Seminar Paper

Advanced Topics in Social Network Analysis

Lecturer: Dr. Tsahi Hayat

Twitter and Viral Trends

Submitted by:

Daniel Arom

22.9.15

Table of Contents

Introduction.....	3
Related Work.....	5
Research Question and Hypotheses.....	7
Methodology.....	8
Results.....	10
Discussion.....	12
Conclusion.....	13
References.....	14
Appendixes.....	15

Introduction

Ever since the introduction of social networks in the early 2000s information and communication of ideas have drastically changed in their framing and mediums and also the delivery method between the sender and intended and unintended receivers. Social networks have many distinctive characteristics that demarcate them from SMS (Short Message Service), VOIP (Voice Over Internet Protocol) and the ordinary voice phone call.

Early social networks prior to Facebook, Twitter, Instagram and LinkedIn included MySpace and Xanga. MySpace which was founded in 2003 was early adaptor to the fundamental concepts we know today like networking, personal framing online and sharing that undermine every social interaction offline and online. Xanga pioneered the highly customizable profile which gave users the option to customize practically any element on their page by accessing the HTML (Hyper Text Markup Language) code of the visual elements they could modify to their own liking and personality, this however required a lot of programming skills and was a set back for the social network's aim of customizability.

In our current state and time in the western world in academic, work and extracurricular activities nearly all of the members interacting are on one social network, if not all that are mentioned. Facebook serves as a social marketing tool for some of its users that frame their lives in a pseudo-happy frame and some let out their problems and grasp Facebook as a personal journal without the confidentiality of a paperback journal. Instagram is visually based social network that aims to extrovert one's creative skills of being a smartphone photographer. Instagram includes tools that help your photography skills and even provide a toolbox of premade filters set to provide digitally taken photographs with an old-school visual of film photography. LinkedIn lets business workers, owners, seekers and entrepreneurs connect to one another in a highly formal manner, something that isn't found in Facebook, Twitter or Instagram.

This research takes the social network Twitter as its content matter due to its significant uniqueness of not being a visually based social network that is prevalent today. Twitter has been known for its straight message transmissions from the sender

to the receiver. On Twitter everyone can interact with everyone from celebrities and politicians to farmers and students. This research aims to answer the concept of correlations and trends between Twitter users and their social network scope. Being branded as the “social network for smarter people”, Twitter is relevant to the intellectual property that composes the electronic knowledge base. Twitter is a witty and smart social network; users have 140 characters or less to compose their message.

In the below section we present the aimed research questions and likely outcomes of the research known as the hypotheses. Afterwards the methodology used to collect the examined data is presented and our results post-data collection. The conclusion of the paper provides a discussion, conclusion and limitations of the aforementioned study.

Related Work

Twitter can be looked upon as a microblogging community. Twitter's users have been observed as sharing their thoughts and seeking out community (virtually). Investigating this phenomenon was found that most posts on Twitter are about a user's daily routine (Finin, Java, Song, Tseng, 2007). Twitter offers distinctive ways of communication that are mainly text based. A Twitter user interested in the statuses of another user signs up to be a "follower" (Arlitt, Gill, Krishnmuthy, 2008). In the latter, the researchers dissected micro-communities quantitatively and found that the follower to following status of the examined accounts are nearly identical with a 1:1 ratio. They concluded their research with the fact that the "Twitterverse" is filled with the usual non-celebrity people that broadcast but celebrities with Twitter accounts are distribution hubs with an uneven ration of followers to following given their popular status in the real world. Twitter serves a word of mouth on the social network sphere, if something tragic or positive happens that is crucial to society, variability of the news flourishes through Twitter.

In modern times Twitter can experimentally serve a source for natural phenomena like earthquakes. Several Twitter accounts report earthquake occurrences. Some examples are that the United States Geological Survey (USGS) feeds tweets on world earthquake information, but such information is not useful for prediction or early warning (Matsuo, Okazaki, Sakaki, 2010). Ever since the publication of this article Twitter has been a source for event evacuation like the Haiti earthquakes and people have used the medium to save many human lives.

