

Complex Adaptive Systems Conference with Theme: Cyber Physical Systems and Deep Learning, CAS 2018,  
5 November – 7 November 2018, Chicago, Illinois, USA

## Text Mining to Understand the Influence of Social Media Applications on Smartphone Supply Chain

Aditya Akundi<sup>a\*</sup>, Bill Tseng<sup>a</sup>, Jiamin Wu<sup>b</sup>, Eric Smith<sup>a</sup>, Subbalakshmi M, Francisco Aguirre<sup>a</sup>

<sup>a</sup>IMSE Department, University of Texas at El Paso, El Paso, Texas -79902, U.S.A

<sup>b</sup>Shantou University, P.R. China

---

### Abstract

Social media applications are being currently utilized to influence the information flow aspect of supply chain management. Verifiable real-time data can provide timely and insightful information about several key aspects of the supply chain of a product and enable it to adapt quickly to ever-changing market conditions. In this paper, information is gathered from Twitter to understand how the tweets about a given smartphone can influence its supply chain and its management. Based on relevant hashtags and keywords found in the latest news about three different smartphone brands (i.e. Apple, Samsung, Huawei), data mining is used to extract and analyze the tweets with the specific hashtags or keywords from Twitter. To reduce the loss of a significant amount of event related information due to Twitter's APIs data access restrictions, the concept of refined hashtags and Keywords is also used to enhance the Twitter crawling model used. Sentiment analysis and opinion analysis were carried out based on the refined hashtags with a goal of analyzing people's emotion towards a specific smartphone brand and on the possibility to predict its influence on aspects of supply chain for enabling real-time adjustments for ensure a robust supply chain model. This effort enabled identifying a new model of smartphone supply chain management with built in social media information flow.

© 2018 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Selection and peer-review under responsibility of the Complex Adaptive Systems Conference with Theme: Engineering Cyber Physical Systems.

**Keywords:** Supply Chain; Data Mining; Sentiment Analysis; Refined Hashtags

---

\* Corresponding author.

E-mail address: [sakundi@utep.edu](mailto:sakundi@utep.edu)

## 1. Introduction

### 1.1. Social Media and Supply Chain

The term social media implies to the use web based and mobile technologies to communicate and interact based on user generated content and dialogue. Enabled by several accessible platforms and characteristics, social media significantly changed the way individuals, organizations and communities communicate [1]. These include: several information content formats such as text, images, video; providing interaction platforms using information sharing and emails; engagement of people across levels (i.e. people who generate information, people who share, and who comment). With the increased popularity of social media, new digital platforms are being developed and used across where individuals interact and share information in the form of news and opinions [2]. Few examples include: Individuals posting reviews on a given products functionality and their experience with a product on Amazon reviews, expressing views and sharing images and videos on Facebook, and Twitter to get updates on current technology and social trends and to keep updated with news across the globe. More emphasis is currently observed in understanding the role of social media on society and its influence in driving a change in institutions, organizations, and societies [2]. In this paper, we explore the use of social media applications in influencing the information flow aspect of supply chain management. With continuous real time data being generated by social media users, this paper analyses on how this data can provide timely and insightful information about several key aspects of the supply chain of a product and enable it to adapt quickly to ever-changing market conditions.

By definition, supply chains are very large in nature incorporating several vendors, distribution centers, suppliers, consumers, manufacturers, manufacturing plants, logistics providers and so on. The use of social media in supply chains can help in identifying trends to make better design and business decisions [3] to include the knowledge gained on industry competition, building customer relations, identifying key performance indicators based on customer feedbacks, and solicit customer recommendations on product functionality for improvement. Few examples of the use of social media in supply chain include [3]: using social media posts for identifying arrival and departure of shipments to customers, knowledge discovery from industry influences and leaders, monitoring reputation among the customers, enabling continuous improvement by finding new suppliers, vendor and consumer base.

This paper provides a summary of an effort by the authors from the information gathered in Twitter to understand how the tweets about a given smartphone can influence its supply chain and its management, based on relevant hashtags and keywords found in the latest news about three different smartphone brands (i.e. Apple, Samsung, Huawei).

