Mining Time-Dependent Influential Users in Facebook Fans Group

Li-Jen Kao[1]
[1]Department of Computer Science and Information Engineering
Hwa Hsia University of Technology
Taipei, 23568 Taiwan
Email: lijenkao@cc.hwh.edu.tw

Yo-Ping Huang[2,3]
[2]Department of Electrical Engineering
National Taipei University of Technology, Taipei, 10608 Taiwan
[3]Department of Computer Science and Information Engineering
National Taipei University, New Taipei City, Taiwan 23741
Email: yphuang@ntut.edu.tw

Frode Eika Sandnes[4,5]
Faculty of Technology, Art and Design
[4]Oslo and Akershus University College of Applied Sciences, Oslo, Norway
[5]Westerdals Oslo School of Art, Communication and Technology, Oslo, Norway
Email: Frode-Eika.Sandnes@hioa.no

***Abstract*—Klout, a famous App, could measure people's social network influence power. Klout score is measured according to the data from past 90 days and an individual who has high Klout score is thought as having high social influence power. Lots of businesses or organizations like to hire high Klout score people to help them to diffuse their brand images. However, Klout score cannot tell us who has high influence power in a specific short time period. For example, it is possible that some of the users might always have high influence power on Monday or on Monday morning. These time-dependent influential users probably have low Klout scores in average but have high influence power in some specific time periods. Businesses should not just know who are the high Klout score users but also they should identify who are the time-dependent influential users because all of them may have some sort of power to influence other users' buying decisions. In this study, a framework based on frequent pattern mining is proposed to find the time-dependent influential users. First of all, the framework will divide a predefined long time period into successive short time segments and then influential transactions that contain Facebook fans' influence power data will be defined in each time segment. From the frequent patterns, the proper time for time-dependent influence users to spread information can be found. A theoretical experiment is given to verify the effectiveness of the proposed framework.**

Keywords—social network mining; frequent itemset; social network influential users mining

I. Introduction

Social network websites are the platforms that users can interact with each other by sharing or exchanging opinions [9]. Not just sharing opinions, the trust relationships can be built among users in this computer network community. Typically, the user builds the relation with someone who they never encountered in person and regardless of geographical locations [4]. It is similar to that people trust the reports from TV news or newspapers, trusting the posters' information on social networks makes the social networks a new type of medium. Users even will spread the ideas on the posters to everywhere of the world. Today, social networks are new information propagation model which the users are both information consumers and information producers [1, 14].

In order to utilize the information propagation power from social networks, lots of businesses or organizations start to build their own fans group to post advertisement to prompt their products. For example, in Facebook fans group, as businesses post new products' information, some of the fans will post comments to share user experiences, and some of the fans add feedbacks for aforementioned comments. If some of the comments, opinions or debates are trusted by fans, not just their buying decisions may be affected but the comments, opinions or debates will be spread by them to everywhere.

The posted comments can affect social network users' thinking and then the comments will be spread out by the users whose thinking is affected. Such a recursive influence propagation way makes the social networks not just social platforms but also marketing powerhouses [15]. For example, the Facebook fans group of Carrefour Taiwan owns 500 thousand fans. Because of fans' influence propagation; many people who are social network users but not the group fans attend the activities held in physical mart.

In fact, influence propagation problems are interesting research topics and have been widely studied in the past [2-3,5-6,9,16,19]. Typically, businesses pay attention on the problem of how to identify the right influential users in social network [1]. If business could know who the influential users are, then businesses can employ every possible way to make the influential users to have more understanding about products and ask them to be market-movers to affect other fans buying decisions making. An important marketing strategy for businesses in social networks is to turn the influence fans as unofficial spokesmen and triggering a cascade of influence by which fans affect other fans or non-fans.