Word of Mouth is a concept that caused virality before the concept was viewed on a screen and plays an important part of social life. This broad reach of eWOM (electronic Word of Mouth) provides consumers with tremendous clout to influence brand image and perceptions (Reynolds, 2006; Urban, 2005 in Zhang, Sobel and Chowdury, 2009). Their findings were that at a ratio of 50:30 tweets of good to bad messages when dissecting tweets that dealt with a product or service. Another interesting find was that 39% of tweet examined had a mention of a brand or organization a crucial factor for marketers using social networks to push their product on consumers. Twitter also elevates academic discourse and attained grades by

students if used in the correct manner with the lecturer and students. This study provides the first piece of controlled experimental evidence that using Twitter in educationally relevant ways can increase student engagement and improve grades, and thus, social media can be used as an educational tool to help students reach desired college outcomes (Junco, Heiberger and Loken, 2010). While Twitter serves as a breeding environment for ideas (Wellman et al., 2014; Dimitrova et al., 2013; Hayat, Samuel-Azran & Galily, 2016), learning (Hayat & Mo, 2015; Mo, Hayat & Wellman, 2015; Hayat & Lyons, 2010) and sometimes hate, an underlying concept must be reviewed. Retweeting, sharing another user's tweet to your follower thus creating virality. Additionally, celebrities and other highly followed users are in a particularly good position to broadcast content for social action (Boyd, Golder, Lotan, 2010). While Twitter equates followers to friends some researchers criticize the company's premise. This view [of the researcher's experiment] should be tempered by our findings that a link between any two people does necessarily imply an interaction between them (Huberman, Romero, Wu, 2008). Twitter has become a source of news many journalist practices today for cross referencing allocated news. Journalists embrace the concept of social media more than they enact the practices (Larcisy, Avery, Sweetser, Howes, 2009). In this 37% of the journalists whom participated used Twitter and that number has been exponentially growing until today. When celebrities for instance tweet after the 1000th retweet virality happens like wildfire. A closer look at retweets reveals that any retweeted tweet is to reach an average of 1000 users no matter what the number of followers is of the original tweet (Kwak, Lee, Park Moon, 2010). Virality on Twitter is a double sided sword since spam (unwanted messages is always present on the internet.) Twitter has a strict spam protocol and it can demotivate users to be overly active, similar to over-liking posts on Facebook. We show that a simple user ranking scheme that penalizes users for connecting to spammers can effectively address the problem by disincentivizing users from linking with other users simple to gain followers (Ghosh, S., Viswanath, B. Et al, 2012).

Research Questions and Hypotheses

Prior research has shown a vast amount of interest on what makes Twitter work and why but lacks a focus on the characterizations of its users. Another void research has shown is that studies do not focus on hash tagging. Hashtag are a pound value '#' on twitter and they work just like Google searching for a term just within Twitter's own database of public profiles. Hashtags are the reason topics trend and reach public spheres and traditional news media and sometimes presidential commenting on an issue. Some prevalent hashtags in history have been #BlackLivesMatter, about the Ferguson injustice and #IsraelUnderFire, about Israel's war on Gaza in 2014. These concepts in the Twitter agora lead to this researches research questions which are: Research Question 1: Does the use of positive or negative hashtags lead to a higher or lower number of favorites and retweets? Research Question 2: When news or trends which gain virality on Twitter with the use of hashtags do they correlate to a certain user's popularity on Twitter?

With the above questions their suggested answers are hypothesized in this research as follows. Research Question 1's educated answer or hypothesis: Positive hashtags on a given subject will lead to more Retweets, while a negative hashtag on a negative subject will lead to more favorites, not retweets. Research Question 2's educated answer or hypothesis: The relevancy of hashtag effects the virality regardless of popularity.

Methodology

This research's methodology was a short process of coding tweets (status updates on Twitter). The subject for data collection was Kim Kardashian. Kardashian is a reality TV star and she became well-known due to an adult film that was released of her, like the similar case of Paris Hilton. Today, Kardashian is a media figure with clothing lines and various products like applications on smartphones, make up and children's dolls. Kardashian has 35.4 million followers on Twitter, 46.5 million followers on Instagram and 2.2 million likes on Facebook. To say the least Kardashian rules social media and has more influence than classical news outlets via the medium of television. For instance, she visited Armenia and brought attention to the Armenian genocide to millions of her followers. Recently, she became the most followed account in the world on the social media platform Instagram.

This research paper's data collection was done in a qualitative manner dissecting Kardashian's latest tweets in chronological order of her public Twitter feed. The tweets were coded from the URL of: <https://twitter.com/KimKardashian>. The number of tweets that were examined were fifty in a latest to oldest manner in a down scale, the excel sheet with the researcher's data collection are attached to this paper. The instruments that were used were freeware software programs. The first software was a web browser by Google. A single website was used for the data collection which was <http://www.twitter.com/> within Twitter the researcher made a pseudo account account to obtain perks of being registered user and accessed Kardashian's tweets. Kim Kardashian's account was access by the URL of <https://twitter.com/KimKardashian?lang=en>. The data collection was entered into Microsoft Excel and statistical analysis was done with IBM's SPSS software at IDC Herzlyia.

