Introduction

Drupal, an open source content management system tool, has been gaining popularity in the past few years. There seems to be a strong community actively contributing and exploring what Drupal has to offer. A network analysis could yield interesting insights on what types of individuals Drupal attracts as well as how the Drupal community interacts on Twitter.

Data Selection

This analysis uses a Twitter dataset of 200 users whose tweets contain the word “Drupal” as an attempt to gather data that reflects how the Drupal community interacts on Twitter. The dataset was collected through the import feature “Import from Twitter Search Network” provided by NodeXL. Due to the huge amount of data on Twitter and the difficulty to collect an exhaustive set of data with NodeXL, only 200 rows of Twitter user information was imported. There was no additional manipulation on how the 200 users are chosen.

Analysis

Headline 1 - Drupal Community Not Well Connected on Twitter

The visualization was acquired through grouping the 200 Twitter users by clusters. Each vertex is color coded by the clustering coefficient and size coded by how many followers each user has. The intergroup edges are combined to provide an overview of the network connection relationships.

The graph shows that Twitter users who express interest in Drupal are not as well connected on Twitter as perceived from Drupal.org community activities. There are many people who are not following Twitter accounts like the drupalplanet and TheWeeklyDrop that provide news tweets on Drupal related topics. Moreover, an interesting finding revealed by the visualization is that it distinguishes the Drupal enthusiasts from the novices. The Twitter community on Twitter appears to be less connected compared to the Drupal.org community activities.
users in the left-most box are mostly freelance developers and job searching agencies that become aware of Drupal development and begin to participate. On the other hand, the clusters in blue or green include individuals who have been working on Drupal for years. Nonetheless, even though the Drupal enthusiasts are more connected as depicted by the triangular intergroup edges area, they are mostly relationships of following certain Twitter accounts that broadcast Drupal news. There does not seem to be a strong social connection in the Drupal community on Twitter.

Headline 2 – Key Connectors Not the Popular Ones

The graph below provides a cleaner view of the network after applying filter to the previous visualization. The intergroup edges are shown, and the nodes are filtered by the betweenness centrality.

The filtered graph shows a remarkable phenomenon that the individuals who have the most connections across different clusters are actually not the people with the most follower/following relationships. Other than the generic accounts like drupalplanet that is maintained by the Drupal.org community group, users like ToddZebert and JohnAlbin, who have only about 2000 to 3000 followers, are the ones that are most connected with different twitter user groups. Active users who have more than 10000 tweets and followers like RaaVi and kristiancarter are not well connected in the Drupal network. However, such occurrence may not be representative as the dataset is not complete. This is worth further investigation.

Headline 3 – Drupal Attracts World-Wide Attention

The visualization below shows the same dataset of 200 Twitter users but grouped by the Twitter accounts’ time zone. The nodes are still color coded by clustering coefficient and size coded by the number of followers each user has. The intergroup edges are combined to depict an overall relationship among the groups.
The combined edges show that the world-wide Drupal network seems more linked. Several clusters have strong connection with people across the continents. For instance, the Twitter users in the light blue cluster are mostly located at California. One user in the group, ToddZebert, the key connector identified previously, is connected with people from France, Australia, Taiwan, and so on. Viewing the Drupal network in this perspective demonstrates that U.S. web developers are not alone in the Drupal hype.

**NodeXL Critique**

Overall, NodeXL provides many useful functionality for network analysis. The concept of displaying data from excel sheets is easy to understand. Features like the “Layout Options” and “Autofill Columns” require little learning effort. However, the experience of using NodeXL has been somewhat exasperating. Other than the known frustration of not being able to redo/undo action and the difficulty of interacting with a large amount of data in a limited screen space, some of the features that would have been helpful are the ability to refresh or change layout for only the selected nodes and the option to resize the boxes for the group-in-a-box layout feature. These features would allow users to make better use of the space given to organize the nodes and reduce edge crossing. In addition, it would be helpful if the initial “Dynamic Filters” box could be bigger. With the current display, it may take users some time to realize that there are more filter options available after scrolling down. Nevertheless, NodeXL is an intuitive and light-weight approach for users who are familiar with Excel to analyze data without investing too much money and time.