Open Access

Open Journal of Big Data (OJBD)
Volume 2, Issue 1, 2016

www.ronpub.com/ojbd
ISSN 2365-029X

Conformance of Social Media as Barometer of Public Engagement

Songchun Moon [A, B]

[A] College of Business, KAIST, 85 Hoegi, Dongdaemun, Seoul, South Korea, moon@kaist.edu

[B] School of Computing Sciences, Newcastle University, Newcastle upon Tyne NE1 7RU, United Kingdom, songchun.moon@ncl.ac.uk

ABSTRACT

There have been continuously a number of expectations: Social media may play a role of indicator that shows the degree of engagement and preference of choices of users toward music or movies. However, finding appropriate software tools in the market to verify this sort of expectation is too costly and complicated in their natures, and this causes a number of difficulties to attempt technical experimentation. A convenient and easy tool to facilitate such experimentation was developed in this study and was used successfully for performing various measurements with regard to user engagement in music and movies.

TYPE OF PAPER AND KEYWORDS

Research paper: S*ocial media, user engagement, user preference, music, movie, Twitter*

1 INTRODUCTION

1.1 Background

Although there are a lot of wordings and hypes around how to analysis user behaviors related to the use of entertainment media, such as movies, music and online games, the status quo on basis of literatures [5, 13, 21, 22] so far is that we are now clearly in the stage of pursuing how to collect such data and from where they might be acquired. Furthermore, there is no standard tools that can be deployed for collecting such data.

Although it is commonly and generally believed that the sort of data could be obtained from public social media services, nowadays even how to access and how to crawl from the sources of Social Networking Services (SNS) are basically unknown or hidden to the people who are interested in big data. Another serious and impending problem imminent in entertainment and games industries is that companies or enterprises involved in the analysis of user behaviours are just not willing to disclose any details about their findings on the correlation between users' particular activities and viewership changes.

Talking with details about this phenomenon prevailing in these industries, G. Ushaw, a lead researcher at the Newcastle University in the United Kingdom, revealed his insight [28] from his unique experiences working in the games industry, which has an amount of commercial interest in scrutinising both application usage (i.e. how often do players visit various sections of a game) and social media to assess popularity. His strong impression was that, while the games industry is pretty assiduous at collecting data, its interpretation of the collected data and especially its reaction to the data was fairly basic. For example, if a part of a map is little used in a multi-player game, the game might add a weapon pick-up there. There is another evidence supporting the expert's view on the unilaterally clandestine attitude of industries appeared

in the article on New York Times in a way that "the unanswered question remains in exactly how much chatter in Twitter does lift TV viewers' engagement of a particular show" [9].

2 MOTIVATIONS

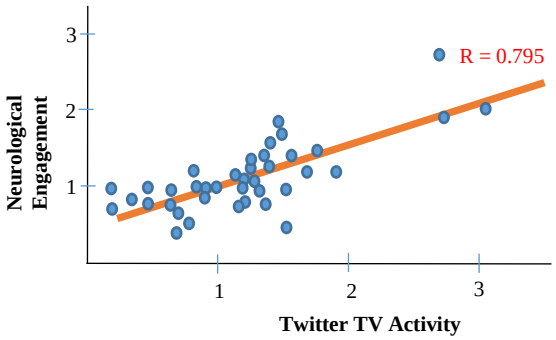
1.2.1 Neurological Finding between Social Media and TV Engagement

Nielsen, a data company best known for its TV ratings, found that the changes in Twitter TV activities are strongly correlated to an extent of 79.5% with neurological engagement shown in Figure 1 [18]. More specifically, the study identified emotion, memory and attention as the three specific neurometrics signals tied to Twitter TV activities. This finding is notable for three reasons.

Firstly, the fact that Twitter TV activities are correlated with the combination of these three neurometrics signals where the program content is engaging viewers through multiple psychological processes. Secondly, the Neuro research of Nielsen has shown that the combination of these same key neurometrics is correlated with sales outcomes in advertisement testing. Thirdly, the Nielsen's TV Brand Effect research has also shown that advertisements perform better on memorability in TV programs with high program engagement. Combined together, these findings suggest that advertising in highly social programs could be an opportunity to drive both advertisement memorability and sales outcomes.