There are also many studies of how to identify influential users. Agarwal et al. [1] proposed a method to identify influential bloggers is according to a set of collectable statistics. Goyal et al. [7] employ propagation graph mining method to find influential users. But now, there is a more easy way to find influential users. Klout, established in 2009, is the application that allows business companies to measure individuals influence scores. Klout score is measured according to the data from past 90 days and an individual who has high Klout score is thought as having high social influence power. Today, businesses try to find influential users according to their Klout scores.

However, there is no method including Klout could tell us who has high influence power in a specific time period. Because even the literature from Klout says that the time for opinions to post is important [18], it is possible that a fans group user may have high influence power in a specific time period but owns low Klout score. Such time-dependent users may also affect others buying decisions and business should not ignore them.

In this paper, we propose a framework based on frequent itemsets mining method to find potential time-dependent influential users. First of all, the framework will divide a predefined long time period into successive short time segments. An influence transaction that contains the fans' influence power data will be defined in each time segment. After influence transactions being collected, the frequent patterns will be found to deduce the time-dependent influential users.

The rest of this paper is organized as follows. Section II introduces the topic of social network mining. The proposed time-dependent influential users mining method is given in Section III. In order to show the effectiveness of the proposed method, experimental results are presented in section IV and Section V concludes the paper.

II. Related Works

This section introduces previous researches related to social networks data mining.

A. Social Networks

Social network users perform various activities. One major activity is to share their opinions or post their products ratings on social network platform. The first example is that a user may rate a movie in Yahoo! Movies. Another good example is Facebook fans post user experiences.

Though posting messages is one of the important activities in social network, the important thing is that those messages can be exchanged or spread very quickly in social network and this characteristic makes the social networking sites as new type of social media [9]. Facebook may be the most famous social networking site, and because its marketing power increase a lot in these years, businesses have utilized Facebook to build fans group to initial activities to promote products, to communicate with their fans, and to collect fans data to analyze users' behavior.

B. Social Networks Data Mining

Since the social media data are vast, noisy, unstructured, and dynamic, analyzing social networks data is a challenge work [9]. For example, the influence and homophily problem in social networks has been studied for many years. However, since most social networks have a mixture of both influence and homophily, it is a challenge to distinguish them.

In order to resolve the aforementioned problem aroused by the characteristics of social network data, several literatures have invented new data mining methods to analyze social network data. Those new data mining methods have applied on community analysis, users' sentiments sensing, improvement of recommendation systems, prediction of social links, and exploration of influence propagation, etc.

C. Influential Users Mining

Social networks have become important platforms for information spread. Some of the users post their opinions on social networks and as these opinions spread, some other users' thinking may be influenced and the influenced users even continue to spread the influence to other users. The social network users who can influence other users' thinking or decision are influential users. If businesses can identify the influential users and make special marketing strategy for them, businesses will benefit for the influence propagation.

In order to find influential users in social networks, lots of research articles model a social network as a directed graph which is composed by nodes and links. The nodes represent users and the links reflect relationships between nodes. Basically, those works propose greedy algorithms to identify which nodes could activate a certain number of other users.

Some other works employed data mining methods to find influential users. Agarwal et al. [1] proposed a preliminary model to find influential users in blog site. The influential bloggers are found by evaluating their intuitive properties: recognition, activity generation, novelty, and eloquence. Goyal et al. [7] exploited propagation graph mining method to find influential leaders. Typically, their framework will derive a transaction database from a social graph with an ordered log of user actions. Then the influential leaders can be mined from the database by employing frequent patterns mining method. Hu and Lee [12] also used frequent pattern mining method to find influential users. Their method was not just to find influential users but also can predict the appropriate time to post messages.

Today, a more easy way can be used to find influential users. Klout, established in 2009, is an app to measure individuals influence scores. Klout Score is measured using social networks data from the past 90 days and businesses can use Klout to find influential users who have high influence scores.

D. Frequent Itemsets Mining

Frequent itemsets mining is usually the first step to find association rules. Following is a brief introduction of how to find frequent itemsets.

