# Overview of Code for Influence Evolution Prototype

March 20, 2013

## 1 Calculating Buzz and Influence

### 1.1 R Code for creating Transition Matrix.

The initial transition matrix is created from a Dirichlet distribution using the `create-trans-from-dirichlet.R` R file.

Once executed this creates the four state transition matrix for the Buzz model. The $\alpha$ parameters of the Dirichlet distribution are adjusted according to our prior knowledge of the transition likelihood. Run `createTransFromDirichlet(out.file)` where `out.file` is the path of the output transition matrix file.

### 1.2 MATLAB Code for creating Emission Matrix

**Note:** these scripts and function require the aggregated weekly observation count data as well as the users those counts refer to. As per the Twitter developer rules of the road [1] user data cannot be made available as a service. Developers must also ensure not to surprise users. Due to these two rules, only anonymised data is available via request from the author [2]. The required input files are outlined in table 1.

To train the Buzz HMM run the following Matlab script: `createAndTrainBuzzHMM.m`. Before running this you may need to change the `dataPath` variable in `initialiseBuzzModel.m` to point to the directory containing your received data files.

[1] https://dev.twitter.com/terms/api-terms

[2] d.simmie@imperial.ac.uk

| | |
|---|---|
| `buzz_model.csv` | `mentions_per_week.csv` |
| `network_nodes_table.csv` | `retweets_per_week.csv` |
| `tweets_per_week.csv` | `unique_interactions_per_week.csv` |

Table 1: Influence Relevant Correlations

### 1.3 R code for analysing the temporal output and producing the influence ranks

To determine the absolute influence rank and temporal value we switch back to R. Set the `data.path` variable in `load-influence-data.R`. Afterwards run `source('create-influence-data.R')`. This will produce the complete buzz and influence output files in the data directory. After this script runs you will have four files in your data directory which are the output files for the model run: `buzz-temporal-output.csv`, `complete-buzz-rank.csv`, `complete-influence-rank.csv` and `influence-ranked-temporal.csv`.