Visualization of Influential Blog Networks Using BlogTracker

Abiola Akinnubi*, Nitin Agarwal*, Ayokunle Sunmola†, Vanessa Okeke‡
*Information Science,
University of Arkansas at Little Rock, South University Avenue, Little Rock, 72204, Arkansas, United States
†Computer Science,
University of Arkansas at Little Rock, South University Avenue, Little Rock, 72204, United States
‡Business Information Systems and Analytics,
University of Arkansas at Little Rock, South University Avenue, Little Rock, 72204, United States

Email: asakinnubi@ualr.edu (Abiola Akinnubi)
Email: nxagarwal@ualr.edu (Nitin Agarwal)
Email: seayokunle@ualr.edu (Ayokunle Sunmola)
Email: vokeke@ualr.edu (Vanessa Okeke)
Corresponding authors: Abiola Akinnubi, Nitin Agarwal

***Abstract*—The advent of web 2.0 and social media blogging has enabled researchers to have access to troves of blog data in the 21^{st} century. While platforms and big corporations like Twitter and Facebook can apply the concept of user networks by leveraging internal tools developed by their teams. Researchers and analysts have had to make use of repetitive ways of analyzing blog networks for collected data. This is due to the limited live databases to store and keep track of blog data and the lack of centralized publicly available tools with such capability. When analyzing blog data, the analyst often wants the capability to model relationships and see blogs that share ideological similarities. This is so because blogs always reference each other when they share similarities in content or when they attempt to reinforce a point of view discussed on the medium. Since the blogosphere is made up of a virtual network of blogs - the blogosphere is defined as the network of blogs and has no limitation in blogs referencing one another. It becomes imperative to have a solution that can allow an analyst to visualize the relationships between blogs based on how influential these blogs are when the analyst tracks the discus on the blogs. We address this by providing users with the capability to visualize and analyze blogs that are influential and how connected these blogs are by a way of network visualization. This demonstration shows how the BlogTracker application analyze and visualizes the blog network.**

***Index Terms*—BlogTracker, Domain Analysis, Network Analysis, Graph Visualization, Web Application, Social Network Analysis, Blog Network Analysis**

I. Introduction

Analyzing blogs and their connectivity to each other has been a repetitive task for researchers. This is due to researchers not having a centralized tool to harness blog data collected over a long period of time. Hence, often time researchers have to re-compute and visualize the same data over again when studying for new information. Since blogs are allowed to reference as many domains as possible that have information with a similar ideology or perspective as the parent blogs. This allows blogs' relationships to be modelled as a network to determine the relationship in content for multiple blogs. Researchers and analysts can learn so much from social media analysis if the blog's relationships can be modelled by using how they reference or link each other as a criterion. We therefore showcase the salient capability of the BlogTracker application in using outlinks referenced by blogs to visualize the blog network for analysts to understand relationships between various domains.

II. Literature Review

Blogs and website presence over the internet has grown exponentially since early 1990 from the basic hypertext pages of an intranet to the public web where debates and publications now occur. It is estimated that over 200 million new web pages are actively maintained daily with new content [11]. Analyzing the influence generated by the various active websites has seen several tools come to the forefront of solving the challenge. [11] demonstrated how online opinions and influence campaigns can be analyzed on blogs by studying the online campaigns using covid-19 as a case study. To analyze blog networks the large number of blogs that exist may become challenging to map to a network but could be achievable if the influential blogs from the data set of study are used. [2] modelled blog network as a virtual community by observing patterns in community activities (structure and conversations) formed by the blogs. The approach used by [2] involved crawling selected blogs and finding the common links between the blogs through posts and comments to create a graph that forms a social network. The influence is then measured by a way of reciprocity as a visualization indicator and betweenness centrality [2]. Since links between blogs are key in finding influence, [10] used hyperlinks for network analysis of educational blog study. [10] uses outbound links

IEEE/ACM ASONAM 2022, November 10-13, 2022
978-1-6654-5661-6/22/$31.00 © 2022 IEEE

as target nodes to a source node from originating blogs. This allows [10] to measure social ties between two links. Their work was able to use the occurrence within a community to measure density and the overall betweenness of communities. The work of [3] also studied for the presence of communities with students' educational blogs by using the participation of students. Their work was able to find the cohesion between groups of the student based on their blogging interactions and the type of content shared between them. [5] modelled relationships of blog networks by applying collective wisdom by using labels of blogsites to model the relationship in graphs between blogs and bloggers. This allows their work to create a weight for each label associated with the blog or blogger until all blogs and bloggers have been visited, this allows their work to generate a cluster of blogs by their shared similarities.