The tweets that Kardashian sent out via her Twitter were measured through several categories that were predetermined to the collection of the metrics. The first category of the coding was named "Tweets" and was composed of a list of the the text inside of the tweet that was dissected, this list was made up of fifty tweets in total. The second category that was measured was the number of retweets that occurred on the specific tweet that was being examined. Retweets on Twitter are the amount of

times that user x takes user y's tweet and posts it on their own timeline. This effect of retweeting a tweet amplifies its viewership online, in some cases tweets which are retweeted many times reach viral statuses from users who are not celebrities. The second category in the coding scheme was "Number of favorites". In this category the researcher noted how many "Favorites" the certain tweet got by Kardashians follower base. Favorites on Twitter are equivalent to likes on Facebook and are noted by a star. The third category was composed of three sub terms "Positive", "Negative" and "Neutral". In this category if a tweet had positive connotation to the researcher's understanding it was marked down with the number 1 as coding. If the tweet had a negative connotation it was marked with the number 2. Lastly if a tweet was neutral, for example a tweeted picture or an emoji (modern day hieroglyphics) it was marked down as with the number 3.

The above measures are underlying categories that exist in every tweet and effect a tweets virality and consumption by the crowd that has engagement with the tweet.

Results

The results of running statistical analysis on Kardashian's tweets show many hypothesized traits that were forecasted in the Research Question section by the researcher. First and foremost an ordinal scale was used to analyze the data in a Spearman test. The highest correlation being 1.0 some results showed to be near close to it. It is important to note that sample size collected in regards to descriptive statistics was very small to begin with, fifty in number.

The results that were important to this research's premise are as follows. There was found to be a very small negative correlation between negative tweets (in their context) to Retweets, as hypothesized "Positive hashtags on a given subject will lead to more Retweets, while a negative hashtag on a negative subject will lead to more favorites, not retweets." Results showed that with more negativity leads to less popularity. The graphs clearly that with positivity comes popularity, once again confirming the research's hypotheses regarding vernacular inside of tweets. Another interesting discovery was that with more favorite to the tweets comes less retweets. By the assessment of this research the cause to less retweets are because the users' engagement with the tweet dealt with compassion and not a spreading of word action which is retweeting in the Twitter universe.

There was one neutral exception in it's classification category which got the most engagement in this study. This tweet was the naked picture of Kardashian with her current to this research's submission pregnant body, this tweet had close to thirty-five thousand retweets. Below is all the data output after analyzing the tweets.

Correlations					
			NumberOfRTs	NumberOfFav	PosNegScale
Spearman's rho	NumberOfRTs	Correlation Coefficient	1.000	.977 ^{***}	-.085
		Sig. (2-tailed)	.	.000	.556
		N	50	50	50
	NumberOfFav	Correlation Coefficient	.977 ^{***}	1.000	-.139
		Sig. (2-tailed)	.000	.	.337
		N	50	50	50
	PosNegScale	Correlation Coefficient	-.085	-.139	1.000
		Sig. (2-tailed)	.556	.337	.
		N	50	50	50

***. Correlation is significant at the 0.01 level (2-tailed).

Statistics				
		NumberOfRTs	NumberOfFav	PosNegScale
N	Valid	50	50	50
	Missing	0	0	0
Mean		4395.80	14915.70	1.46
Median		2256.00	10422.50	1.00
Mode		1837	9068	1
Range		33903	54656	2

Discussion

The preliminary research question that was an aim to answer was if positivity with a tweets composition play a role in the amount of engagement it receives. This question was answered within this research. If a tweet as composed by Kim Kardashian is positive, it will receive more favorites. Retweets on the other hand are much more scarce. The second question mentioned in this research was that if a person's popularity correlates to the user's popularity, the answer to this question is complex. Kim Kardashian which has the most followers nearly on every social network has a high engagement rate with her fans but not every tweets receives the amount of engagement as the previous. Too this the researcher would have recommended a longitudinal study that takes into account more attributes within her tweets while cross examining her social media to other media personas.

Arlitt et al's (2008) studies derive from the premise that Twitter is mainly a text based service, this premise is dated yet somewhat true. Twitter is based on a 140-character limit but many services plug into Twitter. Services like Instagram and Tumblr amplify themselves through Twitter which is the case in Kim Kardashian's case. Many of her tweets are actually Instagram pictures that are being shared via Twitter.

