

Visualizing blog archives to explore content- and context-related interdependencies

Justus Bross
Hasso Plattner Institute
Prof.-Dr.-Helmert-Strasse 2-3
14482 Potsdam, Germany
+49 (0)331 5509 524

justus.bross@hpi.uni-potsdam.de

Patrick Schilf
Hasso Plattner Institute
Prof.-Dr.-Helmert-Strasse 2-3
14482 Potsdam, Germany
+49 (0)176 82026250

pschilfsg@gmx.de

Christoph Meinel
Hasso Plattner Institute
Prof.-Dr.-Helmert-Strasse 2-3
14482 Potsdam, Germany
+49 (0)331 5509 222

office-meinel@hpi-web.de

ABSTRACT

There has been virtually little in the way of user interfaces designed for the exploration and information gathering from large weblog datasets to allow for an integrated and aggregated knowledge collection and information analysis tool. Users have to rely on their own capability to find, select or filter entries and navigate through a blog archive. For weblogs with a large collection of entries this task easily becomes tedious, since current blog interfaces lack fundamental support for facilitating the exploration of their archives. A solution to this problem could be POSTCONNECT, a mature blog-archive visualization tool presented in this paper.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces (GUI, screen design); H.4.3 [Information Systems Applications]: Communication Applications (Information browsers).

General Terms

Measurement, Design, Experimentation, Human Factors

Keywords

Content visualization, weblogs, blogosphere, social spaces, BlogConnect, interdependencies, social knowledge, Information representation, time analysis

1. INTRODUCTION

Since the end of the 90's, weblogs have evolved to an inborn element of the worldwide cyber culture [15]. Compared to around 60 million blogs in the year 2006, the worldwide number of weblogs has increased to a total of more than 133 million in the year 2008 according to two well-known social media tracking sources that are published on a yearly basis [21][24][27]. Technically, weblogs are an easy-to-use, web-enabled Content Management System (CMS), in which dated articles – so-called

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Hypertext '10, June 13–16, 2010, Toronto, Ontario, Canada
Copyright 2010 ACM.

“postings” – as well as comments on these postings are presented in reverse chronological order [4]. Their potential fields of application are highly diverse, beginning with personal diaries, reaching over to knowledge and activity management platforms, and finally to enabling content-related and journalistic web offerings [18]. Some weblogs are creating an enormous and ever-changing archive of open source intelligence. Browsing or exploring this vast pool of data to extract, exploit or leverage (content-related) structures and dynamics in each weblog platform was so far not well supported for the users.

Blog Postings of a blog archive are typically arranged in reverse chronological order but only the latest entries (depends on pre-configuration) are displayed on the front page. However, older content that might still be of value for some users has possibly already been shifted into the non-visible archive at the time of the user's access. The capability of effortlessly exploring a blog archive is therefore essential. In Weblogs, this was so far realized by the plain provision of links to monthly archives as well as a listing of tags for selecting a subset of entries assigned with a specific keyword.

To provide a contemporary and innovative tool for the improved exploration of blog archives, we present POSTCONNECT in this paper: An interactive and well-arranged visualization tool to powerfully explore and browse standard blog systems archives that is the synthesis and application of existing visualization and interaction techniques to the new domain of weblogs. The subsequent section integrates our work into existent research efforts. The implementation details of our visualization tools are elaborated on in section three by discussing its architecture and action sequence of the corresponding application. Section four is solely dedicated to the tool's visualization by clarifying the development phases as well as the functionality of the graphical user interface. Based on insights of the preceding sections, we provide an outlook as well as recommendations for further research with the overall objective to further enhance our visualization tool by extending its functionality and area of application in section five. A conclusion is given in section six, followed by the list of references and an appendix with attendant figures of the research discussed in this paper.

2. RELATED WORK

Related research on the visualization of the blogosphere currently focuses on visualizing and analyzing dependencies within the whole network of the blogosphere instead of analyzing a single blog entity alone.

One approach taken in this regard is link-analysis between different weblogs to track information flow and its continuing spread through the blogosphere. Several scholars compare this dispersion of information with epidemic diseases, where a virus is spread via direct contact to others. Applied to the example of the blogosphere, an interesting topic in one blog is quickly picked up by authors of other weblogs and consequently spread like a virus throughout the blogosphere. The visualization of such a spreading became known as an "infection tree" [1] [11] [3]. Other scholars that were tracking knowledge flows in the blogosphere were for instance Anjewierden et al [2].

Link structure was equally analyzed by Herring and his colleagues to examine dependencies between weblogs [14]. Their study empirically investigates the extent to which (and in what patterns) blogs are interconnected. Their visualization of link patterns in conjunction with quantitative social network analysis as well as qualitative analysis of references and comments found out that the most influential blogs, commonly known as "A-list" blogs, are overrepresented in the network although other groupings of blogs are more densely interconnected.