III. Blog Networks Architecture and Implementation

Our BlogTracker application consists of numerous features but this section will focus on discussing the methodology and implementation procedure for how we are able to generate blog networks. To implement a blog network for "Covid and Vaccine", we proceed by creating a tracker on the BlogTracker Application. We call this tracker "Vaccine Tracker", this will allow us to group and collate blogs that have the keyword vaccine as part of their contents.Figure 1 demonstrates the flow of creating a tracker for the blog network analysis study.

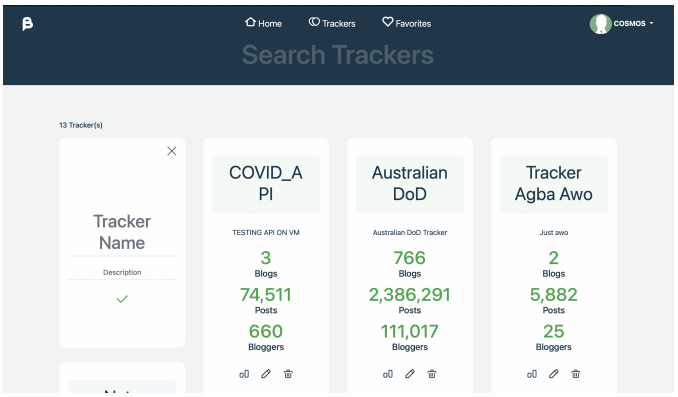

Fig. 1. Demonstrates how to create a tracker for study.

Figure 4 shows the created vaccine tracker with an arrow pointing to the link to proceed to the analytics dashboard. Our approach to extracting outlinks and hyperlinks from blog post content is described in Figure 3 which gives our users an idea of the background processes that BlogTracker goes through during blog domain network extraction. In the next chapter, we demonstrate visual results for one of our selected trackers to help users see an output of what BlogTracker is able to achieve. We are able to support the visualization of network data by extending and leveraging the capability of vosviewer [12]

