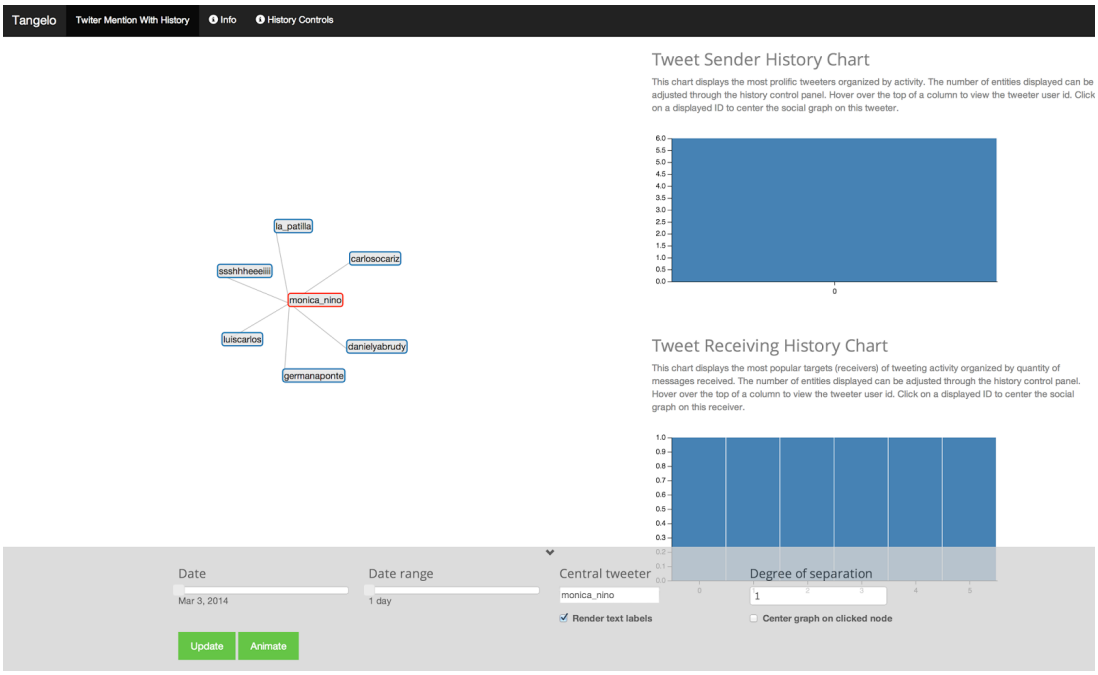
Using the Twitter Mentions Application

XDATA Program
Kitware Team
June 2014

This application is currently hosted at the following URL: http://10.1.98.108:8080/twitter-mention-history/. Upon opening the page in your browser, you should see page with three visualization panes and a configuration panel.

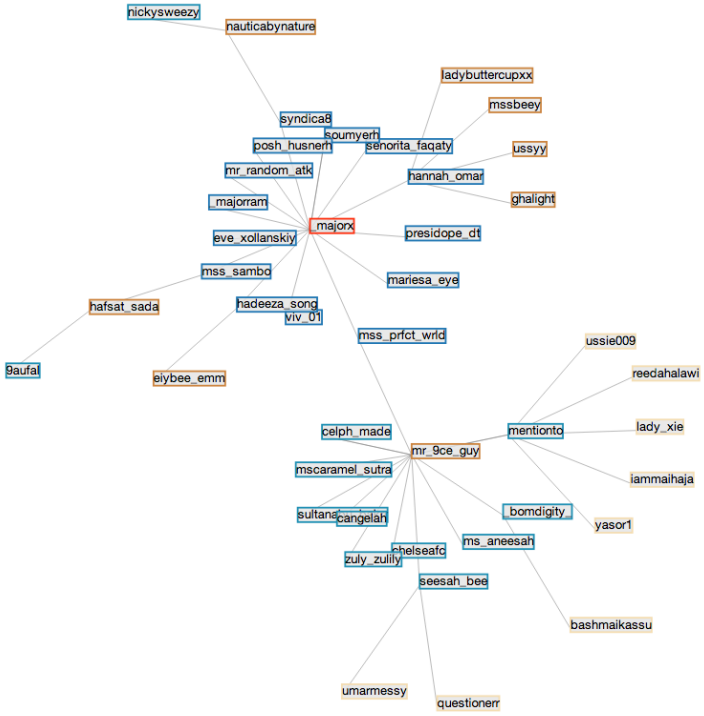
In the example below, the graph on the left shows a tweeter outlined in red and all users to whom this tweeter has sent messages to or received messages from, outlined in blue. On the right are two histograms showing the number of tweets sent (top) or received (bottom) by each user that is displayed in the left-hand graph.