1.2.2 Trends of Multi-Channeling

The American Society of Composers, Authors and Publishers (ASCAP) [23], a music licensing agency, is in one sense fighting for its survival, seeking to change decades-old rules to fit the economics of online music. Furthermore, it is finding ways to distribute more money than ever to its thousands of songwriters. Among the most surprising figures in the ASCAP's report in 2014 [1], ASCAP tracked 500 billion performances of songs, twice as many as it did in the year 2013. Online music is growing extremely fast. According to the ASCAP's report in 2015 [2], the amount of streaming in 2015 through so-called on-demand services like Spotify and YouTube, which let people choose exactly what songs they listen to, increased by 55 percent from the year before.



: Program Segment
R: Correlation

Figure 1: Twitter activities that reveal strong correlations with neurological engagement of TV show viewers (Source [18])

This leads to the observation that there is a very strong connection between the activities of streaming and sales of CD, and this was evidently stated by the recent report from British Phonographic Industry (BPI) [4]. The use of streaming services may be at an all-time high, but CDs and vinyl albums haven not been killed off yet. According to BPI and Entertainment Retailers [4], two thirds of music fans use streaming to find new music and thereafter physically buy their favorites.

According to BPI [20], 13.7bn audio streams are recorded in the UK in 2014. In the meanwhile, the sales of vinyl albums have a sharp increase and reached 1.29 million for the first time over 18 years in 2104 [11]. Furthermore, the drop of CD sales are also becoming slower in the recent years: BPI [20] reports in UK, in comparison with the year before, in 2012 CD sales dropped by 20%, in 2013 sales dropped by 13%, in 2014 sales dropped by 8%, and first half of 2015 - sales dropped by 6%. This suggests that it is "partly down to music fans wanting to "emotionally engage" with an album and build their own physical collection" [20].

1.2.3 Strong Performance Remaining in CDs and DVDs Sales

There is another clue that shows CDs and DVDs still remain important to bring profit even for streaming-centered enterprise such as Netflix (netflix.com). Netflix is a streaming mogul and the top performer in entertainment industry according to S&P 500 index [16], and it owns 65 million steaming users in more than 50 countries in the world [25]. Although the streaming business of Netflix "expects its streaming business to just break even globally through 2016 as it pours billions of dollars into content and an aggressive

expansion", Netflix still churns out millions of dollars in profit each year from DVDs with just 5.3 million DVD subscribers, down from 20 million in 2010 [25].

This indicates that although DVD business is in a slow decline, there is still a huge demand for DVDs. As listeners around the world turn to streaming music, Japan, the second-largest music market after the United States, with slightly less than $3 billion in sales in 2014, has been one important holdout, according to a trade group of the International Federation of the Phonographic Industry [24]. In particular, in Japan by resisting streaming services more than 80 percent of the music sales in the country are still on physical formats like CDs.

1.2.4 Facebook's Controversial Experiment

There is other industry observations on how users' emotion is propagated. Facebook, in an academic paper published in conjunction with two university researchers, reported and admitted that it had manipulated the patterns of user emotion feeds to see how effectively they propagate [3]. For one week in January 2012, Facebook had altered the number of positive and negative posts in the news feeds of some 600,000 randomly selected users to see what effect the changes had on the tone of the posts the recipients then wrote. Facebook found that moods were obviously contagious. The people who saw more positive posts responded by writing more positive posts. Similarly, seeing more negative content prompted the viewers to be more negative in their own posts. This experiment indicates that there could be a strong impact for users in selecting which TV programs or music they are going to be engaged.

1.3 Objectives

At this stage of big data research, there seems for many companies that they will be able to encounter a big business opportunity if they invest appropriate level of efforts for collection and analysis of big data. However, the problem is that the big data is not the data that even big companies are easy to reach. For example, accessing data in social media service sites such as Facebook or Twitter is complicated, since their data is not generally open to public. Even in the very few cases of that social media service sites allow their data to be publicly available, thus data is limited to small portions such that third parties hardly find usefulness associated with the data provided to them. A general trend is that the providers of social media data still try to keep the most data relevant to their own merits with regard to marketing secret. If this trend continues to go on, then the third party companies will never find any opportunity in big data even in a long-term perspective.