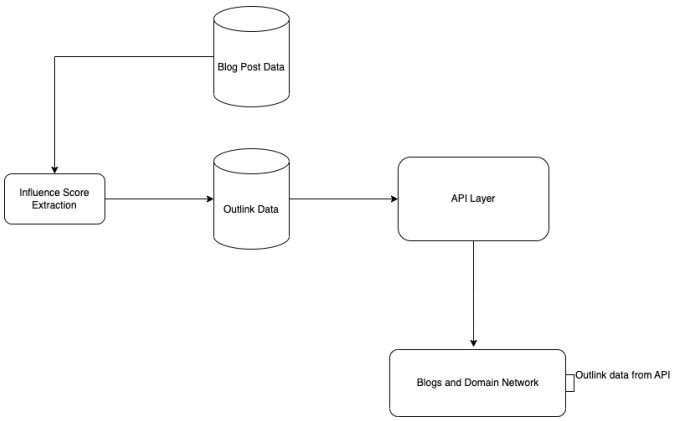

Fig. 2. Implementation flow

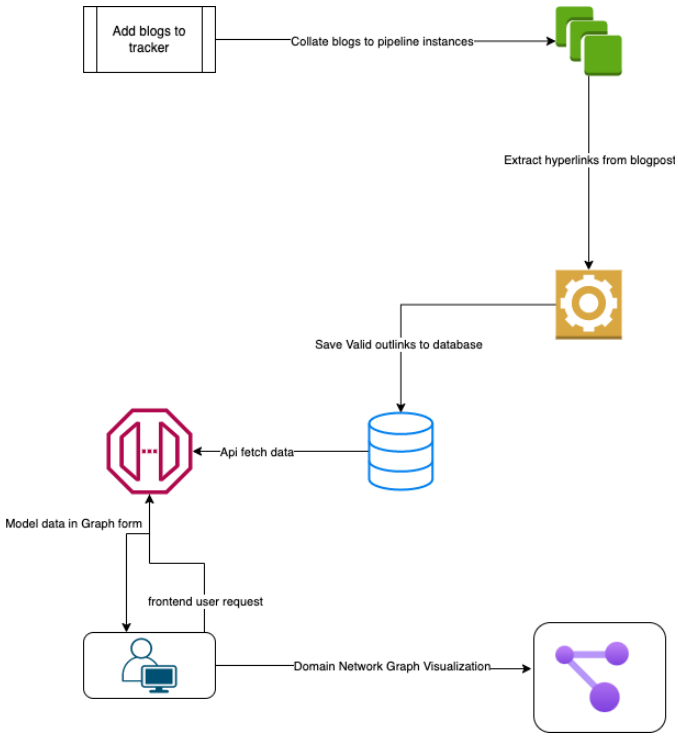

Fig. 3. Background process to blog domain network.

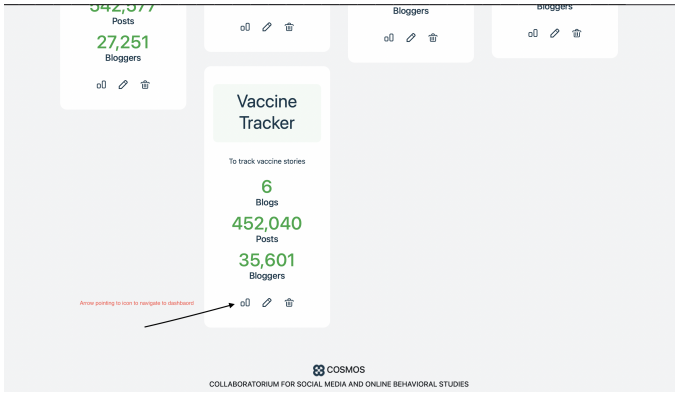

Fig. 4. Arrow points to the icon to click when proceeding to analytic dashboards.

IV. Example Demo: Vaccine and Covid Tracker Analysis

In an example demonstration of the domain (blog) network analysis feature of the BlogTracker application, we looked at the Vaccine events. This tracker contains blog posts with content that has the word vaccine in them since at the data collection stage the user post has specified the search criteria as the vaccine. This helps the data collation and aggregation Boolean script to perform a regex search on various blog data and put all blog posts with vaccine and blog sites name with vaccine under a single tracker name which is unique to each user of our application.

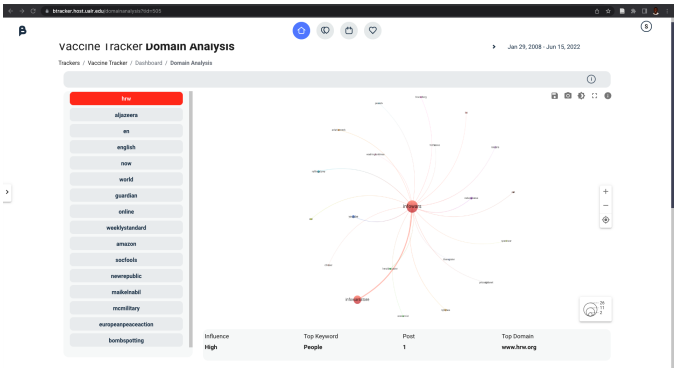

Fig. 5. Visualization of resultant network graphs from a selected tracker and influential domains.

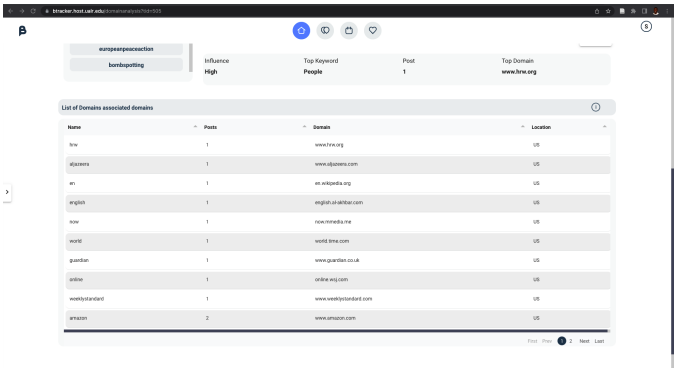

Fig. 6. Table showing other important domain statistics

We have also added the image of an example full-page result of the visualization and analysis generated from the application as seen below inFigure 7. The sidebar shows the list of top influential domains and their ranking in order of influence while the top card panel below the network graph shows other key statistics associated with the selected domain and the top keyword used by the domain. This helps analysts understand some key findings of the domain network and export the visualization in image format.