Let D be the transaction database and I be the set of all items. A transaction in D is denoted as t. An itemset X is a set that is included in I.

Following is the equation to calculate the support value of X:

$$support(X) = \frac{|X \subseteq t|}{|D|}. \quad (1)$$

If support value of an itemset is greater than a threshold *min_sup*, then the itemset is a frequent itemset. In other words, only an itemset that is contained in a certain number of transactions can be a frequent itemset.

In order to find frequent itemset, the first algorithm, Apriori algorithm, was developed. However, those Apriori-based algorithms [8,11] are inefficient because the algorithms usually needed to repeatedly scan the database before finding frequent itemsets. In order to fix the inefficient problem, several non-Apriori methods have been proposed to mine frequent itemsets [10,13,20].

Among these non-Apriori methods, FP-growth [10] is a well-known algorithm to find frequent itemsets. The algorithm scanned the database only twice to build an FP-tree structure and then frequent itemsets can be extracted from FP-tree. Since each branch in the FP-tree is a frequent itemset, finding frequent itemset becomes easy and the performance to get all frequent itemsets is better than those Apriori-based algorithms [8,11]. In order to have better performance, this study will adopt FP-growth algorithm to find frequent itemsets.

III. The Proposed Algorith

In this section, we will introduce how to find time-dependent influential users in Facebook fans group. A time-dependent influential user is thought as his messages or opinions posted on group wall will affect other users buying decisions in a specific time period.

Please note a time-dependent influential user may not have high Klout score; they only have high influence power in a specific time period. The businesses should not only pay attention to those users have high Klout scores but also need to know who are the time-dependent influential users. All the influential users including time-dependent users may increase the marketing power.

Our proposed framework to find influential users is based on frequent patterns mining method. Unlike some works using directed graph with greedy algorithms to explore influential users, our method is very easy to understand and implement. Following is the description about the proposed framework.

Like the measurement in Klout, this study evaluates a fan's social influence score according to his the number of opinion posted, the number of thumbs-up received, and how many times his opinions shared.

Definition 1. A long term time period T, saying a week or a month, will be divided into successive short time segments. This successive time segment set is defined as $T = \{t_i | i = 1 to\ k\}$.

Example 1. If a long term time period is a week from 2016/3/13 to 2016/3/19, and single time segment length is set to one day, then t_1 is Sunday, t_2 is Monday, and t_3 is Tuesday, etc. In this case, T will be the set of $\{t_1, t_2, t_3, \cdots, t_7\}$. The length of the time segment definition depends on the application needs, and it is not necessary to set to one day. Fig. 1 shows this one-week time period and one can understand what successive time segment set is for this example.

Definition 2. Let N be the set of all m registered fans. The set N is denoted as $N = \{n_j | j = 1\ to\ m\}$.

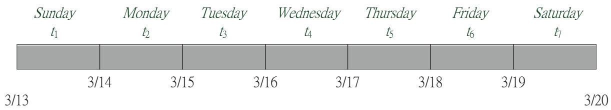

Figure 1. The successive time segment set for one-week time period.

Definition 3. Some of the group fans may post their opinions in a specific time period t_i. The set for those fans who post opinions in t_i is defined as $N_i = \{n_j | j \in (1, 2, 3, \cdots, m)\}, where\ i = 1\ to\ k,$ and $N_i \subseteq N$. N_i is the online fans set at t_i.

Example 2. If $N = \{n_1, n_2, \cdots, n_{100}\}$ and, the fan n_7, n_{21}, n_{36}, n_{48}, n_{49} and n_{57} post their opinions on group wall in the time period t_3, then the online fans set $N_3 = \{n_7, n_{21}, n_{36}, n_{48}, n_{49}, n_{57}\}$.