What matters most is that the initial investment even for a primitive level of simple trials for data crawling is at least beyond some hundreds of thousands of dollars, and in most cases around more than some millions dollars to get a software tool to enable data crawling. It is not at the level that most of the companies can relentlessly afford to conduct their experiments or even preliminary measurements to see if there may be some opportunities for them to be able to deduce significant and meaningful insights from big data so that they would be immersed in further activities related to observations of big data. Moreover, such software tools available in the market are so heavy-weight in terms of functionalities and so slow in terms of speed. Therefore, the observations from the past six years of time lead us to develop a light-weight and at the same time convenient software tool that works with big data, and make the tool public available for any individuals or any organizations that are interested in experiment with big data.

Therefore, the aim of this study is to develop an open platform and present it to the public until the level of source codes. At first, an attempt to access to the big data of Facebook was made as there is no limit in the length of conversations in Facebook. This could for us to be able to facilitate to analysis the results obtained from the big data stores of Facebook more profoundly as they might contain more useful contents owing to the unlimited length of each conversation, compared to the case of Twitter, which limits each message to only 140 characters. However, after finding out that Facebook amended its policy in a way of being not friendly enough in providing its data to third parties, Twitter is considered as an alternative candidate for this study.

With Twitter it was not easy to find a way to get into its API at the first glance. We tried a number of technically different ways to access to the data that Twitter currently stores, and finally one and only one way of accessing its data was found in this study. The source codes of the software tool for accessing Twitter data are mostly developed with Microsoft Visual Studio, Microsoft Access and MySQL, and a fairly large piece of codes is written in C. This tool will also allow checking to see whether there is any correlation between the popularity of music or movies and the frequency of citations amongst users in Twitter. For this aim, music is considered to be better suited to investigation of big insights on the correlation than movies as the influence of individual stars can be more easily grasped in the area of music than that in the area

of movies. In a movie, there usually are a number of stars casted so it is hard to find out exactly how influential any particular star casted is. Therefore, in this study we have a relative more focus on music than on movies.

2 RELATED WORK

It is well established that networks, agencies and advertisers can better understand how audiences of a TV program interact with each other about the TV program by delving into the endless conversations taking place on Twitter. Now, new research [18] shows that Twitter TV activity can also tell us how engaged the general viewing population is with the TV program that it watches. In fact, that sort of activity now stands as a bellwether for general audience engagement. This means that increases in conversation on Twitter during watching a TV program signal that there is a high engagement with the TV program among the general viewing audience.

This new finding is the result of a study [18] that Nielsen conducted. Nielsen analyzed minute-by-minute Twitter activities, in terms of tweets, around live airings of eight prime time broadcasts and cable TV shows with varied levels of Twitter activities and TV ratings. Nielsen also monitored the brain activities of viewers as they were watching new episodes of those programs. Minute-by-minute Twitter TV activities and brain activities were analysed side-by-side across segments of each program to understand whether increases and decreases in Twitter activities were correlated with viewers' neurological response to the TV program.

In its study [18], Nielsen found that three neurometrics signals (emotion, memory and attention) correlated to Twitter TV activities, and that the changes in Twitter TV activities have a big correlation with neurological engagement at a degree of 79.5%. These findings disclosed a fact that Twitter TV activities are tied to the combination of these neurometrics signals and the TV program is engaging viewers through multiple psychological processes. The Neuro research of Nielsen has also shown that the combination of these neurometrics signals is related to with the outcomes of sales in advertisement testing. Furthermore, the fact of that ads perform better on memorability in TV shows with high program engagement was also showed by Nielson in its TV Brand Effect research [18].

While many industry studies have explored the relationship between social TV activities and audience tuning and the relationship between second screen social behaviours and viewer engagement, the study conducted by Nielsen [18] suggests another way for networks, agencies and advertisers to value TV-related conversation on Twitter: this time, not as a measure of social activities in and of itself, but rather as a bellwether for general audience engagement. This insight can be applied to the TV industry to more holistically understand audiences' response to the content and develop further practical insights in the areas of both social TV and consumer neuroscience. Methodology taken in this study is minute-by-minute analysis.

Analysis of minute-by-minute Twitter TV and neuro data was performed for eight prime time broadcasts and cable TV programs ranging in the levels of Twitter activities and TV ratings. Nielsen Neuro monitored the brain activities of nearly 300 participants [18], aged 21 to 54 years old recruited from San Francisco Bay Area, Chicago, and Atlanta using standard Nielsen Neuro recruitment procedures, which include a representative race and ethnicity sampling consistent with the local demographic composition. Neuro data for each program was collected from an equal number of male and female participants who indicated that they regularly watched the given program.

