

Visualization Development for the Virtual Town Square

By Rafael Burgos-Guntin¹, Manuel Perez-Quinones², Heather Price³, Andrea Kavanaugh²
University of Puerto Rico-Mayaguez¹, Virginia Tech², Radford University³

Abstract

Data collection has become more and more efficient as time goes on, with repositories accumulating hundreds of gigabytes in some cases. The question we face is “how do we analyze all of this data?” One answer is using visualizations as a mean of displaying data in a plot or graph in which we can compare single or multiple aspects of the data at hand. We are developing these tools as part of the Virtual Town Square to better understand and present the interests and/or concerns of the people of the New River Valley in VA.

Introduction

Data visualization consist in representing and communicating efficiently data with the use of graphs and schemes. The idea behind this project is to develop a few visualizations for the Virtual Town Square website (VTS). The VTS website’s goal is to recreate the old town square, a place where everyone in town came to hear and discuss the local issues. In other words what VTS is trying to do is bring that to the 21st century by creating a interactive news page, specifically for the community of New River Valley, VA.

The way the website works is with help of a local-news aggregator, which as its name states, compiles all the news it can find regarding the New River Valley and stores them in a local database. The local-news aggregator collects information and news not only from regional newspapers, but also social media. The process of representing all this data is the challenge at hand; how can users know what people are saying about a particular topic without having to scroll through a bunch of web pages.?

Background

Visualization within the context of VizBlog can help promote online discussion and participation by highlighting popular topics and allowing information discovery (VizBlog 2007). Our prior work with VizBlog help users identify related blogs based on content analysis.

For the Virtual Town Square, we want to take his a step further by including topic analysis of Twitter and Facebook posts. In addition, we want to take advantage of Twitter and it’s mobile use for VTS. Wohn and Na (2011) found that tweets about specific television shows tended to come from mobile devices. The use of mobile devices within our site would allow us to analyze not just the content but the location of the tweet.

Visualizations in VTS

We created two visualizations for the VTS website. These were intended to present clearer data from tweets regarding the New River Valley. One visualization is a Treemap showing frequency of words in tweets and the other one is a Connected graph showing the relationship between people and their interests.

The first visualization is a Treemap (Figure 1). The Treemap visualization is based on a frequency count of words in tweets. The words come from the tweet’s VTS database. Based on the frequency counts, the program produces a map like that shown in Figure 1. The visualization shows each word with a box of size relative to frequency of occurrence in the tweet collection. The more often the word appears the larger the box. With the application of this visualization we can know in seconds what people are tweeting about.

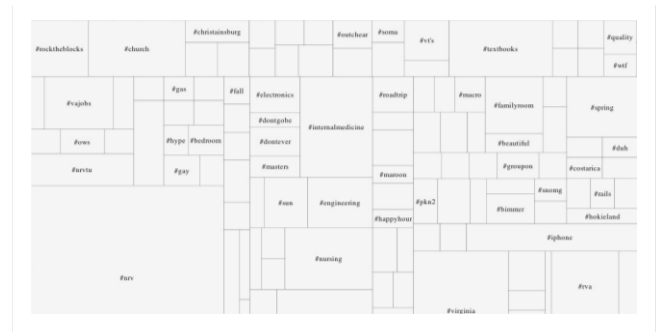


Figure 1: Treemap Visualization

The goal of the second visualization, Connected Graph (Figure 2) is to understand the connections between the people and their interest. This

visualization also uses the data from the tweets collected in VTS. In this visualization there are more aspects in consideration than in our previous visualization, we are working with frequencies of appearance as in the Treemap visualization, but also we are each of the tweets to each other.

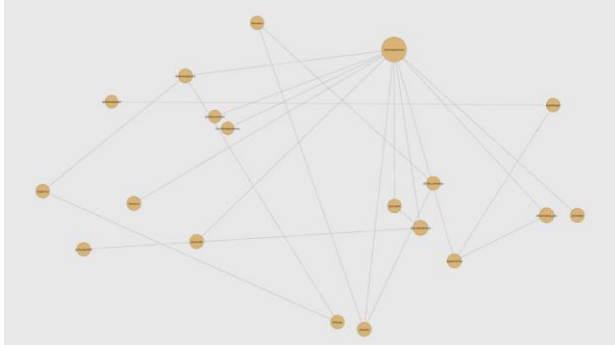


Figure 2: Connected Graph Visualization

The way we did this was with a system of nodes and edges. The nodes are “the whom, the what and the where” of the data, in other words the #hashtags and @mentions of the tweet data. The edges are the connections that indicate relation with two or more nodes. In the graph the connections are determined other mentions. For instance if a tweet includes #Blacksburg and #jobs there will be a connection between two nodes, one for #Blacksburg and one for #jobs. With the plot produced, the user will be able to explore the connections between #hashtags and @mentions as well as the frequencies each of one has.

Implementation

we used the Processing programming language for the implementation of the visualization. Processing is a programming language and development environment created to serve as a software sketchbook and to teach computer programming fundamentals within a visual context at first. This language later evolved into a development tool for professionals. Being the open source language that it is today, there are tens of thousands of students, artists, designers, researchers, and hobbyists who use Processing for prototyping, and production.

Evaluation

In order to test our product we presented both visualizations to the partners for the VTS. The partners have been active contributors during the

process of enhancing the VTS experience by presenting different points of view. The partners group consists of town leaders, community activist and local business owners.

The feedback from group was mostly positive. The feedback on the Treemap visualization indicated that it was overwhelming. One of the partners said “it’s too much information at one time”. One reason for this comment is that we presented all the tweet data in one view. One fix for this problem is to use time frames. This way you will be able to see the Treemap from a specific day or perhaps a month, giving more control to the user and reducing the complexity of the graph. Once this is done we plan to present the visualizations to focus groups to get some more feedback.

Conclusions

Our work in visualization development for the Virtual Town Square has all but just begun. We have developed two main visualizations at this point. Both of them will allow the user to have an active and efficient view of the social media surrounding the New River Valley community. The next step would be to integrate other social network visualizations this way having a complete perspective of what the people are concerned about. The main challenge we face is creating a program people would use on a daily basis, without feeling overwhelmed by the amount of information or complexity of the structure.

References

- Seungwon Yang, Andrea L. Kavanaugh (2011) "Collecting, analyzing and visualizing tweets using open source tools." In dg.o '11 Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times, pp. 374-375. ACM New York, NY, USA.
- Manuel A. Pérez-Quñones, Andrea Kavanaugh, Uma Murthy, Philip Isenhour, Jaideep Godara, Spencer Lee, Alain Fabian (2007) "VizBlog: a discovery tool for the blogosphere" In dg.o '07 Proceedings of the 8th annual international conference on Digital government research: bridging disciplines & domains, pp. 314-315. ACM New York, NY, USA.

Wohn, D. Y., & Na, E. K. (2011). Tweeting about TV: Sharing television viewing experiences via social media message streams. *First Monday, 16*(3).