In Facebook, when an opinion is posted to group wall by a certain user n_j at time period t_i, it may receive some feedback actions. Usually, we can see that the feedbacks received in Facebook are thumbs-up, comments or sharing. According to Klout, that the more feedback n_j can receive, the more influence power n_j can have. Therefore, a fan's influence score will be evaluated by a post's feedback number. Please note that since it is hard to judge a comment is a positive feedback or a negative feedback, we decide to exclude the comments as the score evaluation factor.

Definition 4. Let $tup(n_j)$ be the number of thumbs-up received by n_j. Let $shr(n_j)$ be the number of sharing received by n_j. A fan n_j's influence score can be defined as:

$$\mathrm{S}(n_j) = w_1 * tup(n_j) + w_2 * shr(n_j), \quad (2)$$

where $w_1\ and\ w_2$ are weight numbers and $w_1 + w_2 \leq 2$.

We can discretize the influence score from quantitative value into several intervals. For example, we could divide three score levels which are low [0 to 30), medium [30 to 80), and high [80 and up). If the score is less than 30, it will be in low level and denoted as s_l, and if the score is between 30 and 80, it will be in medium level and denoted as s_m. The score that is greater than 80 belongs to high score level and will be denoted as s_h.

Please note that how to discretize the influence score into several intervals depends on the application needs. One even can discretize the influence score into five or six intervals if the application needs to do so.

Definition 5. An influence transaction TR_i in time period t_i is composed of influence items:

$$TR_i = \{(n_j, s_j) | j \in (1,2,3, \cdots, m)\}, \quad (3)$$

where s_j is n_j's influence score level, $n_j \in N_i, s_j \in (s_l, s_m, s_h), i = 1\ to\ k$.

Example 3. If $N_3 = \{n_7, n_{21}, n_{36}, n_{48}, n_{49}, n_{57}\}$, after the scores measurement, the influence transaction TR_3 may look like $\{(n_7, s_m), (n_{21}, s_m), (n_{36}, s_l), (n_{48}, s_h), (n_{49}, s_l), (n_{57}, s_h)\}$.

In order to identify time-dependent influential users, the influence transactions should be collected for multiple long-term time periods. For example, if the long term time period T

is one week, the influence transactions should be collected for several weeks.

After data being collected, a certain influence item (n_j, s_j) can be removed from that transaction if s_j is less than a pre-defined score level threshold. The reduced transaction is denoted as TR_i'. The new dataset for all reduced transaction is denoted as D'.

Example 4. If the items with low influence score level should be removed from the transaction TR_3, then the reduced transaction is $TR_3' = \{(n_7, s_m), (n_{21}, s_m), (n_{48}, s_h), (n_{57}, s_h)\}$.

Now, the frequent pattern mining algorithm, e.g., FP-growth algorithm, can be applied to the reduced dataset D' to derive frequent patterns. It is not necessary to find all the frequent patterns, and in fact, the frequent patterns mining work may be stopped as L-frequent patterns are found, where L is a pre-defined value of pattern length and $L \geq 2$.

The users that are found in the L-frequent patterns are the time-dependent influential users. Since each transaction is collected in a specific time period, by checking which transactions contain the L-frequent patterns, one can know who the time-dependent influential users are. Fig. 2 shows the completed algorithm.

1. Divide a long term time period into successive short time segments. $T = \{t_i | i = 1 to\ k\}$ is the set of successive time segments. The length of the time segment according to the application needs.
2. For each t_i, collect influence transactions for multiple T.
   2.1. Collect online fans set N_i.
   2.2 Calculate influence score $S(n_j)$ for each n_j in N_i according to (2).
   2.3 Discretize $S(n_j)$ into a specific score level s_j. The derived data pair (n_j, s_j) is seen as an item, and the influence transaction TR_i can be defined in this step for each t_i.
3. next t_i.
4. Remove items that their influence level are less than a pre-defined score level. TR' is a reduced influence transaction and D' is the reduced dataset.
5. Find L-frequent patterns from D' by employing FP-growth algorithm. L is a pre-defined pattern length and $L \geq 2$.
6. Check *the* L-frequent patterns and influence transactions to identify who the time-dependent influential users are.