3 TWITTER CRAWLER: DEVELOPMENT AND EXPERIMENTAL EVALUATION

This section presents the development design of our Twitter crawler, and the experimental evaluation of its prototypical implementation.

3.1 Development of *SMCrawler*

One main purpose of this work is to develop a light-weight software tool, which is easy to use and fast when crawling big data of social media. Therefore, we develop a crawler for Twitter, which contains only very essential components: social media application programming interface (API) module, database storing module, and Twitter engine search modules. The role of API module is to access Twitter through API interfaces and get the data out of Twitter. The data crawled in this way will then be stored in relational databases for the analysis. The data crawler developed in this paper is named *SMCrawler*.

Our *SMCrawler* system is comprised of five main parts: Twitter API Authorisation Manager, Microsoft (MS) Access Database Manager, MySQL Database Manager, Twitter Search Manager and Monitor Display management. The structure of *SMCrawler* is depicted in Figure 2.

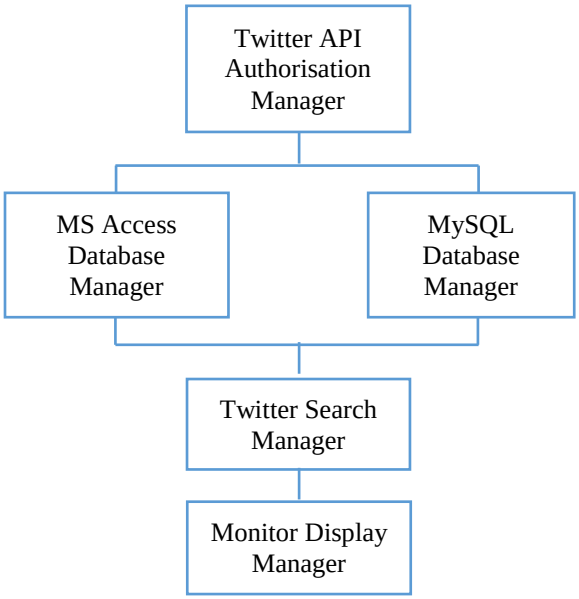


Figure 2: Structure of *SMCrawler*

The Twitter API Authorisation Manager checks whether the details of login to Twitter is valid and the two databases managers (MS Access Database Manager and MySQL Database Manager) are needed to store the data results retrieved from Twitter. The data crawled from the stores of Twitter are stored into two different databases: MS Access database and MySQL database. Twitter Search Manager is the real data crawler that searches and collects data from Twitter. Monitor Display Manager is used to manage the display of data in the monitor.

The data stored in the Microsoft Access database are very similar to Excel documents as the Microsoft Access database provides the primitive form of data. The MySQL database uses traditional relational tables. The purpose of the MySQL database is for the convenience of data manipulations, and it provides and reflects a preference of choices depending upon the interest of data manipulators. Most MySQL workbenches provide a data manipulation platform, which makes the manipulation of databases easy and convenient for users.

The flow of operations of *SMCrawler* is depicted in Figure 3 for the understanding of the details with regard to the processes taken inside *SMCrawler*. To have a clear understanding for the readers of this paper with regard to the way how the search engine developed in this paper works, the key idea on the principle of operations in search engine is presented here. The *SMCrawler* starts to work with a search keyword supplied by a user in a box *Keyword*, and the

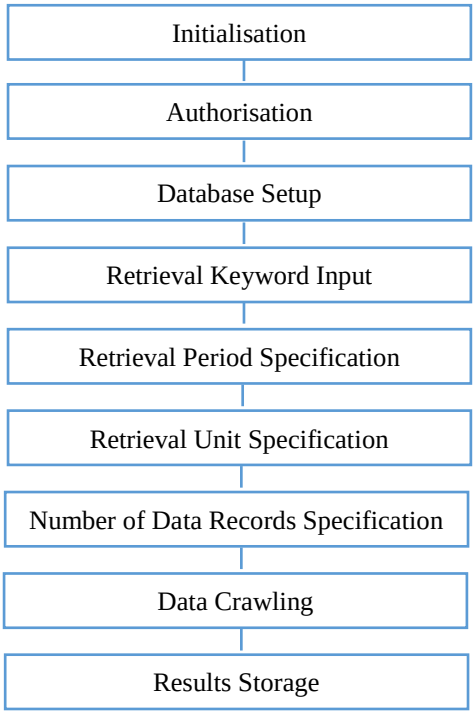


Figure 3: Flow of Operations in *SMCrawler*

input language currently supported is English and Korean. Other input languages can also be easily added. The search engine then crawls and collects Twitter data records from the open source data files of Twitter on basis of the two parameters, Keyword and Language. The collected data are then arranged in a sequence of *User ID*, *User Name*, *Date and Time Created*, *Language*, *Content of Tweet Message of 140 Characters Long*, and *Keyword*.

