

Big Data Velocity,
or Where and Why is this Data Coming so Fast
Searching twitter for: realdonaldtrump

Tidewater Big Data Enthusiasts
Chuck Cartledge
Developer

Monday 11th April, 2016 14:12

Contents

List of Figures	i
List of Tables	ii
1 Introduction	1
2 Local results	1
3 Global results	3
A Misc. files	8

List of Figures

1	A sample connection graph started from one tweet.	4
2	A Pareto graph of connections from the sample graph.	5
3	A connection graph including all discovered tweeters.	6
4	A Pareto graph of connections including all discovered tweeters.	7

List of Tables

1	Most well connected tweeter of sample tweets.	1
---	---	---

1 Introduction

This is an automated report generated by searching twitter using the search token:

realdonaldtrump

2 Local results

An unattended search on twitter as made for the search term; realdonaldtrump. The results are reported (see Figures 1 on page 4 and 2 on page 5). Information about the 30 most connected tweeters was also collected (see Table 1).

Table 1: Most well connected tweeter of sample tweets. The entries are ordered by the number of edges that are incident to each node. Only 30 most connected nodes are shown.

#	Edges	Name	Description	Location
1	9,211	1flyc	Motivational Use Only: Psalms 51:1-3 Took A Couple Losses, Nun I Can't Get Back, I'm Still Alive!.. #iJWTM	Countin'Up
2	1,162	02LowNoma	Addictive Affairs-Indy	Indiana
3	417	00bonton		
4	289	03cammie	Just because u cant see air dosnt mean to stop breathing and just because u cant see god dosnt mean to stop beliving!!!! FOLLOW ME!!!	
5	148	_032149073445	straight as a door knob strictly dickly here to just social an meet people	nc
6	111	byron36501663		
7	23	MicahRapier	Bitter clinger.	Modesto, CA
8	8	Drake	Views From The 6	Paradise
9	8	LilTunechi	IM YOUNG MONEY http://t.co/wMwkyzCu	Mars
10	8	wizkhalifa	Taylor Gang	Pix Burgh

(Continued on the next page.)

Table 1. (Continued from the previous page.)

#	Edges	Name	Description	Location
11	6	ArianaGrande	dangerous woman	Honey-moon ve.
12	6	BarackObama	This account is run by Organizing for Action staff. Tweets from the President are signed -bo.	Washington, DC
13	6	jtimberlake	The Official Twitter of Justin Timberlake	Memphis, TN
14	6	katyperry	Growing...	
15	6	KevinHart4real	My name is Kevin Hart and I WORK HARD!!! That pretty much sums me up!!! Everybody Wants To Be Famous But Nobody Wants To Do The Work	Philly/LA
16	6	MeekMill	I just wanna see the sun shine tomorrow #dreamchasers #hustler	Killadel- phia
17	6	Tip	#HustleGang #GrandHustle #BankRollMafia IG: @Troubleman31 https://t.co/HlSoeEz7Rm kingofdassouthbookings@gmail.com	Atlanta, GA
18	4	0208Luna	lang not English, data ignored	lang not English, data ignored
19	4	aliciakeys	Passionate about my work, in love with my family and dedicated to spreading light. It's contagious!;-) http://t.co/OxWhtNsqoE http://t.co/yEkbY8wWEz	New York City
20	4	AvrilLavigne	Get 'Avril Lavigne' on iTunes now: https://t.co/m8yd5aoLeN	
21	4	bigtimerush	Kendall Schmidt, Carlos PenaVega, James Maslow and Logan Henderson. Download our album 24/seven here: http://t.co/gy64F46aJR	

(Continued on the next page.)

Table 1. (Continued from the previous page.)