Chin and Chignel [7] examine how communities can be discovered through interconnected blogs as a form of social hypertext. Their proposed model detects communities in blogs by aligning centrality measures from social network analysis with parameters of the "Sense of Community" model (SOC) of Mcmillan and Chavis [19]. A similar approach of Efimoa et al. was noted in the work "Finding "the life" between the buildings [...]" [8].

Attitudes form the basis of the work of Chau and Xu[6] to reveal and visualize communities of similar interest or thoughts, also known as blog groups or blog rings. They could, for instance, show that blogging (or advertising) about a product on the Web would not necessarily attract people's attention in the blogosphere. This is quite different from other media formats (such as TV), in which recurrent commercial product advertisements often result in higher sales.

"Vizblog" is another visualization technique to reveal similarities between blog entries in different blogs [26]. Through association and content analysis, blog entries are linked to each other to form clusters of related content. By manipulating the graph and filtering content, their visualization let users to navigate and explore online discussions. It can furthermore promote participation by highlighting 'the buzz' of popular topics (see also [22]) and laying out the structure of conversations (see also [30]).

Tirapat et al. have a similar approach in revealing "buzz" in the blogosphere [28]. Their method uses information-retrieval techniques to associate blog entries to topics, collected by crawling an authoritative resource on this subject area. They visualize and systematically analyze several blogs in the same domain of interest (movies) in order to assess to what extent the "buzz" of blogs correlates with public opinion. The collected information is represented in terms of a topic map, which is subsequently visualized in three different types of views, each one designed to communicate a different aspect of the data.

There are other rather loosely related projects that are dedicated to the visualization of social networks in general and not the blogosphere in particular: We like to mention MultiVis[25],

Viszter [12], Prefuse [13] as well as the work of Shen et al. [23], Trier [29] and Gloor et al. [9] in this regard.

On top of academic researchers that are - as presented before - interested in analyzing the interconnected blogosphere for a multitude of reasons like revealing hot topics (buzz), content or community clusters, the spread of knowledge (infection tree) as well as centrality or emotion analysis, similar efforts are undertaken by commercial providers in offering blog intelligence services. Well-known services are among others Blogscope¹, Dig², Technorati³ and Blogpulse⁴.

However, the existing academic work discussed beforehand differs substantially compared to our approach. While it is their main objective to map or reveal interdependencies within the whole of the network (blogosphere or at least fractions of it), we in turn strongly believe that the basic requirement for any of these efforts lies in understanding the content- und contextual structure of each single entity of a weblog beforehand. Modeling and mining the vast pool of textual data within each single weblog to extract, exploit and describe meaningful knowledge to ultimately leverage content-related structures or clusters is the higher-level aim of our project presented in the following.

There are two existing research initiatives our project could currently be compared to. The first one is the iBlogVis system presented by Indramato et al. [17]. Instead of analyzing and visualizing the blogosphere as a whole, their work focuses on supporting content exploration of individual weblog archives and providing the history of user interaction in these archives. The biggest shortcoming of this tool however is the absence of standard data sets and user tasks.

The second one is the IN-SPIRE system presented by Gregory and his colleagues [10]. This tool allows users to harvest blogs, view them by thematic content, isolate key words of interest, run queries, visualize changes in content over time or isolate bloggers of interest. However, again this is an approach focusing on the whole of the blogosphere - or at least fractions of it.

3. IMPLEMENTATION

POSTCONNECT was implemented as an Adobe-Flex⁵ application with the aid of Flare⁶. Flare is an ActionScript-library specialized on data visualization for web applications. It serves as a gateway between the capture of blog data and their graphical illustration. A PHP-script collects the attendant data from the *Wordpress* database at frequent intervals and provides it to the Flex-application in the JSON-format.

3.1 Architecture

Figure 1 displays an overview of the core components of POSTCONNECT. The *Controller* constitutes the central element that manages the flow of information between the single components.

¹ <http://www.blogscope.net/>

² <http://digg.com/>

³ <http://technorati.com/>

⁴ <http://www.blogpulse.com/>

⁵ <http://www.adobe.com/de/products/flex/>

⁶ <http://flare.prefuse.org/>

The *SourceManager* serves as the gateway to Wordpress by loading the blog data into the application. For that purpose it administers multiple sources of data for the different blog characteristics posts, categories and tags. Each source obtains its data from exactly one JSON-string generated by the PHP-script. The *FilterManager* filters the blog data according to defined rules and determines the amount of objects to be visualized. Each object can be filtered by any of its properties. The corresponding filtering rules are defined similarly to the ones of a SPAM-filter, for which the property to be filtered, the target value to pass the filter as well as an operator (e.g. containedness of a keyword in a string) need to be determined. The visualization of the graph that contains the objects to be displayed is realized by the *Visualization* object.

data in an adequate form ready to be visualized. The *Controller* delivers the data to the *SourceManager*, whose respective sources - filled with blog-data - modify it with presentable objects. The non-processed data is meanwhile saved as a copy to avoid repetitive loading processes - for instance in case of user interaction during runtime of the implementation. When the data is completely loaded, the *Controller* is notified via a callback function. The modified data-container is then handed-over from the *Controller* to the *FilterManager*.