After the development of the crawler SMCrawler, it was thoroughly and repeatedly tested: we use it to crawl Twitter data almost every data from May to December 2016.

3.2 Measurement of Correlations on Effect of Users' Engagement on Social Media

This section evaluates the prototypical implementation of our *SMCrawler* by the measurement of the correlations on the effect of users' engagement on social media. For the social activities of Twitter users related to the music industry, we focus on the two British pop stars Elton John and Adele, and for the users' activities in Twitter associated with the movie industry, we choose two TV programs in UK: *Blue Bloods* and *Edge of Tomorrow*.

In the arena of music, the UK still uniquely maintains its prominent status in the popularity and quality of music [29]. Especially in the pop music, prominent singers like Beatles, Elton John and Adele are from the UK. The British signer Adele is almost solely phenomenally outstanding even under the circumstances of slowdown of album sales in music tracks.

The true winner throughout the entire past thirty years is clearly Elton John as he has taken the position of the topmost singers in UK, which are traditionally nominated once a year. Elton John was nominated more than ten times in past thirty years and no other singers can match the level of his success in the performance of singers in the world in the whole history of the pop music. He is still doing so superbly and even he is fairly senior enough at this time of his life. Therefore, if we pick one candidate who can neck-to-neck peer to Adele at this time, it would still be Elton John.

With this attention in mind, we want to measure the frequencies of the social activities of users on Twitter associated with British pop stars Elton John and Adele. We use *SMCrawler*, to access Twitter data almost every data from May to December 2016. The crawling results have shown that the density of the user engagement for Adele is consistently at least 20 times higher than Elton John. This indicates the significant impact of social media on the effect of the user engagement in the entertainment industries. A detailed discussion on this research result will be given in the next section.

The Twitter data obtained related to the TV programs (*Blue Bloods* and *Edge of Tomorrow*) showed the impact of social media on the movie industry, which will also be discussed in the next section in detail.

3.3 Problems of Visualization

As Twitter only shows data in a scrolling fashion, taking screenshots was just impossible in its nature from the start, consequently capturing such data inevitably required instantaneous actions of taking snapshots at the experimenter side. Although a high resolution camera was used to capture the result of display, the quality of the snapshots of the scrolling Twitter Data was not satisfactory. Therefore, these snapshots of the data crawled from Twitter will not be displayed here. Instead, we will describe and analyze the results obtained from our experiments in detail in the next section.

4 Analysis of Twitter Data

4.1 Methodology of Analysis of Tweets

The key technique deployed in analyzing the tweets data garnered from Twitter is on basis of keywords. Since the data crawled from the open social media source consists of text only, general perceptions might be that there would be some decent natural language techniques and some formalized mathematical frameworks associated with the techniques in analysing such data. This might be true if the open source were from some different social media platforms such as Facebook other than Twitter for our experiments.

However, since Twitter permits a very limited length of text (only 140 characters), the analysis on these pieces of text simply turned out to be straightforward and intuitive in a way that such complicated methodologies or algorithms are not mandated for the analysis of the Twitter text messages. Therefore, in this paper, we focused on the density of tweets only. If the parameters such as the influence of tweets other than the factor of density might be adopted for this type of research, it is absolutely necessary to incorporate a novel mathematical framework and some natural language processing techniques for the analysis of tweets. A detailed discussion on this issue is given in the part of future considerations in the section of conclusions.

4.2 Analysis of Music Tweets

The results obtained from Twitter's tweets evidently showed that the density of tweets data on basis of time duration absolutely depends upon the popularity of singers.

Tweets data about Elton John is very sparsely distributed in a way that for every minute there are only one or a couple of instances of tweets related to Elton John. In the meanwhile, the results obtained for Adele were fairly much denser in a degree of there are 20 to 30 tweets in every minute, which are associated with Adele. This phenomenon has been constantly and consistently observed from the Twitter data samples of 24 hours on 16th of May 2016.