### 1.2. Twitter as an Information Platform

Twitter is a social media tool that allows its users to publish information in the form of short messages up to 140 characters or less. One can also use Twitpic and other similar resources to send and share video and pictures in Twitter [2]. Messages in Twitter can help provide information about a wide range of supply chain events such as, Twitter messages for notifying the arrival and departure of a shipment from a warehouse, coordinating shipment handling and transportation events [1]. Compared to other social media platforms, Twitter provides free access to its users to capture data and with around 328 million users of Twitter reported in the first quarter of 2017, we use Twitter posts to capture and fetch tweets based on keywords referring to Apple, Samsung, and Huawei smartphone brands through hashtags. These three brands were selected for analysis based on their wide acceptability and use.

## 2. Tweet Mining Model

Text mining, an application of Natural language processing, is the application of analytical methods to text data to derive information from a given collection of text. Using algorithms, one can learn from huge amount of texts by summarizing the main theme and to identify the ones of most interest. Applications of text mining include clustering, text summarization, and link analysis [5]. Text mining methodologies have been successfully utilized to analyze large

amount of texts in the fields of business, health sciences, manufacturing and educational domains [5]. To apply text mining for analyzing tweets (unstructured text data) in this context of study fetched from Twitter, Fig. 1. illustrates the tweet mining methodology developed by the authors.

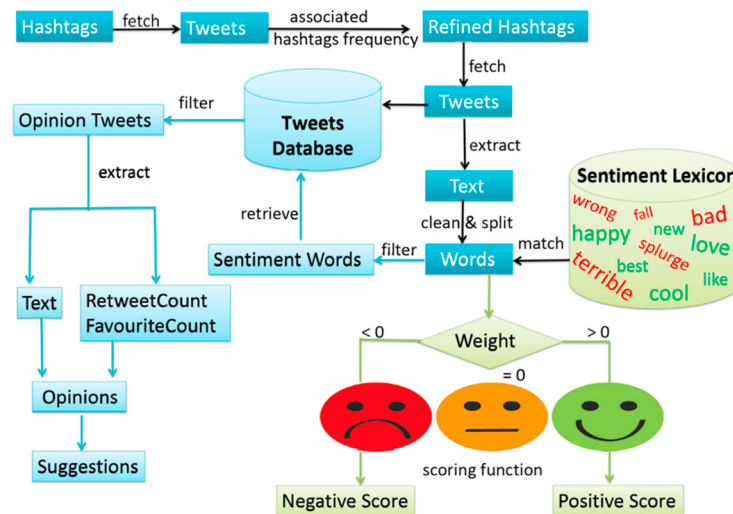


Fig. 1. Tweet Mining Framework used

The remainder of this paper explains the steps followed using the illustrated framework.

### 2.1. Hashtag Selection

The interactions on Twitter, a microblogging platform are mostly unstructured with the limitation of 140 characters per tweet. Hashtags, a crowdsourcing tagging system, facilitate as information markers to aggregate information from a same context.

Users in Twitter use hashtags to post information about and to join an already ongoing discussion [4]. To identify the hashtags relevant to the smartphones in question, hashtags “#iPhone” for iPhone, “#Samsung” for Samsung phones, and “#Huawei” for Huawei phones, were first used to fetch the tweets. Based on the fetched tweets, the hashtags were further refined to capture more tweets. The goal of refining hashtags is to collect additional relevant tweets that are associated to the phone brand in question without the use of relevant hashtag. For example, the tweet “Tim cook blames rumor mill for the bad sales lately” does not have #iPhone but it still indicates the declining sales of iPhone. Another goal for refining hashtags is to reduce the loss of a significant amount of information because of Twitter’s APIs data access restrictions which complicates fetching relevant tweet corresponding to an event. For example, the tweet “I love the high #tect thing as #iPhone so much that I will definitely buy the next one” had #iPhone but was not captured at the first attempt due to the rate limit in Twitter. However, this tweet could have been captured using #tect, which doesn’t have a direct relation to iPhone. Fig. 2. illustrates the refined hashtag selection approach.