Similar to the *SourceManager*, the data container is then processed by each of the *FilterManager*'s filters. The difference here lies in the fact that filters are organized sequentially, where each output serves as input for the following one. The last filter then determines the final amount of data, out of which the graph

is built. The graph contains merely nodes till the initial generation of the graph, while edges are momentarily ignored. The edges are therefore not compiled until the user requests the visualization of a post's dependencies within a weblog. First reason for this procedural method lies in the improvement of performance, and the corresponding ambition to make POSTCONNECT equally executable on older computer systems. Since the visualization of bundled edges necessitates *B-Spline-curves* that are quite computing-intensive, the amount of curves to be rendered at once was minimized with this method. Secondly, the initial generation with all available

edges and nodes would reduce the understandability of the diagram greatly. This problem originates in the way of looking at how standard blog-data can most meaningfully be visualized. Section 4.2 will elaborate further on this issue.

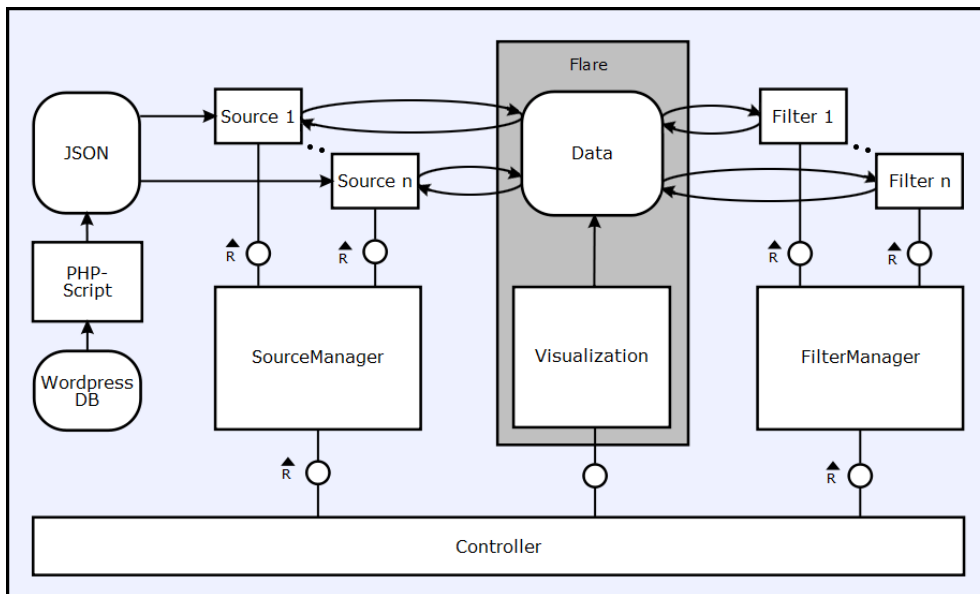


Figure 1: Architecture

Flare already defines the connection to a *Data*-instance that serves as a container for the objects to be displayed. The Filters and sources continuously modify the data object as shown in figure 1. Within the framework of Flare, the *data*-container organizes the blog-data in groups that originate from a data source of the *SourceManager*, whereas the graph is generated by the default groups *nodes* and *edges*.

The current implementation does not rely on the shapes and the layout-templates as predefined by Flare, since they did not fully meet the demands for the designated field of application in the blogosphere. This is why POSTCONNECT defines custom layout classes on top of to the components discussed here. These classes are however not relevant for the overall comprehension of the functionality, because of which they are dealt with shortly in section 4.2.

3.2 Action Sequence

At the beginning of the application the *Controller* initializes a visualization-instance that references an empty data object at this point. It is then handed over through every single component and meanwhile filled with content until it finally provides the blog

4. VISUALIZATION

The interactive visualization technique "citation patterns"⁷, a joint project of *Eigenfactor Project*⁸ and *Moritz Stefaner*⁹ based upon the EIGENFACTOR™ METRICS, were taken as the basis for our visualization tool.

4.1 Development phases

The first conceptual development step displayed in figure 2.A still showed fundamental similarities to the original visualization technique of "Eigenfactor citation patterns". We arranged the

⁷ <http://well-formed.eigenfactor.org/radial.html>

⁸ <http://eigenfactor.org/>

⁹ <http://moritz.stefaner.eu/>

typical blog components posts, categories and tags on several parallel rings in a circular arrangement. Posts are directly linked with each other whenever they exhibit a similar tag or category. A specific color-code indicates which one of the two is the reason for the linkage. However, it soon became apparent that this basic model was not an appropriate arrangement for the visualization of blog data, since one post can be arranged into more than one category because of which the hierarchical edge bundling technique could not be used. A precise mapping was therefore not possible and often resulted in confusion. Since a category has no coercive dependency to a post on an adjacent ring in the same segment of the circle, the visualization consequently had to be adapted. The same applies for the tag-ring.