The results of this experiment show that the number of tweets for Adele turns out to be at least 20 times greater than that for Elton John. This conclusion might be considered as another evidence that proves that Adele is now the top singer, in addition to the fact that Adele has the biggest hit albums in the world at this time and one of her songs particularly holds a phenomenal success of selling 8 million copies in a

year, which previously can never be seen in the past two decades. The measurements for the popularity of the pop musicians Adele and Elton John have been observed with different time and duration in our experiments, and their results appear almost the same patterns throughout the entire trials of measurement in our experiments.

4.3 Analysis of Movie Tweets

To see the impact of social media on the movie industry, two most popular television dramas have been chosen for comparison: *Blue Blunt* and *Edge of Tomorrow*, which are played at exactly the same frame of time on 8pm each Wednesday on Sky, which is one of the prominent television broadcasting stations in U.K., for the period of one week, from the 16th of May, 2016 to the 23th of May, 2016. The results collected from tweets in a duration of two days prior to the televising on Wednesday have displayed an obvious evidence with regard to the existence of the strong correlation between the number of tweets and the number of audience or viewers.

The density of tweets obtained for *Blue Bloods* turned out to be significantly populated than that for *Edge of Tomorrow*. Consequently, it is highly likely that the probability of the audience engagement, during the time of televising, in *Blue Bloods* would subsequently be higher than that in *Edge of Tomorrow*.

4.4 Support from Previous Traditional Media Survey for our Results

A very strong correlation between one's popularity and the density of tweets has been supported by several previous media survey research reports from the media research organisations such as Nielson [19] and Mental Gloss [17] from the UK. According to both Nielson NNTV Program Report [19] and Mental Gloss [17], *Blue Bloods* is ranked as one of the top TV shows in 2015. It was ranked at the ninth and sixth position respectively in those two reports in 2015. In contrast, *Edge of Tomorrow* failed to be included in the top spots even though it is still on air until these days. This indicates that the preference of viewing of audiences for TV shows is predictable before the launching of actual on-airing of TV programs.

It would have been more meaningful if the experiments with more Twitter data for TV viewing were possible as the statements confirmative to our results came out as year-end reports [17] [19]. However, the privilege for accessing Twitter big data is set by Twitter, and only a certain limited portion of data is open to public. For our experiments, the data public available from Twitter is only in the period of one week.

5 Conclusions

It was able to see the impact of social media on the popularity of an artist in the entertainment industry through the activities in the social media. In this study, we show such impact from the social website Twitter. What was found in this study is when users are more engaged with a target in mind, whether the target is a certain genre or a certain entertainer, with social media it is more likely that the users tend to exert their activities of follow-up to that target in a way of e.g. purchasing CDs relevant to the target as an explicit expression associated with their deep loyalties to the target.

In this study, it was a sort of fortunate in that it was able to approach one of the prominent social media, Twitter. However, it must be noted that there seems to be very few social media that are willing to allow their big data for public to be able to freely touch with. For example, Facebook once widely opened their API to the public but very recently it had withdrawn from continuation of that sort of user-friendly support to the public. As many other social media tend to maintain or change their stances, in this study it was only able to experiment with Twitter.

The study on the predictability of the audience engagement on TV shows with social media prior to the actual televising needs to be investigated in more detail in a way how far back in time such predictability remains trustworthy. Such research could be explored with *SMCrawler* if social media open up a vast amount of their data to the public.

5.1 Future Considerations

There is an issue to be investigated further in this sort of research in social media: In case the data crawled from social media are lengthy text, for instance more than 140 characters, the analysis on text data may not be straightforward as in the case of Twitter. Subsequently, some novel methodologies or techniques on basis of natural language processing are needed for analysis of such lengthy text. We will need such techniques if other social media also open their data to the public.

Social big data is closely related with cloud computing. Therefore, future research will include how applying or adapting the techniques developed in could computing to the computing of social big data for e.g. the efficient data management [15][26], for the

distributed storage and processing of social data [12] and the reorganization of spam text [30]. Further research also involves the software engineering perspectives of adopting cloud computing [6] in enhancing the software design and implementation capability of the tool as the social media big data is usually provided on the cloud computing platform.