#	Edges	Name	Description	Location
22	4	britneyspears	It s Britney Bitch! Get #PieceOfMe tickets now! https://t.co/wmqgzcAxcc	Los Angeles, CA
23	4	ddlovato	New album #CONFIDENT out now!!! http://t.co/yBrykn2kbK	DAL-LAS/LA
24	4	Eminem	Southpaw: Music From and Inspired by the Motion Picture: http://t.co/T1Z3XZyGb7	Detroit
25	4	iamwill	i.am...i.can...i.will	
26	4	iCarly	This is the official Twitter for iCarly!	Seattle, WA
27	4	jennettemccurdy	gimme a break, I'm just trying to figure everything out here.	
28	4	justinbieber	Let's make the world better. Join @bkstg and add me on @shots 'justinbieber'. OUR new single COMPANY out now. OUR new album PURPOSE out NOW	
29	4	kendricklamar	untitled unmastered. https://t.co/CQTjn8lblS	CPT — Booking: mgmt@txdx.com
30	4	ladygaga	jazz art pop punk actress fashion magazine editor and columnist!	

(Last page.)

3 Global results

The local results have been incorporated into a growing global picture (see Figures 3 on page 6 and 4 on page 7).

21160 unique nodes, 28677 edges

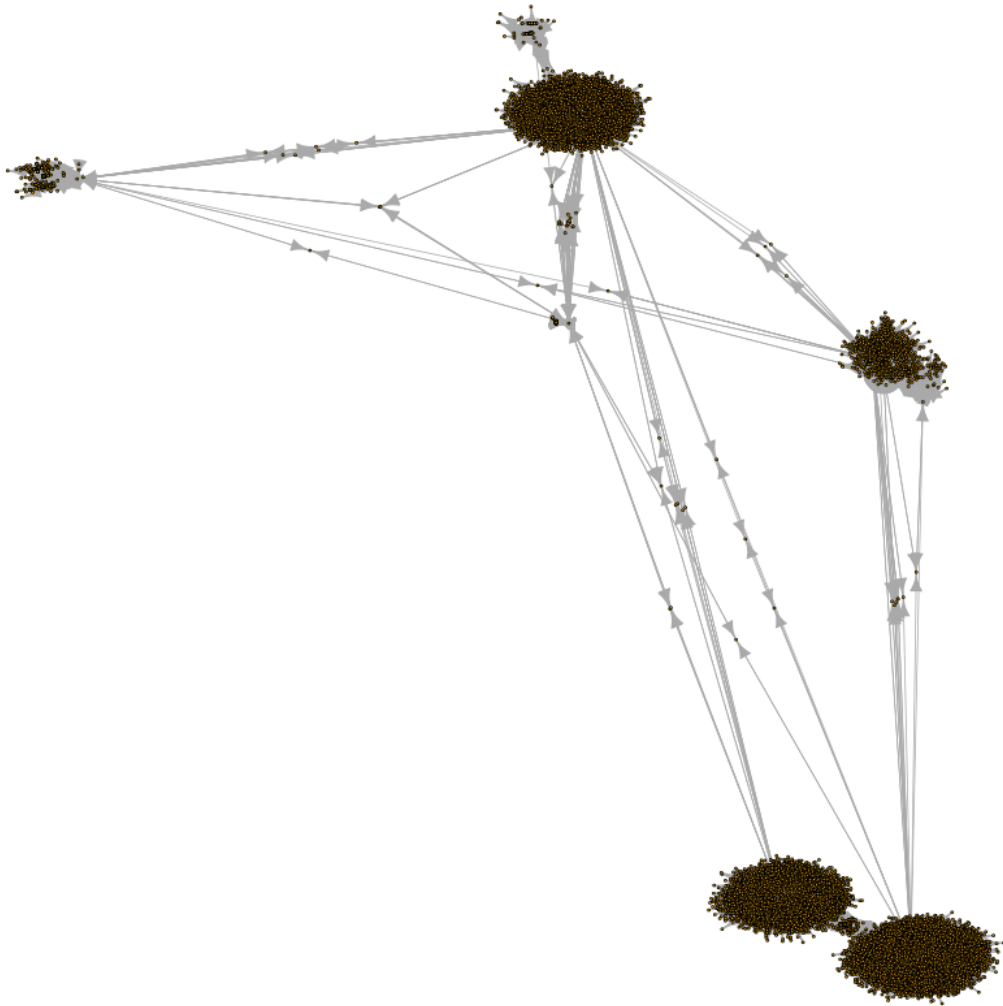


Figure 1: A sample connection graph started from one tweet.