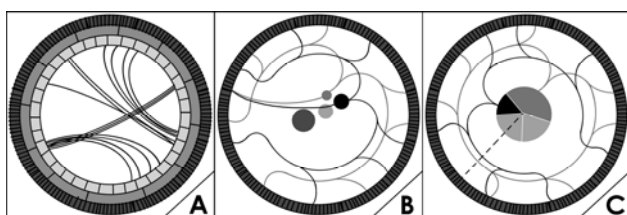


Figure 2: Development phases of visualization

Figure 2.B shows the consequence of this insight. The outer ring merely displays the posts, arranged in chronological order of its publication date. Compared to a clock-face this would mean that the first post is situated at 12:01 o'clock position and the last post published in a blog at 11:59 o'clock position. The categories of a blog are illustrated as bubbles in the circle's center area. Their size depends on the number of post published within each respective category. Here, the edges represent the common tags of an "active" (clicked-on) post and its related categories. Posts are therefore not directly connected among each other but instead through their common categories. The "strength" of a connection (number of common tags) is displayed via altering transparency of the edges.

Similar to figure 2.A, the hierarchical edge bundling technique [16] could also not be used in this development stage. Basically, bundling was nonetheless still employed here in order to preserve the readability of the visualization. In doing so, bundling is realized with control points dependent on the post's and category's position that pools all edges in regular intervals and opens out into the respective category-bubbles. However, the problem is that categories and edges can cut across each other. Furthermore, in- and outgoing edges of a single category might cut across each other as well. Then again, temporal relevance of certain topics can be meaningfully illustrated through the accumulation of posts with similar tags in a certain segment of the circle.

The final development stage (figure 2.C) is our solution to the above-mentioned limitations. The pie-chart-like visualization of categories at the circle-center circumvents that any of the edges might cut across another object. The diagram is furthermore greatly cleaned in its appearance, since there is now only a single connection from the "active" (clicked-on) post into the center. This stands in contrast to the active post being connected to every single category it is related to. These connections that were still existent in figure 2.B did not produce any advantage whatsoever, since affiliation of a post to categories can better be highlighted in terms of color (see figure 3).

4.2 Functionality and User Interface

What you see in figure 3 (appendix) is the central user interface of POSTCONNECT. Within this single graphical representation of a blog's entire content, contextual interdependencies and structural characteristics can be easily and meaningfully be controlled and explored. Using the dataset of the German weblog IT-Gipfelblog¹⁰, the user interface reads as follows:

All posts of a single weblog are arranged as segments on the outer circle of our visualization in chronological order. To start exploring content and contextual interdependencies in a weblog you first choose a post of interest. This "active" post, marked as a single yellow segment on the outer circle and furthermore indicated by the straight line into the center (see letter "A" in figure 3), is the one for which POSTCONNECT now generates visual relationships with the remaining content of the blog. These relationships are based upon *tags* given for single posts. Since tags are a common method in the blogosphere to summarize a post's content, it was the most straightforward determinant for the identification of contextual relationships.

The relationships that can be explored with POSTCONNECT are the ones between the "active" post, all those other blog-posts in which similar tags have been chosen, and the categories in which these posts have been classified. The content-related categories usually chosen and predefined by any weblog-administrator are represented in the pie-chart-like central point of *POSTCONNECT* (see letter "B" in figure 3). Every category has a unique color code and differs in dimension. Categories, in which more posts were published, are thus represented bigger within the pie-chart than those with fewer posts. Categories related to the "active" post are highlighted with a unique color-code, otherwise they remain grayish. Related posts that are connected to the "active" post are highlighted as a black segment on the outer circle representing all blog posts available. All non-active posts remain – similarly to the inactive categories – grayish.

By activating a post to explore its dependencies within a blog-platform, you get all tag-based connections to other posts as sinuous lines (see letter C in figure 3) that have the same color-code as the categories into which these related posts are grouped. While the more highlighted lines indicate a connection based upon a multitude of tags and thus a strong content- or contextual-related dependency, the transparent lines show a lower degree of relationship.

A listing with all those posts (figure 4), categories (figure 5) and tags (figure 6) related to the currently activated post (marked yellow) is provided in the top right area of the user interface. Detailed information of the activated post like title, author, timestamp or current ranking is furthermore provided here.

