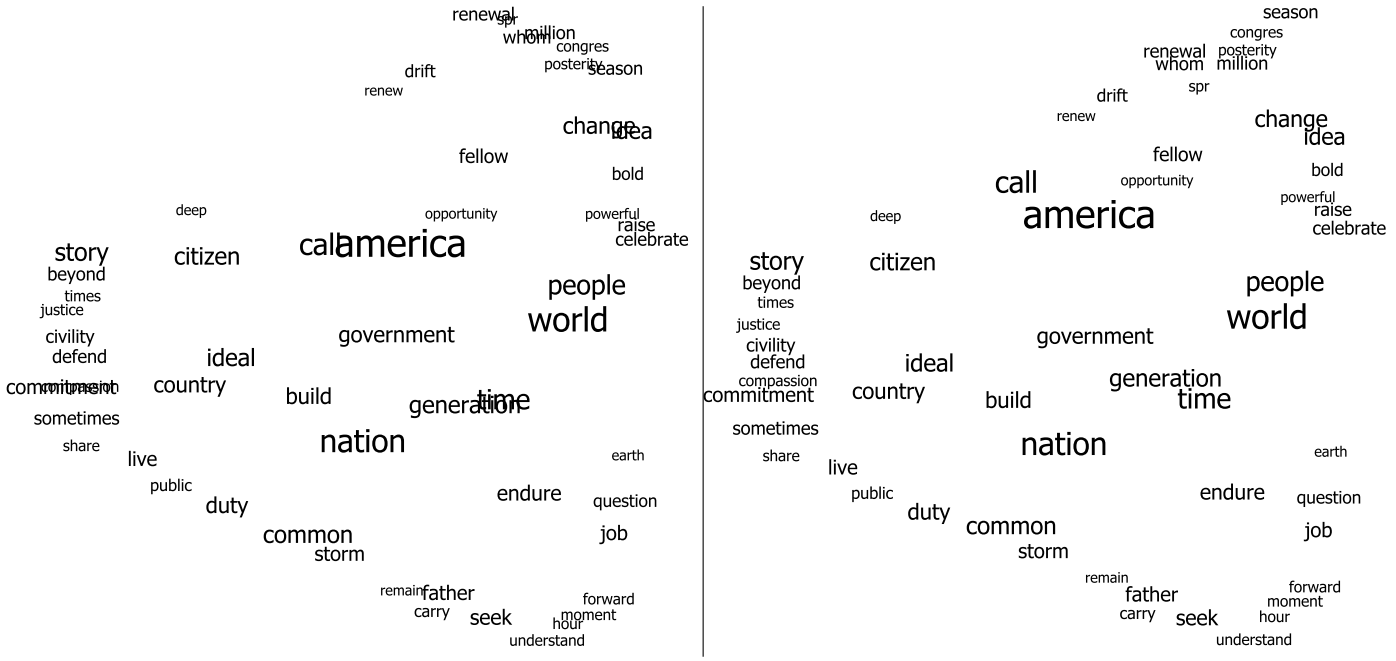

Fig. 5. Tag clouds visualizing the inaugurals of the last three presidents of the United States. On the left is the tag cloud created with distance harmonization on the right the tag cloud with the FTA applied additionally.

argued to use the rectangle which is the nearest to the center of gravity of the cluster as core rectangle. Starting with the core rectangle all rectangles overlapping with the selected one are moved to resolve the overlapping. Also rectangles which overlap the rectangles to be moved are moved in the same direction. Repeating the clustering and moving iteratively results in an overlap-free visualization of the rectangles.

In our approach the definition of overlap was slightly changed since it is sufficient to display the terms and not their bounding boxes in a readable way. The visibility of a term is not affected much by a small overlap. Therefore two terms, i.e. their respective rectangles, were only considered to be overlapping if the area of overlap was greater than 10% of the area of the smaller rectangle.

In Figure 5 an example of overlap removal via FTA is depicted. Here again the inaugurals of the US presidents are visualized. The figure on the left shows the result with distance harmonization. In contrast to Figure 3 nearly all terms are readable already. On the right is the result where FTA has been applied additionally. Obviously all terms are visible here, but unfortunately, the resulting tag cloud is distorted.

V. Evaluation

We already provided a visual evaluation of our approach and showed how distance harmonization and FTA improve the readability and aesthetics of the tag cloud. In addition a quantitative evaluation expresses how distance harmonization increases the stress while decreasing the size and the overlapping. Therefore we use four evaluation criteria.

At first the total space $\text{size}\,(P)$ required to display the tag cloud positioning $P := P(T)$ for all terms $t \in T$ is considered. As mentioned in Section IV-C an automatic scaling of distances is applied. The scaling factor is based on the minimal required space (e.g. the sum of all term rectangles), with the aim of keeping the size of the final visualization at an appropriate level.

The overlapping measure is defined by the sum of all overlapping areas. Here we don't take into account that an area can contain information from more than two terms. If this is the case the view deteriorates, since positioning more than two terms on top of each other further reduces their visibility. Here the function $\text{overlap}\,(t_i, t_j)$ calculates the overlapping rectangles between the corresponding rectangles of terms t_i and t_j and returns 0 if they don't overlap.

$$\text{overlap}\,(P) = \sum_{i<j} \text{overlap}\,(t_i, t_j)$$

Afterwards the amount of white space is calculated using the total space and the overlapping measure.

$$\text{white}\,(P) = \text{size}\,(P) - \sum_{t \in T} \text{size}\,(t) + \text{overlap}\,(P)$$

Finally stress E, already defined in Section IV is evaluated. Note that for all kinds of harmonizations we use the scaled distances to measure the stress. Therefore an increase of stress must be expected, as MDS uses harmonized distance values.