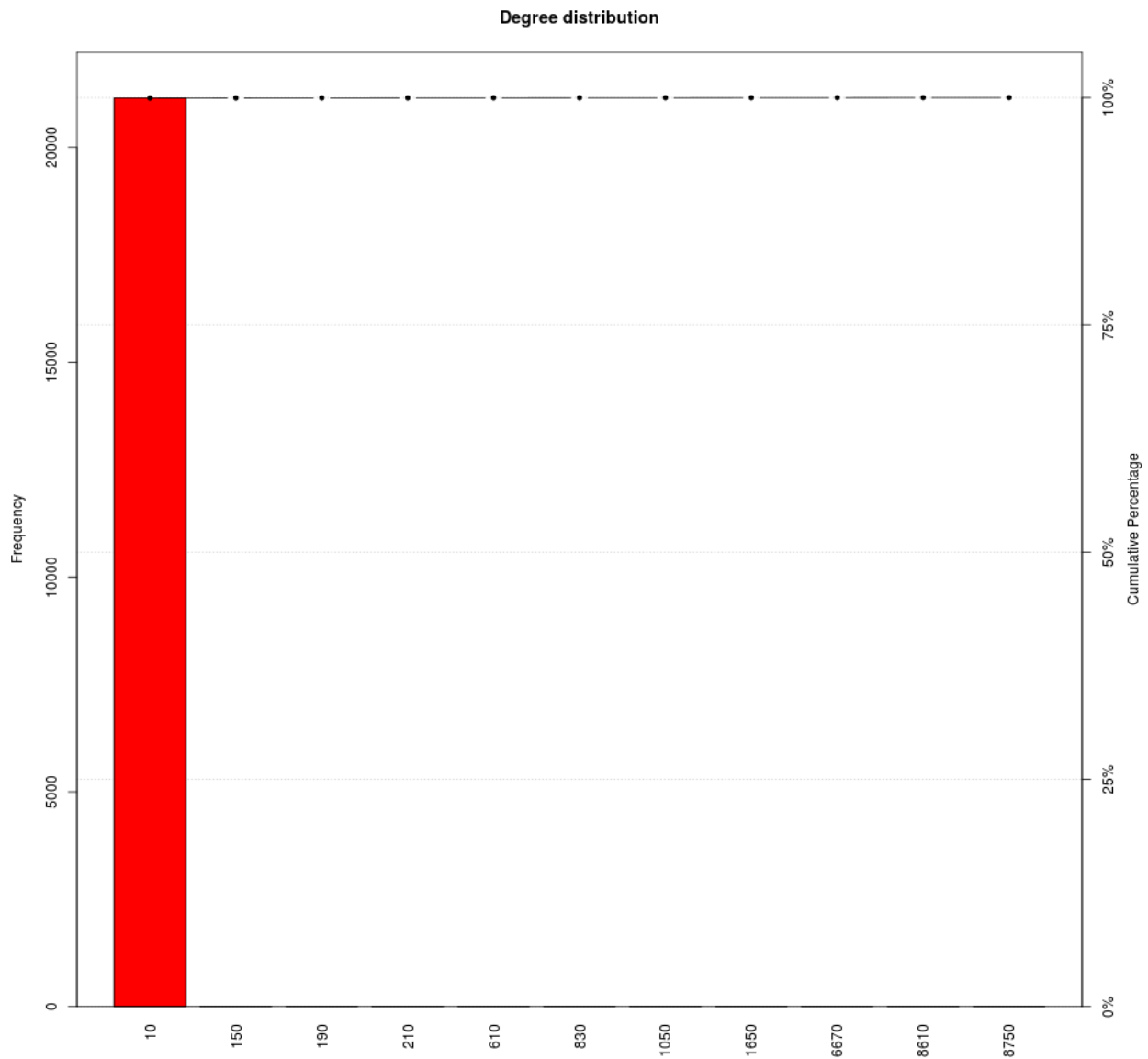


Figure 2: A Pareto graph of connections from the sample graph.

