Will Webberley
Supervised by Stuart Allen & Roger Whitaker

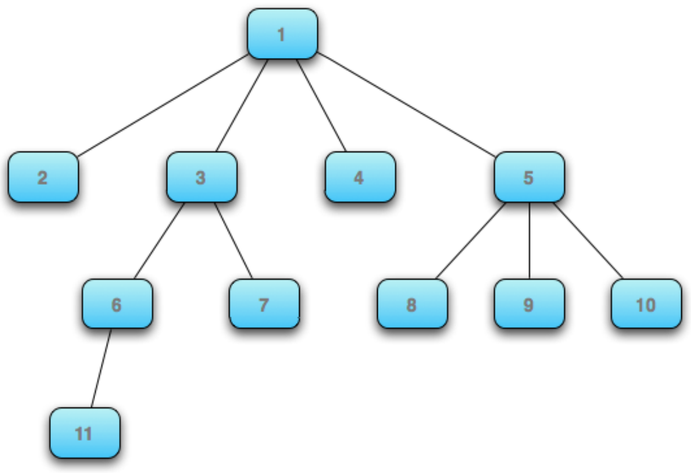
Cardiff School of Computer Science & Informatics
Poster Day 2011

Retweeting: Information-Forwarding in Twitter

What's happening?

The use of Twitter as a **microblogging** site has grown rapidly over the past few years. Users are able to **share** information with others by **retweeting** posts.
Retweeting is effectively the **forwarding** of a message from one user to others. It can be seen as a form of **voting** (or **filtering**) mechanism as users will usually only retweet other's posts that are **interesting** or useful to them. The retweeting user will have (in some cases, vastly) increased the size of the **audience** for the post.
I believe that this filtering should be taken further so that users only view the information they actually **want to see** via some form of **tweet-retrieval system**.

Timeline @Mentions **Retweets** Searches Lists



Example of a retweeting tree. Nodes represent users at different points of the retweet chains. These chains are illustrated by each of the branches. The user at the root is the original tweeter. This retweet has had 10 total retweets down 7 different chains. the longest path-length is 3. Users 1, 3 and 5 have had the most retweets and so can be seen as more influential - maybe by having more followers.

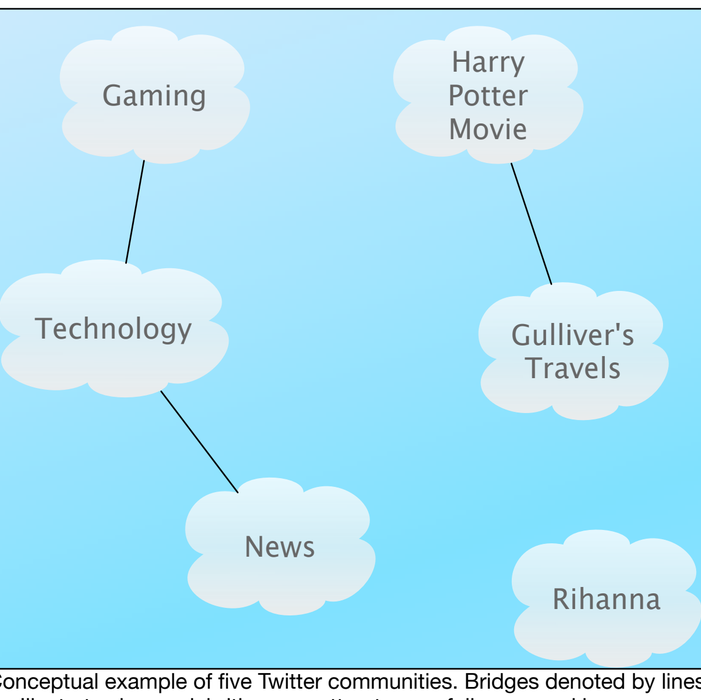
1. The Retweet Structure

As more and more users retweet a post - including retweeting a retweet of a post - a retweet **chain** emerges.

A map of the different paths a retweet chain can take can be illustrated by a retweet **tree**.
The **path-length** of a retweet is defined as the length of a particular chain.