The Network Graph

To perform new queries from the Twitter dataset, set the desired start date (with the slider entitled “Date”) and then set the “Date range”, which specifies how long a snapshot of social network activity should be investigated. After these controls are adjusted, then click the green “Update” button in the lower left corner of the interface. The twitter database is immediately queried and a new social network graph, centered around the selected tweeter using the supplied date ranges, is built dynamically and rendered.

A larger example of the network graph is shown in the figure below. The central tweeter is shown in red surrounded by first blue, then tan nodes, depending on the distance from the central tweeter. As the network distance from the central tweeter increases, the alternating blue and tan colors become gradually less saturated to indicate the increasingly distant connections.

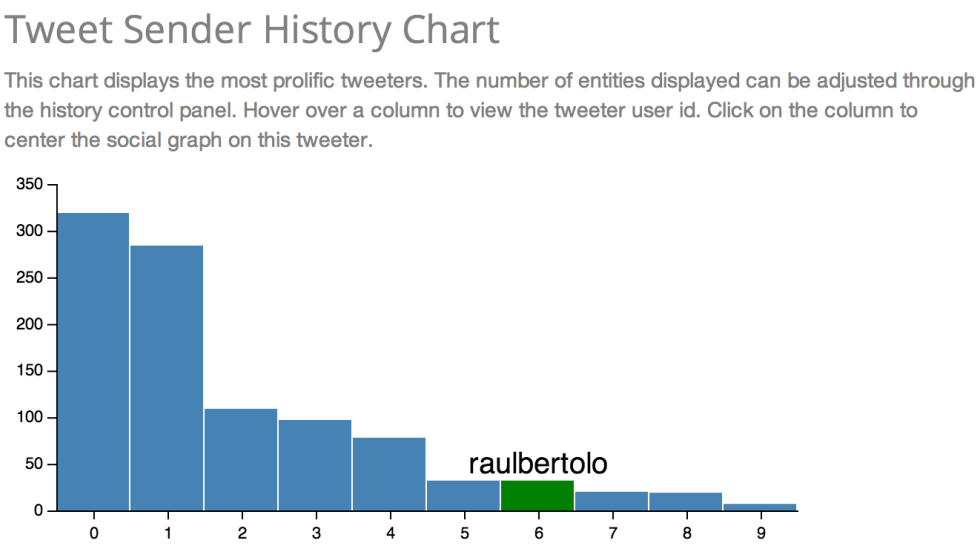


The History Graphs

As previously mentioned, the temporal network graph displays only communications that occurred during the time span selected by the UI controls at the time of the last query. In contrast, long term views of activity are provided by the bar graphs along the right side of the interface. These two graphs accumulate the history of all tweeters displayed since the beginning of an analysis session. For the purposes of our analysis, the social network is divided into message senders and receivers. Since Twitter consists of only public broadcast messages, we define a receiver as someone who has been addressed by name in the content of a tweet. For example, if user ID *bill* sends the tweet "@james, I will meet you at five." Then *bill* is considered a sender and *james* is considered a receiver.