Furthermore, from the visualization and dashboard displayed in Figure 7the table allows the user to see other key information about other hyperlinks referenced by the selected

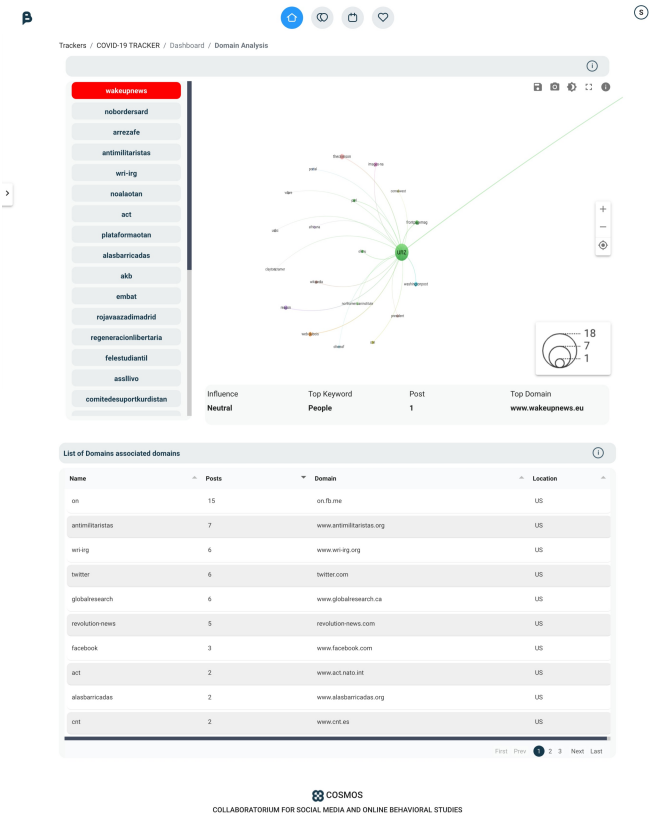

Fig. 7. Full page network visualization of a selected domain network.

tracker which the user can click on the paginated icon to see more items.

Finally, our BlogTracker application is capable of many other salient advanced social network and blog data analyses but this work focuses on showing one of our latest developments in extending the capability of BlogTracker.

V. Conclusion and Future Works

In this paper, we looked at the current unavailability of solutions and the need for a tool with the capability of blog network analysis and visualization given collected blog data under a tracker. To solve this challenge, we developed the capability to analyze and visualize blog networks by applying graphs theory behind the scenes into the existing BlogTracker solution. This will give analysts outside of traditional large corporations who use BlogTracker the ability to analyze and visualize the relationships between blogs based on their influence. Also, our BlogTracker solution is constantly undergoing iterations to solve diverse social network analysis needs now and in the future. We also understand that our tool must be user-friendly and easy to navigate. Hence, Our BlogTracker tool is available to serve various needs in analyzing blog data. Furthermore, one of BlogTracker's utilities is demonstrated via a case study on an existing "Covid and Vaccine" tracker. Analysis of the network of blog data collected through the BlogTracker application shows the influential top blogs and

their overall statistics of outlinks and the number of posts generated for a selected date range. This helps analysts understand how these blogs contribute to the discus in a selected tracker when combined with other analyses available in our BlogTracker application.

In future works, we are working on other types of network analyses possible with blog data, and we hope to add a blogger network feature in the future.