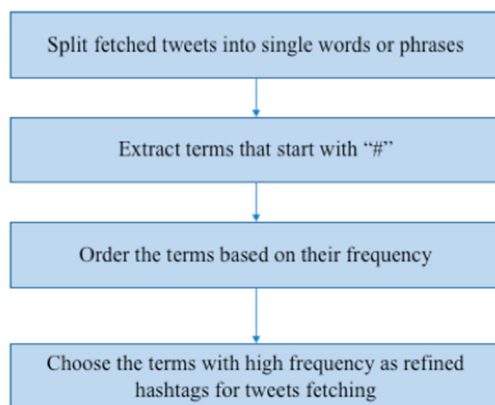


Fig. 2. Refined Hashtag selection approach

Table 1 portrays the original hashtags and refined hashtags identified for the analysis.

Original Hashtags	Refined Hashtags	Frequency	Relationship
#iPhone	#Apple	2849	Company
	#IOS	1819	Operating System
	#iPhone7Plus	105	Phone
	#iPhone8	5016	Phone
#Samsung	#GalaxyS8	827	Phone
	#GalaxyNote8	246	Phone
	#Android	614	Operating System
#Huawei	#Huaweip10	114	Phone
	#Huaweip10Plus	45	Phone
	#Android	327	Operating System

Fig. 3. shows the relation between original and refined hashtags. The size of the nodes in fig. 3. represent the word frequency where bigger the node size, higher the frequency. The length of the links between the nodes represent the connection wherein the longer the length, weaker the connection between the nodes.

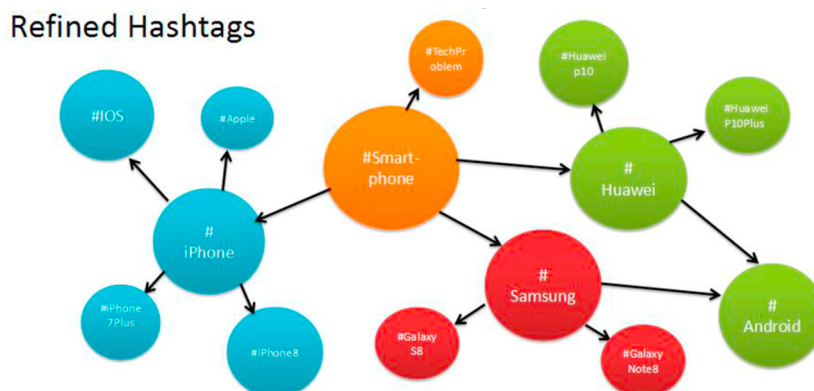


Fig. 3. Refined hashtags used for tweets fetching

## 2.2. Tweets Fetching and Cleaning

Considering the advantages and disadvantages of Twitter API, both search API and streaming API were used to capture the tweets, which are the stored in a database to be analyzed. For cleaning the fetched tweets, text is first extracted to identify and remove irrelevant notions in text such as Punctuations, Stop words, White space, URLs, Expressions, and Numbers. Tweets once cleaned are then split into words for analysis. Table 2 portrays the refined hashtags and the number of tweets fetched based on the refined hashtags.

Table 2. Fetched tweets based on refined hashtags

Hashtags	Tweet Number	Time	Language
#iPhone	15150	2017/7/11	English
#IOS	10000	2017/7/13	English
#Samsung	10000	2017/7/14	English
#Huawei	6254	2017/7/15, 2017/7/16	English
#Android	10000	2017/7/17	English

## 2.3. Sentiment Analysis

Sentiment analysis is the process of identifying the viewpoint and the tone lying behind a corpus of words for understanding the attitude expressed by an individual online. Application of sentiment analysis include its use by Obama administration to understand public opinion on policy changes and announcements ahead of presidential election [6]. Similarly, sentiment analysis was used to understand consumer attitude and react accordingly by Expedia Canada when they observed negative feedback on the music used in their television advertainment [6]. The words of each hashtag were matched to sentiment lexicon. Sentiment lexicon, a resource for sentiment analysis is a database of lexical units for a language with sentiment orientations.

Words from texts were matched with SENTIWORD 3.0.0. [7], a sentiment lexicon with positive or negative score for each sentiment word. For # iPhone 650 sentiment words were retrieved. Table 3 and Table 4 show only the sentiment words with a frequency higher than 136 words. For #Samsung, 1188 sentiment words were retrieved and illustrated in Table 3 are the sentiment words with a frequency higher than 150. 496 sentiment words were retrieved for #Huawei and Table 4 shows sentiment words with a frequency higher than 200.