252358 unique nodes, 458371 edges

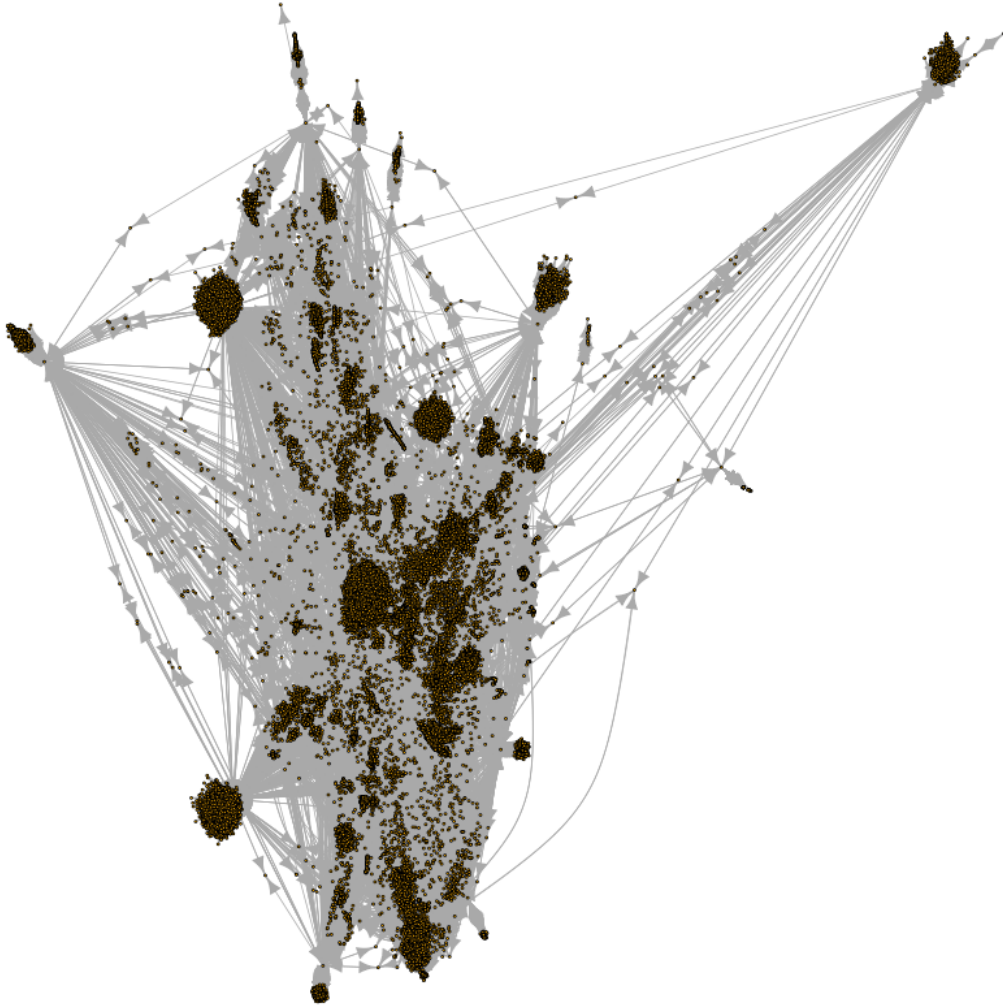


Figure 3: A connection graph including all discovered tweeters.