This work affects interactive media by the claim that a social network doesn't always work by its text-book. For instance, Facebook is becoming a search engine when it opened it just for friendly connections like MySpace and Xanga. Social networks are always changing they electronic organism to the eyes of this researcher. Books have been written on Twitter and countries have been freed from tyranny via the social media, it's in the hands of the users not the developers. This research had various constraints, the first being that only one single celebrity was targeted for analysis. Future studies should compare a celebrity to the usual folk who compose Twitter just like the population in the non-virtual world.

Conclusion

This research's importance is of vast significance to the study of new media within the academic world. This research premised that with a user's popularity and framing of a tweet popularity increases. The methodology used was a Spearman analysis of a select tweets and significant results were derived. The research's hypotheses were proved to be true by the results gathered by the researcher in this specific case.

References

- Boyd, D., Golder, S., Lotan, G. (2010). Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. *Proceedings of the 43rd Hawaii International Conference on System Sciences 2010*. 1-10.
- Chowdury, A., Jansen, J., Sobel, K., Zhang M. (2009). Twitter Power: Tweets as Electronic Word of Mouth. *Journal of the American Society for Information Science and Technology*, 60(11), 2169-2188.
- Dimitrova, D., Gruzdt, A., Hayat, Z., Mo, G. Y., Mok, D., Smale, L., Robbins, T., Wellman, B. & Zhuo, X. (2013). NAVEL Gazing: Studying a Networked Scholarly Organization. In Evangelos Kranakis (Ed.), *Advances in Network Analysis and its Applications*. Berlin: Springer.
- Ghosh, S., Viswanath, B., Kooti, F., Sharma, N., Korlam, G., & Benevenuto, F. et al. (2012). Understanding and combating link farming in the twitter social network. *Proceedings Of The 21st International Conference On World Wide Web - WWW '12*. <http://dx.doi.org/10.1145/2187836.2187846>
- Hayat, T., Samuel-Azran, T., & Galily, Y. (2016). Al-Jazeera Sport's US Twitter followers: Sport-politics nexus?. *Online Information Review*, 40(6), 785-797.
- Hayat, Z., & Lyons, K. (2010). The Evolution of the CASCON Community: A Social Network Analysis. *CASCON 20th Annual International Conference*. Toronto, ON, Canada.
- Hayat, Z., & Mo, G. Y. (2015). Advice giving and receiving within research networks. *American Behavioral Scientist*, 59 (5), 582-598.
- Heibergert, G., Junco, R., Loken, E. (2010). The effect of Twitter on college student engagement and grades. *Journal of Computer Assisted Learning*, 27, 119-132.
- Huberman, B., Romero, D., & Wu, F. (2008). Social networks that matter: Twitter under the microscope. *First Monday*, 14(1). <http://dx.doi.org/10.5210/fm.v14i1.2317>
- Java, A., Song, X., Finin, T., & Tseng, B. (2007). Why we twitter. *Proceedings Of The 9Th Webkdd And 1St SNA-KDD 2007 Workshop On Web Mining And Social Network Analysis - Webkdd/SNA-KDD '07*. <http://dx.doi.org/10.1145/1348549.1348556>
- Krishnamurthy, B., Gill, P., & Arlitt, M. (2008). A few chirps about twitter. *Proceedings Of The First Workshop On Online Social Networks - WOSP '08*. <http://dx.doi.org/10.1145/1397735.1397741>

- Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media?. *Proceedings Of The 19Th International Conference On World Wide Web - WWW '10*. <http://dx.doi.org/10.1145/1772690.1772751>
- Lariscy, R., Avery, E., Sweetser, K., & Howes, P. (2009). An examination of the role of online social media in journalistsâ. *Public Relations Review*, 35(3), 314-316. <http://dx.doi.org/10.1016/j.pubrev.2009.05.008>
- Mo, G. Y., Hayat, Z., Wellman, B. (2015). How Far Can Scholarly Networks Go? Examining the Relationships between Distance, Disciplines, Motivations, and Clusters. *Studies in Media and Communications*, 9, 107 - 133.
- Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes Twitter users. *Proceedings Of The 19Th International Conference On World Wide Web - WWW '10*. <http://dx.doi.org/10.1145/1772690.1772777>
- Wellman, B., Dimitrova, D., Hayat, Z., & Mo, G. Y. (2014). Networking Scholars in a Networked Organization. In Daniel Brass (Ed.), *Research in the Sociology of Organizations: Contemporary Perspectives on Organizational Social Network Analysis*. Bingley, UK: Emerald. Pp. 479-497.