Table 3. Phone Hashtag Sentiment Words Retrieval

Word Match	Frequency	Positive-Score	Negative-Score	Word Match	Frequency	Positive-Score	Negative-Score
<b>#iPhone</b>				<b>#Samsung</b>			
New	517	0	0.125	Sale	517	263	0.125
Will	516	0.125	0	Will	516	260	0.125
Touch	383	0.375	0.125	Get	383	244	0.375
Look	298	0	0.375	Just	298	241	0.625
Fingerprint	274	0	0.625	Profit	274	227	0.875
Face	220	0	0.5	News	220	218	0.125
Prototype	203	0.125	0.375	Win	203	211	0.125
Recognition	197	0.125	0	Smart	197	208	0.5
Like	161	0.5	0	Gold	161	206	0.125
Price	159	0.625	0	Best	159	199	0.75
Release	139	0.125	0	Poorly	139	171	0

Table 4. Phone Hashtag Sentiment Words Retrieval

Word Match	Frequency	Positive-Score	Negative-Score
<b>#Huawei</b>			
Mate	305	0.5	0
Smart	265	0.5	0
Chicken	255	0	0.125
Yet	245	0	0.125
Will	239	0.125	0
Release	225	0.125	0
New	214	0	0.125

For hashtag #IOS 918 sentiment words were retrieved. Fig. 4 shows the sentiment words with a frequency higher than 200. 650 sentiment words were retrieved for #Android and in Fig. 4 the words with frequency higher than 200 are shown.

<b>#IOS</b>				<b>#Android</b>			
wordmatch	Frequency	Pos-Score	Neg-Score	wordmatch	Frequency	Pos-Score	Neg-Score
like	3602	0.375	0	like	1519	0.375	0
goal	3522	0.125	0	goal	1435	0.125	0
free	1293	0	0.625	amp	595	0.125	0
new	761	0	0.125	news	583	0.125	0
amp	601	0.125	0	free	500	0	0.625
get	376	0.375	0	new	472	0	0.125
featured	281	0	0.125	just	284	0.125	0
puzzle	237	0	0.125	will	247	0.125	0

Fig. 4. System Hashtag Sentiment Words Retrieval

Sentiment scoring is used for each hashtag for identifying their polarity i.e. identify a positive and a negative score for each hashtag. Sentiment scores identified for each hashtag are shown in Fig. 5.

Hashtags	Positive Score	Negative Score	Total Score
#iPhone	0.4479	0.2967	0.1512
#Samsung	0.5108	0.2550	0.2558
#Huawei	0.4302	0.1788	0.2514
#IOS	0.6755	0.2192	0.4563
#Android	0.4847	0.1869	0.2978

Fig. 5. Sentiment score of Hashtags

The polarity sentiment score, see Fig. 6., when mapped illustrates that #Samsung has the highest positive score while #Huawei and # iPhone have the lowest; #iPhone has the highest negative score, while #Huawei has the lowest; and #IOS has a higher positive score to # Android.

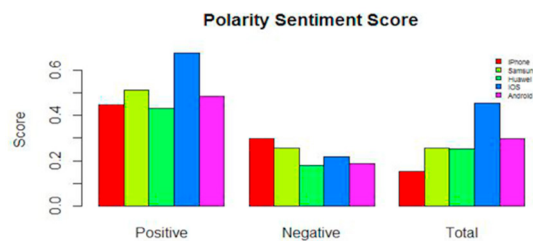


Fig. 6. Comparison of Polarity Sentiment Score

Based on the results from sentiment words retrieval, the top 10 words with higher frequency were selected for opinion analysis. Opinion tweets i.e. the retrieved tweets containing the top sentiment words were used to understand consumer opinions on smart phones which could be used for making useful suggestions for better decisions in supply chain. Fig. 7 shows the opinion tweets.