Table I shows the measurement for the example considering the three inaugurals. It contains the measurements for MDS without adjustments (pureMDS), MDS on harmonized distances with and without the Force-Transfer-Algorithm. In comparison the values for the standard alphabetical sorted table tag cloud and the inside out tag cloud can be seen in the last two rows. The percentage of overlapping is already decreased

Used method	overlap	whitespace	size	stress(E)
pureMDS	$37.7 \cdot 10^3$	$582 \cdot 10^3$	$635 \cdot 10^3$	0.251
harmonized	$5.97 \cdot 10^3$	$491 \cdot 10^3$	$576 \cdot 10^3$	0.347
harmonized,FTA	$2.94 \cdot 10^3$	$461 \cdot 10^3$	$550 \cdot 10^3$	0.359
sorted	$0 \cdot 10^3$	$82.1 \cdot 10^3$	$173 \cdot 10^3$	
insideout	$0 \cdot 10^3$	$151 \cdot 10^3$	$242 \cdot 10^3$	

TABLE I
QUANTITATIVE EVALUATION OF THE INAUGURALS VISUALIZATION, WITH DIFFERENT TAG CLOUD VIEW. ALL MEASURES, EXCEPT THE STRESS, ARE IN PIXELS.

by distance harmonization and further by the FTA. The same behavior can be seen for the white space values. On the other hand, distance harmonization decreases the size of the tag cloud, in comparison to pureMDS. And also FTA afterwards only slightly increases the size again. The stress on the other hand is increased by harmonization as well as FTA. Finally, with only moderately increasing stress we could achieve next to zero overlapping and also decreases the size and the white space of the tag cloud.

VI. CONCLUSIONS

In this work we have presented a distance aware tag cloud visualization as well as methods to improve the readability, usefulness, and aesthetics of the resulting tag cloud. As usual, the importance of a term is determined by its number of occurrences and visualized by the term's font size. In addition relations between terms based on their cosine distances in the document space are represented by their positioning in the tag cloud. To position the terms, a multi-dimensional scaling approach with stress minimization was used. The distances of terms in the tag cloud are adjusted to remove overlaps and large amounts of empty white space. Since the proportions of the distances are preserved, their information content is maintained. For a further fine tuning of the positioning we used the FTA algorithm to remove the remaining overlap of the term's bounding boxes.

ACKNOWLEDGMENT

This work was partially supported by the DFG Research Training Group GK-1042 "Explorative Analysis and Visualization of Large Information Spaces".

REFERENCES

[1] S. Bateman, C. Gutwin, and M. Nacenta, "Seeing things in the clouds: the effect of visual features on tag cloud selections," in *HT '08: Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*. New York, NY, USA: ACM, 2008, pp. 193–202.

[2] F. B. Viégas and M. Wattenberg, "Tag clouds and the case for vernacular visualization," *interactions*, vol. 15, no. 4, pp. 49–52, 2008.

[3] S. Milgram and D. Jodelet, "Psychological maps of Paris," *Environmental psychology*, pp. 104–124, 1976.

[4] C. Seifert, B. Kump, W. Kienreich, G. Granitzer, and M. Granitzer, "On the beauty and usability of tag clouds," in *Information Visualisation, 2008. IV'08. 12th International Conference*, 2008, pp. 17–25.

[5] S. Bateman, C. Gutwin, and M. Nacenta, "Seeing things in the clouds: the effect of visual features on tag cloud selections," in *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia*. ACM, 2008, pp. 193–202.

[6] M. Martín-Merino and A. Muñoz, "A new mds algorithm for textual data analysis," in *ICONIP*, 2004, pp. 860–867.

[7] Á. Blanco and M. Martín-Merino, "A partially supervised metric multidimensional scaling algorithm for textual data visualization," in *IDA*, 2007, pp. 252–262.

[8] M. Martín-Merino and Á. Blanco, "A local semi-supervised sammon algorithm for textual data visualization," *J. Intell. Inf. Syst.*, vol. 33, no. 1, pp. 23–40, 2009.

[9] W. Cui, Y. Wu, S. Liu, F. Wei, M. X. Zhou, and H. Qu, "Context preserving dynamic word cloud visualization," in *2010 IEEE Pacific Visualization Symposium (PacificVis)*. Taipei: IEEE, März 2010, pp. 121–128.

[10] M. R. Berthold, N. Cebron, F. Dill, T. R. Gabriel, T. Kötter, T. Meinl, P. Ohl, C. Sieb, K. Thiel, and B. Wiswedel, "Knime: The konstanz information miner," in *Studies in Classification, Data Analysis, and Knowledge Organization (GfKL 2007)*. Springer, 2007.

[11] J. Kruskal and M. Wish, *Multidimensional scaling*. Sage Publications, Inc, 1978.

[12] M. Cox and T. Cox, "Multidimensional scaling," *Handbook of data visualization*, pp. 315–347, 1988.

[13] J. Sammon, "A nonlinear mapping for data structure analysis," *IEEE Transactions on Computers*, vol. 18, no. 5, pp. 401–409, 1969.

[14] I. Borg and P. Groenen, *Modern multidimensional scaling*. Springer New York, 1997.

[15] X. Huang and W. Lai, "Force-transfer: a new approach to removing overlapping nodes in graph layout," in *ACSC '03: Proceedings of the 26th Australasian computer science conference*. Darlinghurst, Australia, Australia: Australian Computer Society, Inc., 2003, pp. 349–358.

[16] X. Huang, A. Sajeev, and W. Lai, "A scalable algorithm for adjusting node-node overlaps," *Computer Graphics, Imaging and Visualization, International Conference on*, vol. 0, pp. 43–48, 2006.