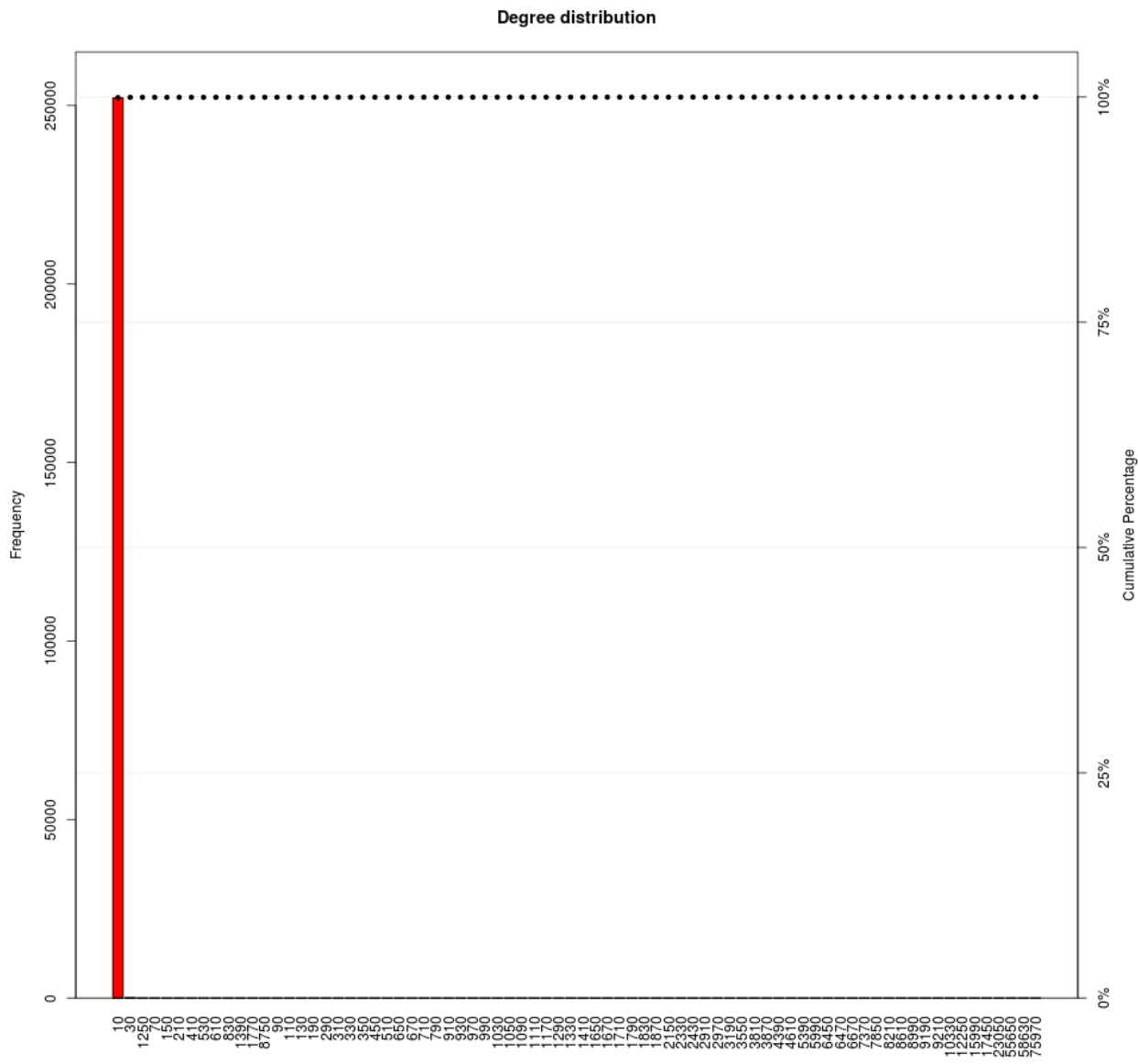







Figure 4: A Pareto graph of connections including all discovered tweeters.

A Misc. files

These are the files that were used to create this report:

1. Makefile - the makefile that orchestrates the R files used to capture the raw data and create this report 
2. hashTables.R - an R support file used by the test.R program 
3. myLibrary.R - an R support file used by the sampleMaker.R program 
4. samplemaker.R - an R program used to create the \LaTeX file used to create this report 
5. test.R - an R program used to discover and record connections (uses a postgres database) 
6. twitterLibrary.R - an R support file used by the test.R program 