Text	RetweetCount	FavouriteCount
RT @JonylveParody: "Steve Jobs would never have released a product with a color shiny, mirror-like finish." <b>#iPhone</b> #iPod <a href="https://t.co/wPg7mhdKZk">https://t.co/wPg7mhdKZk</a>	56	342
<b>#iPhone 8</b> "Touch ID on the back" version. Not so <b>bad</b> as I thought at first.. :D <a href="https://t.co/6RwpY1Q8eE">https://t.co/6RwpY1Q8eE</a>	121	311
RT @WellsAdams: Remind me when the <b>#iphone 8</b> comes out to <b>splurge</b> and get the most memory humanly available. This "storage almost full" nigh...	81	915
The <b>new</b> @Apple iphone will cost \$1200+? If so, my 7+ will be my last <b>#iPhone</b> . Can't justify spending. Wish there was another viable option./New #iPhone available fall2017 will cost \$1200;some young people now question if they can afford it to keep the newest iPhone./Only a company like #Apple can have half of the world saving money for the upcoming #iPhone cause it's going to be damn expensive!	526	127
Starting to really <b>like</b> the glass back design in white. So premium <b>#iPhone8</b>	61	89
Text	RetweetCount	FavouriteCount
<b>#Samsung</b> Will Get More <b>Profits</b> Than Apple For The First Time In Its History <a href="https://t.co/4LK1f7DeUZ">https://t.co/4LK1f7DeUZ</a> #news/#SAMSUNG On Track To Replace Apple As Most Profitable Tech Company, Thanks To Strong <b>Sales</b> Of The S8	5401	409
<b>#huawei</b> and #kfc made a <b>smart</b> phone – yes, really <a href="https://t.co/ytUdDgVkrx">https://t.co/ytUdDgVkrx</a> <a href="https://t.co/Xq8BGAFbAr">https://t.co/Xq8BGAFbAr</a> /KFC X HUAWEI the collaboration that you never imagine happen!!! #lifestyle #huawei #kfc #food #trend #chicken... <a href="https://t.co/g6VcbcaAzQ">https://t.co/g6VcbcaAzQ</a>	16152	774

Fig. 7. Opinion Tweets

It is observed that: the sentiment towards the glass black design of iPhone8 varies among the consumers where in, few think it is an expectational design where as others had a negative opinion on it. Though both the tweets recount is similar, but the favorite count of the second tweet was much higher indicating a negative sentiment towards the design of iPhone 8. The sentiment on iPhone 8 touch ID was positive and the sentiment on its memory storage was negative. Also, there was a positive sentiment observed in the tweets towards the venture of KFC Huawei smartphone.

### 3. Discussion and Future Work

All in all, the analysis of the tweets fetched portrayed that the sentiment on Samsung and Huawei were positive when compared to iPhone. In the case of Samsung, most of the tweets were about Samsung making more profits comparative to apple based on their new release of galaxy 8. Huawei was mostly in the tweets that featured the collaboration with KFC in coming up with anew smartphone. Finally, for iPhone most of the tweets were on a negative sentiment about the price and design of iPhone 8. However, the sentiment towards IOS system was observed to be better in comparison to Android system. These insights can help manufacturers making better decisions, for instance, apple considering lowering the price of iPhone 8, or in modifying glass black design, and increasing memory capacity to better accommodate their customers. For phone component suppliers, considering the sentiment towards Samsung and Huawei are better than to that of iPhone, suppliers can produce and meet the demand of Samsung and Huawei components based on the changing trends from consumer sentiments.

This work helped to identify a conceptual model of smart phone supply chain management with built in social media analysis, where information flow is faster and diversified compared to the traditional supply chain. Fig. 8. illustrates the conceptual framework identified.