Even though POSTCONNECT focuses on the visual representation of a single blog's content, we consider it to be crucial to include the bigger picture in which a single weblog is embedded: a segmented and independent public that dynamically evolves and functions according to its own rules and with ever-changing protagonists, a network also known as the "blogosphere" [14]. A single weblog is embedded into this network through the usage of hyperlinks as well as its so-called "blogroll", a blogosphere-

¹⁰ <https://www.it-gipfelblog.de>

internal referencing system [5], while single posts are interconnected with the rest of the blogosphere via track- or pingsbacks as well as referrers. Since content- and contextual relationships are not limited to the boundaries of a single weblog platform, these in- and outgoing (content-related) linkages of single posts with the rest of the blogosphere needed to be included in our visualization. POSTCONNECT highlights the segments of single posts with inherent out- or ingoing connections to the blogosphere with the colors green (for incoming links like referrers or pingsbacks) and red (for outgoing linkages like trackbacks or common links) on the outer circle of our visualization (refer to letters "E" and "D" in figure 3).

Especially for weblogs that are already in operation for several years and that have published hundreds or even thousands of different postings in their platform, a visualization of all posts at the same time does not seem reasonable. POSTCONNECT therefore offers a variety of tools, summarizing listings and tables, as well as filtering mechanisms that facilitate the exploring-experience for weblogs with a large amount of content greatly. Positioned at the top right area of the POSTCONNECT user interface (refer to figure 3), these utilities provide extra information for the current actions performed.

The filtering mechanism currently expanded in the screenshot of figure 3 is another powerful tool to find only the information you are interested in. Especially within the datasets of large weblogs you can hereby preselect and limit the data available for exploration and analysis. By filtering the content with regards to authorship, timestamp as well as selecting specific posts, categories or tags of interest, the analysis is laid out more clearly. You can furthermore zoom into an area of interest within the visualization to explore taglines in more detail. The zoomed-in section can furthermore be dragged and therefore easily moved into the center of the user interface.

Menu tab "Blogosphere" provides more detailed information about the in- or outgoing linkages of single posts within a blog of interest and the rest of the blogosphere. This tab thus indicates the name of the referring blog including URL and ranking according to a service like Technorati, as well as more specific information about post within these referenced blogs like author, timestamp, topic, and much more.

5. FURTHER RESEARCH

The visualization tool presented in this paper visualizes data of Wordpress blogs alone. Even though Wordpress is one of the most commonly used blogging software worldwide [20], it is of strategic interest for us to make our tool usable for most of the other popular weblog platforms. This is why the Flex-application was implemented independently of Wordpress so that the only component to adapt for a different system would be the data-collecting PHP-script. Additional blog-characteristics (e.g. author information) could also be added with minor extra effort. We would for instance consider it to be highly interesting to add an extra "author" variable to the existing dataset that could – when meaningfully paired and weighted with other variables (ranking or voting of posts for instance) – provide some sort of reputational score. The available dataset could as a consequence be filtered to obtain only posts with a predefined minimum reputational score.

We also expect from our continuing developments on PostConnect that it can eventually provide intelligent advises and deductions on the basis of automatic blog-archives analyses. Blog administrators could for instance make use of this "blog intelligence service" to obtain insights about how to manage content more effectively through rearrangements of categories or the deletion of redundant tags or posts. On the basis of their preceding search or interactions patterns, users in turn could be provided with automatic recommendations that indicate which articles or post might be interest for them.

Another major objective is to develop our visualization tool towards a maturity level, with which it could be offered as a free add-on for blogging communities like the one of Wordpress. On top of the basic browsing- and exploring functionality as explained in the preceding sections, many single weblogs could form a community-group of blogs that could use our tool to browse the aggregated archives of all weblogs connected to the group. The points through which single blogs would be interconnected are the colored in- and outgoing posts (green and red for referrers, track- and pingsbacks) that were already introduced with the "blogosphere" POSTCONNECT menu tab in section 4.2. We call the further development of the research described in this paper as BLOGCONNECT. Another extension of such a community could be achieved with results from an ongoing blog-crawler project, which is capable of collecting the same type of data that is necessary for POSTCONNECT. With this data crawler-data however, POSTCONNECT would not necessarily have to be installed on each physical machines hosting a weblog, but could provide almost the same browsing functionality from a central web-enabled instance.

6. CONCLUSION

We discussed the rationale, related research efforts and implementation details of our interactive visualization tool for decisively facilitating the exploration of blog archives for the users compared to existing solutions. The enhancements achieved through POSTCONNECT can be grouped into those from a user or reader perspective and the administrator- or author perspective of such a weblog platform.

For administrator, our visualization offers fundamental alleviation in optimizing the content structure of a weblog, since missarrangement of categories as well as redundancy or overcastted temporal accumulation of posts (and therefore topics) can more easily be identified and the editorial staff consequently more quickly be advised to change it. For users, POSTCONNECT greatly optimizes the exploration of blog archives that were so far confusing and hard to browse explore.

Next to the continuing efforts we like to make in further developing POSTCONNECT, there are several promising directions to follow up on the research presented in this paper. We expect immense synergies from our independently undertaken blog-crawler-project in order to achieve fundamental insights about content-related interconnections of the whole blogosphere. Based upon these synergies we plan to launch a service called BLOGCONNECT that can provide the same insights as POSTCONNECT, but on a much bigger scale.