Experiments have demonstrated that the longest path-lengths are of around 6, and some tweets can be **retweeted hundreds of times** in total. Typically users with more followers are more likely to have their posts (and retweets) retweeted.
A single post can therefore **propagate** quickly through the Twitterspehere.

As people keep posting and retweeting, groups of users - or **communities*** - of users with similar interests can develop. Members of communities tweet often about the given **topic**, and have many friends and followers of the same community.
Tweets can therefore be retweeted between communities via **bridges** - users with friends and followers in more than one community.



Conceptual example of five Twitter communities. Bridges denoted by lines. Illustrates how celebrities may attract many followers and become a community themselves, and acute communities around movies and news stories.

2. Why Are Retweets Important?

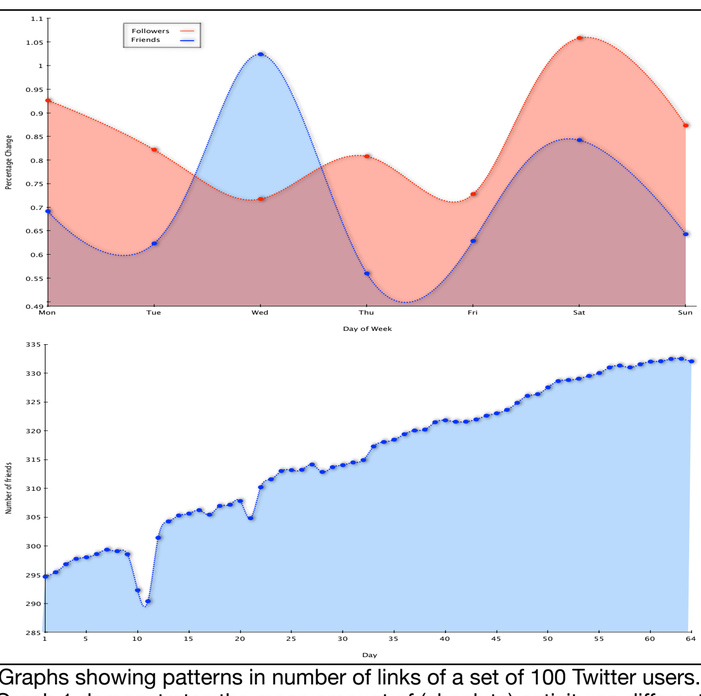
People use Twitter as a means to acquire information they are **interested** in. Twitter enables users to **elect who they follow** (and, to an extend, who follows them), thus giving them control over the information they receive (a primitive information-retrieval system). Not all the information from friends is always interesting or useful, however.

What's needed is a more effective **information-retrieval system** in which a user can view (nearly always) **only** the posts they are interested in. This may involve the viewing of information from people who aren't friends of the user (perhaps through retweeting).

The **quality** of the retrieved information can be partly based on;

- users currently followed by the user
- **content** of tweets retweeted by the user
- frequently viewed or retweeted topics
- frequently used (or viewed or retweeted) **hashtags**
- how frequently users add and remove friends.

#posterDay2011



Graphs showing patterns in number of links of a set of 100 Twitter users. Graph 1 demonstrates the mean amount of (absolute) activity on different week days and Graph 2 shows the change in mean number of friends over a 64-day period.

3. Next: Analysis of Friend/Follower Graphs

The next stage is to study the **graph formed from the links** (friend/follower) between users, and the dynamics of these links. The more friends a user has, the more information the user will be exposed to.

Preliminary results have shown roughly a **0.2% average daily increase** in total number of friends - showing users tend to want to widen their information input. If this rate is greater than the rate of users joining Twitter, will all users eventually be linked?

A user changing his/her friends usually indicates they want a **change in the information** they receive and the frequency that this changes shows how often the user changes what he/she wants to view.

Some users have many followers and post frequently, whereas others have few followers but many friends. These users are therefore **information providers** and **information consumers** (i.e. information-retrieval becomes more important) respectively

Any effective information-retrieval system must take into account the **changing views** and **interests** of the user, so understanding the patterns and their frequencies are clearly vital.

*Java A, Finin T, Song X, Tseng B. *Why We Twitter: Understanding Microblogging Usage and Communities.* (SNA-KDD Workshop 2007).