The top graph list all message senders that have ever been present in the network graph and orders them according to the number of communications they generated, starting with the most active at the left and progressing in order of decreasing activity. Each time a new query of Twitter data is performed, the set of observed senders and the respective message counts are

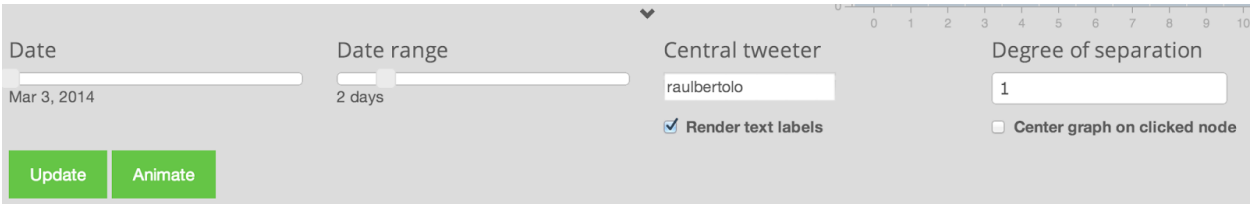
updated. The count of messages is displayed on the vertical axis to the left of the bars as shown below.



The history graph can be explored by hovering the mouse over any of the columns of the histogram and the Twitter account corresponding to the bar will be displayed above it. By clicking on the column, a user can elect to "explore this entity further". The node link graph on the left will be updated to *center* on the selected tweeter by executing a new query using the current date criteria. Therefore, the user can use the history graphs to move through the social networks and follow interesting entities simply by clicking on the entity's bar in the history chart.

The top of the two history charts show the most active senders, the bottom chart performs identically but display the most active receivers instead.

You can raise or lower the control panel below by clicking on the arrow above it  . The control panel contains the parameters useful for controlling how to explore the social networking aspects of the Twitter dataset.

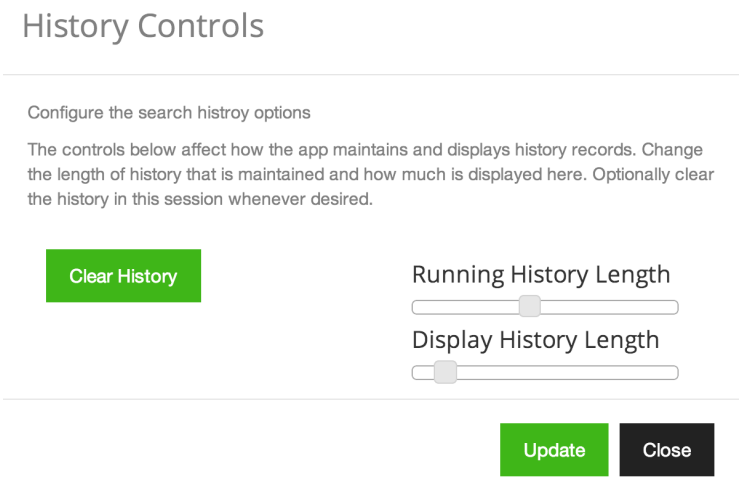


Again, the date slider controls which date is displayed in the node link graph, while the date range slider determines how many days of activity are considered. The "Central tweeter" box controls which user the graph is centered on and will be updated automatically as you click on elements in the visualization panes.

Toggling the “Render text labels” box will switch between the default graph view with the user names embedded to a simplified view showing only circles for each tweeter. By increasing, the number in the “Degree of separation” box, the application will query the tweets sent by the receivers as well. This allows you to explore the complex network of messages sent visually. As the graph becomes larger, you can click and drag an empty area to scroll the view over areas outside of the view pane.

History Controls

Clicking on the “History Controls” button on the top of the window will open a new pane with additional configuration to control the number of messages that are displayed on the history charts as well as the number of that are cached in memory. Optionally, the user can manually clear the history to reset the data. The user can clear the history at any time he or she chooses. The top slider, entitled “running history length” contols how many query iterations an entities record of being discovered states around in the local history “memory”. When an entity’s last appearance was longer ago than the running history length, the entity’s history is removed.



The ‘display history length’ slider does not affect how the entities are stored in the history mechanism, but it does control how many are displayed for rendering. Through use of these two controls, a number of different investigative scenarios can be pursued using these simple controls.