7. REFERENCES

- [1] Adar, E. and Adamic, L.A. Tracking Information Epidemics in Blogspace. *Web Intelligence*, (2005).
- [2] Anjewierden, A., Hoog, D., R., B., R, and Efimova, L. Detecting knowledge flows in weblogs, 13th. International Conference on Conceptual Structures (ICCS, (2005), 1-12.
- [3] Aschenbrenner, A. and Miksch, S. Blog Mining in a Corporate Environment. 2005. http://www.citeulike.org/user/domakesaythink/article/2382823?citation_format=IEEEtran#.
- [4] Bross, J., Acar, A.E., Schilf, P., and Meinel, C. Spurring Design Thinking through Educational Weblogging. 2009 International Conference on Computational Science and Engineering, Ieee (2009), 903-908.
- [5] Bross, J., Quasthoff, M., Berger, P., Hennig, P., and Meinel, C. Mapping the blogosphere with rss-feeds. The IEEE 24th International Conference on Advanced Information Networking and Application, IEEE (2010).
- [6] Chau, M. and Xu, J. Studying Customer Groups from Blogs. *fbe.hku.hk*, , 850-856.
- [7] Chin, A. and Chignell, M. A social hypertext model for finding community in blogs. *Conference on Hypertext and Hypermedia*, (2006).
- [8] Efimova, L., Hendrick, S., and Anjewierden, A. Finding'the life between buildings': An approach for defining a weblog community. *Internet Research* 6, 1997 (2005).
- [9] Gloor, P., Laubacher, R., Zhao, Y., and Dynes, S. Temporal visualization and analysis of social networks. NAACSOS Conference, June, Citeseer (2004), 27-29.
- [10] Gregory, M., Payne, D., McColgin, D., Cramer, N., and Love, D. Visual Analysis of Weblog Content. ICWSM'2007, PNNL-SA-54063, International Conference on Weblogs and Social Media'07, Boulder, CO, United States (US). (2007).
- [11] Gruhl, D., Guha, R., Liben-Nowell, D., and Tomkins, A. Information diffusion through blogspace. *Proceedings of the 13th international conference on World Wide Web*, ACM New York, NY, USA (2004), 491-501.
- [12] Heer, J. and Boyd, D. Vizster: Visualizing online social networks. *Proceedings of the 2005 IEEE Symposium on Information Visualization*, (2005), 33-40.
- [13] Heer, J., Card, S., and Landay, J. Prefuse: a toolkit for interactive information visualization. *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM New York, NY, USA (2005), 421-430.
- [14] Herring, S., Kouper, I., Paolillo, J., et al. Conversations in the blogosphere: An analysis" from the bottom up. *Proceedings of the 38th Hawaii International Conference on System Sciences (HICSS'05)*, Ieee (2005), 107b-107b.
- [15] Herring, S., Scheidt, L., Bonus, S., and Wright, E. Bridging the gap: A genre analysis of weblogs. *Proceedings of the 37th Hawaii International Conference on System Sciences (HICSS'04)*, (2004).
- [16] Holten, D. Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. *IEEE Transactions on Visualization and Computer Graphics* 12, 5 (2006), 741-748.
- [17] Indratmo, J. and Gutwin, C. Exploring blog archives with interactive visualization. portal.acm.org, (2008).
- [18] Kircher, H. Web 2.0 - Plattform für Innovation. *it - Information Technology* 49, 1 (2007), 63-65.
- [19] Mcmillan, D. and Chavis, D. Sense of community: A definition and theory. *Journal of Community Psychology* 14, 1 (1986), 6 - 23.
- [20] Mintert, S. and Leisegang, C. Liebes Tagebuch ... Sieben frei verfügbare Weblog-Systeme. *iX-Archiv* 7, (2008), 42-53.
- [21] Parker, G. Power to the People - Social Media Tracker Wave 4. *Critical Studies in Media Communication* 4, 2009. <http://universalmccann.bitecp.com/wave4/Wave4.pdf> .
- [22] Qamra, A., Tseng, B., and Chang, E. Mining blog stories using community-based and temporal clustering. *Proceedings of the 15th ACM international conference on Information and knowledge management*, ACM (2006), 67.
- [23] Shen, Z., Ma, K., and Eliassi-Rad, T. Visual analysis of large heterogeneous social networks by semantic and structural abstraction. *IEEE transactions on visualization and computer graphics* 12, 6, 1427-39.
- [24] Smith, T. Power to the People: Social Media Tracker Wave 3. Retrieved on September 2, 2008. http://www.goviral.com/articles/wave_3_20080403093750.pdf.
- [25] Sun, J., Papadimitriou, S., Lin, C., Cao, N., Liu, S., and Qian, W. Multivis: Content-based social network exploration through multi-way visual analysis. *Proc. SDM*, (2009), 1063-1074.
- [26] Tauro, C., Pérez-Quñones, M., Isenhour, P., Ahuja, S., and Kavanaugh, A. VizBlog: Discovering Conversations in the Blogosphere. *Technology demonstration at Directions and Implications of Advanced Computing-Conference on Online Deliberation*, University of California, Berkeley, (2008).
- [27] Technorati.com. State of the Blogosphere 2009. October 19, 2009 at 6:00 am. <http://technorati.com/blogging/feature/state-of-the-blogosphere-2009/>.
- [28] Tirapat, T., Espiritu, C., and Stroulia, E. Taking the Community's Pulse, one Blog at a Time. *cs.ualberta.ca*, ACM New York, NY, USA (2006), 169-176.
- [29] Trier, M. Towards a Social Network Intelligence Tool for visual Analysis of Virtual Communication Networks. In K. Meissner and M. Engeli, *Virtuelle Organisationen und Neue Medien*. TUDpress Verlag, Dresden, 2006, 331-342.
- [30] Viégas, F. and Donath, J. Social network visualization: Can we go beyond the graph. *Workshop on Social Networks, CSCW*, Citeseer (2004), 6-10.