Other important issues are the security in big data [7] and privacy in big data [27], since even under the presumption of anonymity in revealing big data, the vulnerability of inference with regard to the data revealed may still be widely open. Recovery issues including disaster recovery [8] must also need to be addressed, since the data crawled must be guaranteed to be safe even under the circumstances of mobility and portability in positioning the data. Another issue worth investigating is the data quality in big data [14] as the data obesity due to unnecessary redundancy and subsequent inconsistency prevalent in these days becomes very serious.

We are going to definitely investigate how to improve the clarity of visualisation for data crawled from social media in the coming occasion of research as a sequel to this research paper.

REFERENCES

[1] ASCAP, “Annual Report 2014”, [Online]: http://www.ascap.com/~/media/files/pdf/about/annual-reports/ascap_annual_report_2014.pdf, 2014.

[2] ASCAP, “Annual Report 2015”, [Online]: http://www.ascap.com/~/media/files/pdf/about/annual-reports/2015-annual-report.pdf, 2015.

[3] R. Booth, “Facebook Reveals News Feed Experiment to Control Emotions”, [Online]: m https://www.theguardian.com/technology/2014/jun/29/facebook-users-emotions-news-feeds, Guardian, June 2014.

[4] British Phonographic Industry, “Streaming Drives Boost in UK Music Consumption in First Half of 2015”, White Paper, retrieved from https://www.bpi.co.uk/home/streaming-drives-boost-in-uk-music-consumption-in-first-half-of-2015.aspx, April 2016.

[5] S. Carey, “How World of Warcraft Maker Blizzard Entertainment Uses BI and Analytics to Unlock Business Value of Gameplay Data”, Computerworld, UK, March 2016.

[6] V. Chang, “A Cybernetics Social Cloud”, Journal of Systems and Software, 30 Pages, May 2016.

[7] V. Chang, Y. Kuo and M. Ramachandran, “Cloud Computing Adoption Framework: A Security Framework for Business Clouds”, Future Generation Computer Systems, Vol. 57, pp. 24-41, 2016.

[8] V. Chang, “Towards a Big Data System Disaster Recovery in a Private Cloud”, Ad Hoc Networks, Vol. 35, pp. 65-82, 2015.

[9] V. Goel, “Study Backs Twitter as Barometer of Public Moods”, The New York Times, p. 17, March 2015.

[10] V. Goel, "Facebook Tinkers with Users’ Emotions in News Feed Experiment, Stirring Outcry", The New York Times, [online]: http://www.nytimes.com/2014/06/30/technology/facebook-tinkers-with-users-emotions-in-news-feed-experiment-stirring-outcry.html, June 29, 2014.

[11] S. J. Griffiths, “UK's first official vinyl chart launched as sales rise”, BBC News, [Online]: http://www.bbc.com/news/entertainment-arts-32251994, April 13, 2015.

[12] S. Groppe, J. Blume, D. Heinrich, and S. Werner: "A Self-Optimizing Cloud Computing System for Distributed Storage and Processing of Semantic Web Data", Open Journal of Cloud Computing (OJCC), 1(2), Pages 1-14, 2014, [Online]: https://www.ronpub.com/publications/ojcc/OJCC-v1i2n01_Groppe.html

[13] B. Harvey and B. Friend, “Building Audience: Courting and Keeping Customer in a Media and Entertainment Industry Awash in Data”, White Paper, Vision Critical, 2014.

[14] S. Haustein, “Grand Challenges in Altmetrics: Heterogeneity, Data Quality and Dependencies”, Scientometrics, Vol. 108, No. 1, pp. 413-423, 2016.

[15] U. Marjit, K. Sharma, P. Mandal: "Data Transfers in Hadoop: A Comparative Study", *Open Journal of Big Data (OJBD)*, 1(2), Pages 34-46, 2015, [Online]: https://www.ronpub.com/publications/ojbd/OJBD_2015v1i2n04_UjjalMarjit.html

[16] J. McDuling, "Netflix is the best performing stock in the S&P 500 this year", Quartz, [Online]: http://qz.com/395298/netflix-is-the-best-performing-stock-in-the-sp-500-this-year/, April 30, 2015.

[17] Mental Gloss, “The Most Watched TV Shows of 2015: That Aren't Football”, [Online]:

http://mentalfloss.com/article/73122/most-watched-tv-shows-2015-arent-football, December 2015.