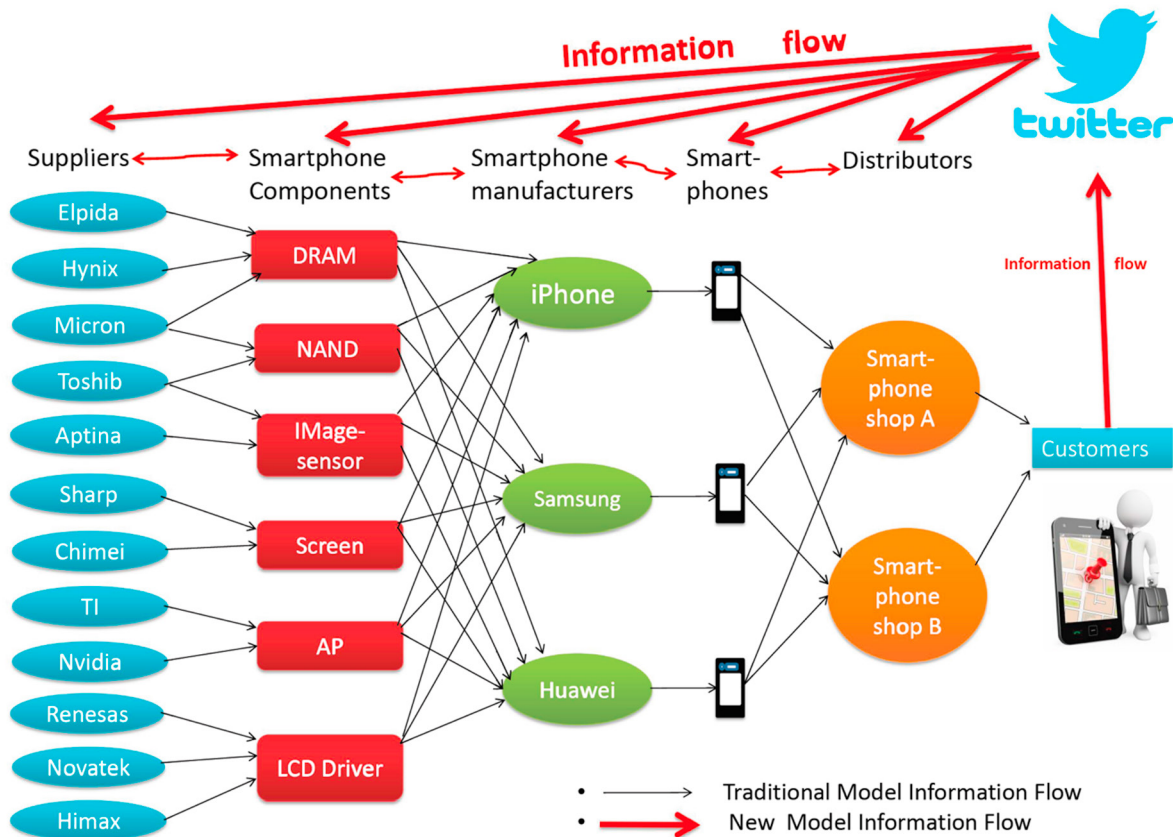


Fig. 8. Smartphone supply chain management model with built in social media information flow

The framework identified is based on hashtag refinement that applies only to hashtags, however the authors note that non-hashtag words also contain important potential information and will be used as a future work to improve the framework. Though it is difficult to verify the integrity of information fetched from Twitter on whether a specific tweet is true or false, the identified framework can be used as a baseline in developing new models and algorithms to facilitate better supply chain decisions by verifying tweets based on a user's credibility, fetching multiple sets of tweets for validating results, and in applying machine learning techniques to improve the accuracy and reliability of tweet mining models.

## References

- [1] O'leary, Daniel E. "The use of social media in the supply chain: survey and extensions." *Intelligent Systems in Accounting, Finance and Management* 18.2-3 (2011): 121-144.
- [2] Housley, William, et al. "Interaction and transformation on social media: the case of Twitter campaigns." *Social Media+ Society* 4.1 (2018): 2056305117750721.
- [3] Rusch, Ed. "Using Social Media In The Supply Chain." *Manufacturing Business Technology*, 18 Dec. 2015, [www.mbtmag.com/article/2014/08/using-social-media-supply-chain](http://www.mbtmag.com/article/2014/08/using-social-media-supply-chain).
- [4] Giglietto, Fabio, and Yenn Lee. "A hashtag worth a thousand words: Discursive strategies around# JeNeSuisPasCharlie after the 2015 Charlie Hebdo shooting." *Social Media+ Society* 3.1 (2017): 2056305116686992.
- [5] He, Wu, Shenghua Zha, and Ling Li. "Social media competitive analysis and text mining: A case study in the pizza industry." *International Journal of Information Management* 33.3 (2013): 464-472.
- [6] Bannister, Kristian. "Sentiment Analysis: How Does It Work? Why Should We Use It?" *Brandwatch*, Brandwatch, [www.brandwatch.com/blog/understanding-sentiment-analysis/](http://www.brandwatch.com/blog/understanding-sentiment-analysis/).
- [7] SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining <url: <http://sentiwordnet.isti.cnr.it/>>