8. APPENDIX

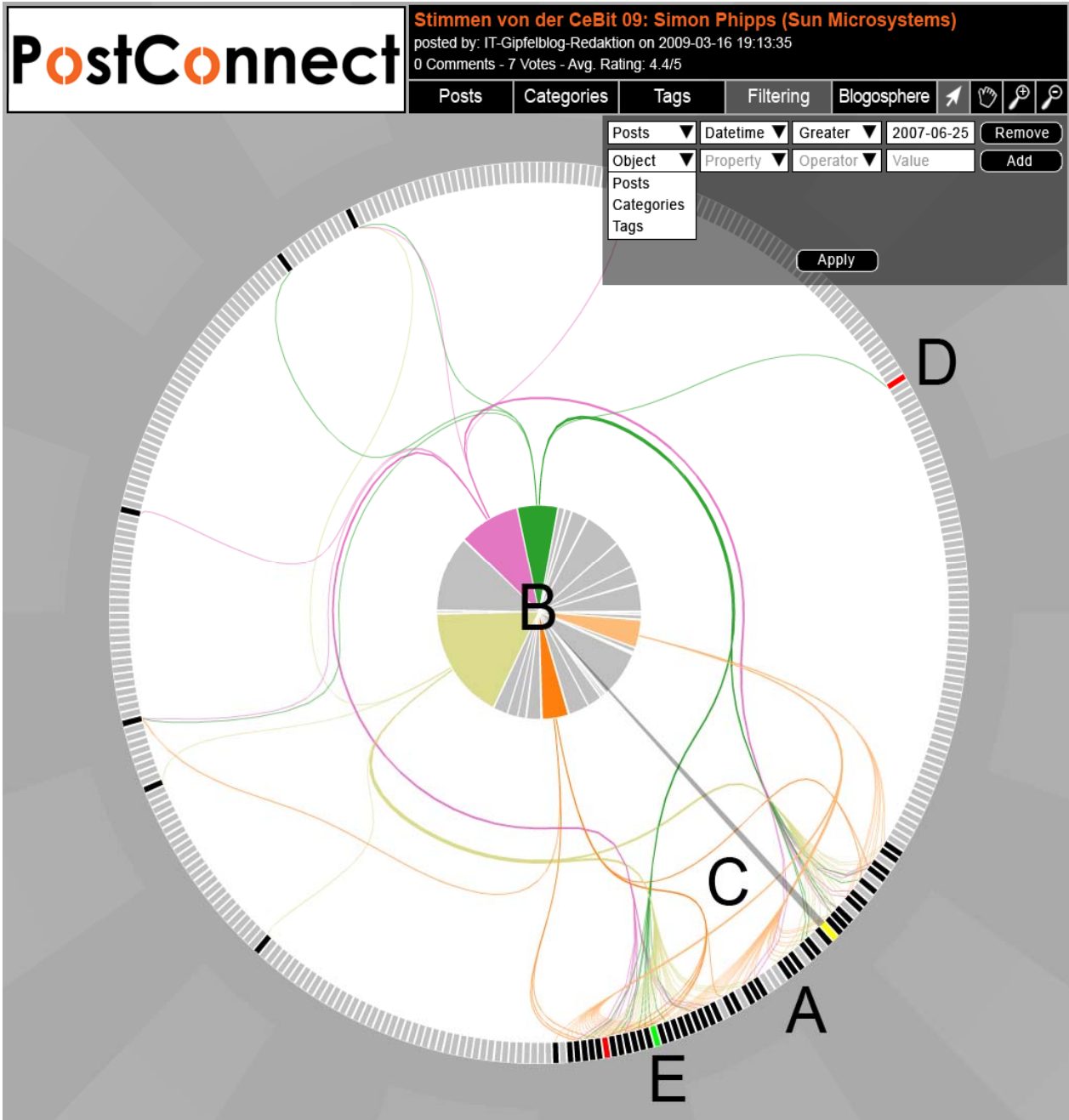


Figure 3: *PostConnect* User Interface

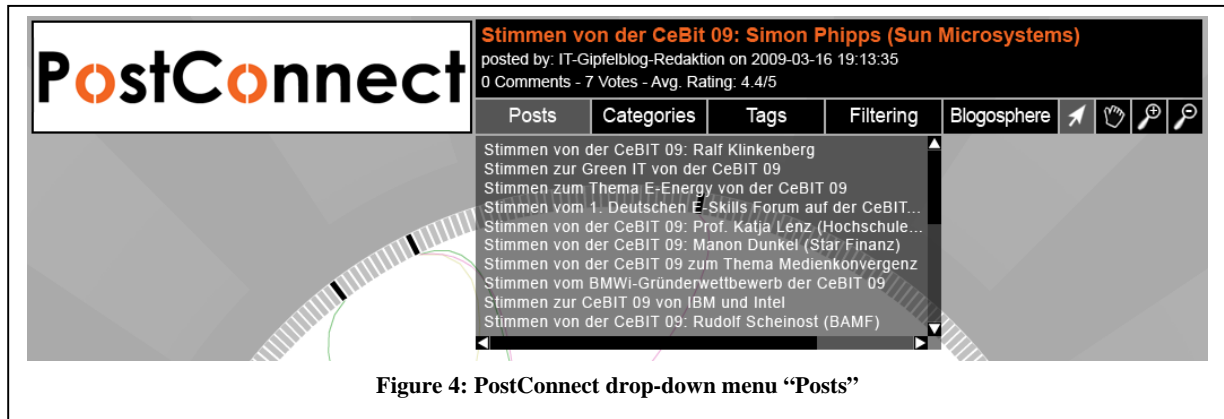


Figure 4: PostConnect drop-down menu "Posts"