[18] Nielson, “Social TV: A Bellwether for TV Audience Engagement”, White Paper, [Online]: http://www.nielsen.com/us/en/insights/news/2015/social-tv-a-bellwether-for-tv-audience-engagement.html, March 2015.

[19] Nielson, “Tops of 2015: TV and Social Media”, [Online]: http://www.nielsen.com/us/en/insights/news/2015/tops-of-2015-tv-and-social-media.html, December 2015.

[20] Newsbeat, "Streaming hasn't killed off CDs and vinyl yet, says BPI", BBC, [Online]: http://www.bbc.co.uk/newsbeat/article/35058533/streaming-hasnt-killed-off-cds-and-vinyl-yet-says-bpi, Dec. 11, 2015.

[21] D. Palmer, “How Data Scientists Have Changed the Way Games Are Designed at Candy Crush Developer King”, Computing Magazine of The UK, March 2015.

[22] A. Rajpurohit, “Interview: Vince Darley, King.com on the Serious Analytics behind Casual Gaming”, Kdnuggets Magazine, March 2015.

[23] B. Sisario, “Ascap Topped $1 Billion in Revenue Last Year, Lifted by Streaming”, The New York Times, [Online]: http://www.nytimes.com/2015/03/04/business/media/ascap-topped-1-billion-in-revenue-last-year.html, March 2015.

[24] B. Sisario, “Music Streaming Service Aims at Japan, Where CD Is Still King”, [Online]: http://www.nytimes.com/2015/06/12/business/media/line-music-a-new-streaming-service-aims-at-japanese-market.html, June 2015.

[25] E. Steel, “Netflix Refines Its DVD Business, even as Streaming Unit Booms”, retrieved from [Online]: http://www.nytimes.com/2015/07/27/business/while-its-streaming-service-booms-netflix-streamlines-old-business.html?_r=0, July 2015.

[26] P. Teli, M. V. Thomas, and K. Chandrasekaran: "An Efficient Approach for Cost Optimization of the Movement of Big Data", *Open Journal of Big Data (OJBD)*, 1(1), Pages 4-15, 2015, [Online]: https://www.ronpub.com/publications/ojbd/OJBD_2015v1i1n02_Teli.html

[27] O. Tene and J. Polonetsky, “Big Data for All: Privacy and User Control in the Ages of Analytics”, Northwestern Journal of Technology and Intellectual Property, Vol. 11, No. 5, pp. 239-273, 2013.

[28] G. Ushaw and S. Moon, “Intimate Interview on Games Industries in the UK”, Colloquial Seminar, Claremont Tower, School of Computing Sciences, Newcastle University, UK, April 2016.

[29] Wikipedia, “Music of the United Kingdom”, [Online]: https://en.wikipedia.org/wiki/Music_of_the_United_Kingdom., June 2016.

[30] M. YR, S. S. Nayak, K. Chandrasekaran: "Cognitive Spam Recognition Using Hadoop and Multicast-Update", *Open Journal of Big Data (OJBD)*, 1(1), Pages 16-28, 2015, [Online]: https://www.ronpub.com/publications/ojbd/OJBD_2015v1i1n03_YR.html

ACKNOWLEDGEMENT

This research was partially supported by SBS, Seoul Broadcasting System, a major media company, in Korea.

AUTHOR BIOGRAPHY

Dr. Songchun Moon is currently a Professor at College of Business at KAIST in Korea and a Professor at School of Computing Sciences at Newcastle University in the U.K. He finished his PhD in Computer Science at University of Illinois at Urbana-Champaign in January 1985. Since then he has been teaching in Department of Computer Science at KAIST, Cambridge University, Edinburgh University and Queen's University of Belfast and also in School of Business and Economics at National University of Ireland, Galway. He has developed first multi-user relational database management system called IM in Asia in 1990 and first-ever distributed database management system DIME in Asia in 1992. He published more than 70 international journal papers and was the chair of the program committee at DASFAA'93 and a program committee member in ACM SIGMOD, IEEE Data Engineering and VLDB. He was invited as the Distinguished Scholar at the Hungarian Academy of Sciences. He has been PhD Program Director at KAIST and has been and is still serving as a director of European Association of Information Technology for the past 25 years. His research areas include enterprise data management, database engines, cyber security and privacy, analysis of social media big data and big data management.