ACKNOWLEDGEMENTS

This research is funded in part by the U.S. National Science Foundation (OIA-1946391, OIA-1920920, IIS-1636933, ACI-1429160, and IIS-1110868), U.S. Office of Naval Research (N00014-10-1-0091, N00014-14-1-0489, N00014-15-P-1187, N00014-16-1-2016, N00014-16-1-2412, N00014-17-1-2675, N00014-17-1-2605, N68335-19-C-0359, N00014-19-1-2336, N68335-20-C-0540, N00014-21-1-2121, N00014-21-1-2765, N00014-22-1-2318), U.S. Air Force Research (FA9550-22-1-0332), U.S. Army Research Office (W911NF-20-1-0262, W911NF-16-1-0189), U.S. Defense Advanced Research Projects Agency (W31P4Q-17-C-0059), Arkansas Research Alliance, the Jerry L. Maulden/Entergy Endowment at the University of Arkansas at Little Rock, and the Australian Department of Defense Strategic Policy Grants Program (SPGP) (award number: 2020-106-094). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations. The researchers gratefully acknowledge the support.

REFERENCES

[1] X. Jin and Y. Wang, "Research on social network structure and public opinions dissemination of micro-blog based on complex network analysis," *Journal of Networks*, vol. 8, no. 7, p. 1543, 2013.

[2] A. Chin and M. Chignell, "Identifying communities in blogs: roles for social network analysis and survey instruments," *International Journal of Web Based Communities*, vol. 3, no. 3, pp. 345–363, 2007.

[3] A. Jimoyiannis, P. Tsiotakis, and D. Roussinos, "Social network analysis of students' participation and presence in a community of educational blogging," *Interactive Technology and Smart Education*, vol. 10, no. 1, pp. 15–30, 2013.

[4] W. H. Hsu, A. L. King, M. S. Paradesi, T. Pydimarri, and T. Weninger, "Collaborative and Structural Recommendation of Friends using Weblog-based Social Network Analysis." in *{AAAI} {Spring} {Symposium}: {Computational} {Approaches} to {Analyzing} {Weblogs}*, vol. 6, 2006, pp. 55–60.

[5] A. Abraham, Aboul-Ella Hassanien, and V. Snášel, "Computational social network analysis: trends, tools and research advances," 2009.

[6] C. S. Dempwolf and L. W. Lyles, "The uses of social network analysis in planning: A review of the literature," *Journal of Planning Literature*, vol. 27, no. 1, pp. 3–21, 2012.

[7] Li-mei Che, Yun-gang Wei, B. Sun, Q. Qiu, and Nai-xin Zhan, "Micro-Blog Social Network Analysis:" Digu" Network as an Example," in *2010 2nd {International} {Symposium} on {Information} {Engineering} and {Electronic} {Commerce}*. IEEE, 2010, pp. 1–3.

[8] M. Jamali and H. Abolhassani, "Different aspects of social network analysis," in *2006 {IEEE}/{WIC}/{ACM} {International} {Conference} on {Web} {Intelligence} ({WI} 2006 {Main} {Conference} {Proceedings})({WI}'06)*. IEEE, 2006, pp. 66–72.

[9] Yao-Jen Chang, Yao-Sheng Chang, Shu-Yu Hsu, and Chiu-Hui Chen, "Social network analysis to blog-based online community," in *2007 {International} {Conference} on {Convergence} {Information} {Technology} ({ICCIT} 2007)*. IEEE, 2007, pp. 2193–2198.

[10] E. Z. F. Liu, R. C. Shih, and Y. L. Tsai, "Hyperlink network analysis of the educational blog," *British Journal of Educational Technology*, vol. 42, no. 2, pp. 25–29, 2011.

[11] A. Akinnubi, N. Agarwal, Z. Stine, and S. Oyedotun, "Analyzing online opinions and influence campaigns on blogs using BlogTracker," in *Proceedings of the 2021 {IEEE}/{ACM} {International} {Conference} on {Advances} in {Social} {Networks} {Analysis} and {Mining}*, 2021, pp. 309–312.

[12] A. Perianes-Rodriguez, L. Waltman, and N. J. van Eck and, "Constructing bibliometric networks: A comparison between full and fractional counting," *Journal of Informetrics*, vol. 10, no. 4, pp. 1178–1195, 2016. [Online]. Available: https://doi.org/10.1016/j.joi.2016.10.006.;https://www.sciencedirect.com/science/article/pii/S1751157716302036