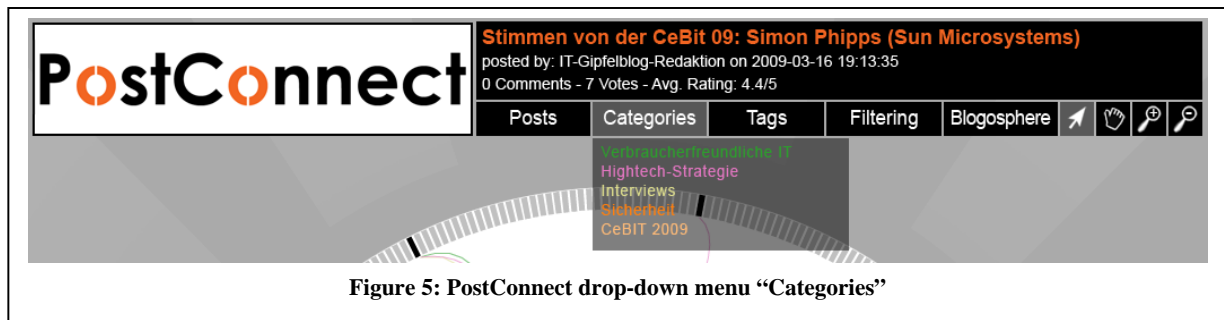


Figure 5: PostConnect drop-down menu "Categories"

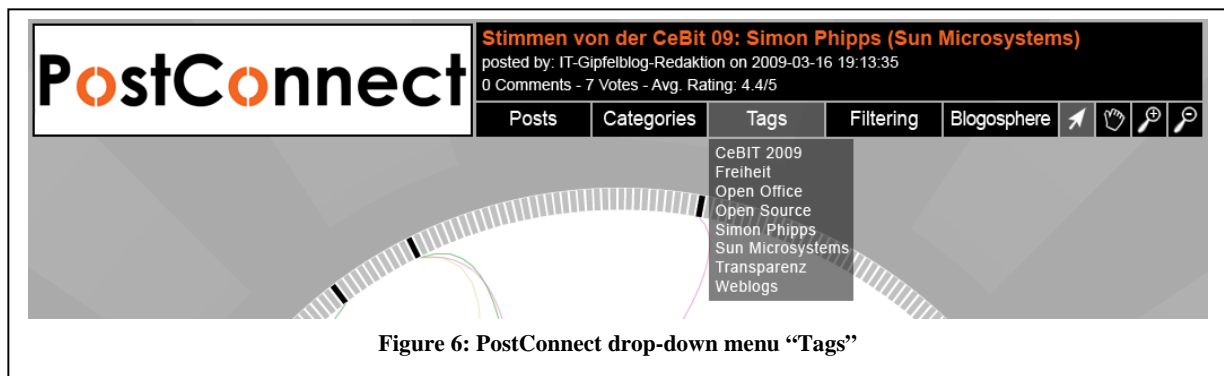


Figure 6: PostConnect drop-down menu "Tags"