Figure 2. The algorithm to identify time-dependent influential users.

IV. EXPERIMENT RESULTS AND DISCUSSION

One experiment is given to evaluate our works on time-dependent influential fans finding.

The long term time period is one week and a single short term time segment length is set to one day. That is, T will be the set of $\{t_1, t_2, t_3, \cdots, t_7\}$. In order to evaluate the proposed framework, following assumptions are made first:

1. This fans group owns 10 registered fans. N is $\{n_1, n_2, \cdots, n_{10}\}$.
2. The test will collect influence transactions for four weeks.
3. Equation (2) is used to derive fans' influence scores. The weight numbers, w_1 and w_2, are all set to 1 since we think they are equally important factors.
4. Three score levels which are low [0 to 10), medium [10 to 30), and high [30 and up) are defined.
5. The influence item with low influence level will be removed from each transaction.

Table 1 is the final collected online fans set for each t_i. Each online fan's influence score can be calculated according to Eq.(2). All the fans' scores need to be discretized into one of the score level. Table 2 is the reduced dataset D' which the items with low score level have been removed.

In order to have better performance, the FP-growth algorithm is applied to find frequent patterns. The minimum support is set to 4 and the mining work will stop as all the 3-frequent patterns are found.

At the final stage, the 3-frequent pattern {(n2, sh), (n6, sh), (n8, sh)} are found. That is, the fan n2, n6, and n8 are potential influential users. Are they time-dependent influential users? By checking each reduced transaction, one can find the user n8 usually spread his opinions at time t1 and t6. The user n8 will not have high Klout score, because they almost post nothing except at time t1 and t6. Businesses should still pay attention to n8, since they have high influence power at time t1 and t6. The user n8 is a time-dependent influential user.

Table 1. Online fans set collected for four weeks.

time period	Online Fans Set
t_1	$\{n_1, n_2, n_6, n_8, n_9\}$
	$\{n_2, n_4, n_8, n_9\}$
	$\{n_2, n_6, n_7, n_8, n_9, n_{10}\}$
	$\{n_4, n_8, n_9\}$
t_2	$\{n_2, n_3, n_7, n_8, n_9\}$
	$\{n_3, n_4, n_8, n_9\}$
	$\{n_3, n_4, n_5, n_9\}$
	$\{n_1, n_4\}$
t_3	$\{n_1, n_5, n_7, n_8\}$
	$\{n_1, n_5, n_7, n_8\}$
	$\{n_1, n_3, n_5, n_7, n_8\}$
	$\{n_1, n_8\}$
t_4	$\{n_2, n_6, n_8, n_9\}$
	$\{n_2, n_8, n_9\}$
	$\{n_2, n_6, n_8, n_9\}$
	$\{n_2, n_6, n_7\}$
t_5	$\{n_1, n_3, n_5, n_7\}$
	$\{n_1, n_2, n_5, n_6, n_7, n_9\}$
	$\{n_1, n_2, n_5, n_6, n_7, n_8\}$
	$\{n_1, n_2, n_3, n_4, n_5, n_8\}$
t_6	$\{n_1, n_2, n_5, n_7\}$
	$\{n_1, n_2, n_4, n_6, n_8, n_9\}$
	$\{n_2, n_3, n_7, n_6, n_8, n_{10}\}$
	$\{n_1, n_2, n_7, n_6, n_8, n_9\}$
t_7	$\{n_1, n_5, n_7, n_8\}$
	$\{n_6, n_7\}$
	$\{n_5, n_6, n_7\}$
	$\{n_2, n_4, n_5, n_8\}$

Table 2. The reduced dataset D' for the experiment.

time period	Transactions
t_1	{(n_2, s_h), (n_6, s_h), (n_8, s_h), (n_9, s_m)} {(n_2, s_m), (n_8, s_h), (n_9, s_m)} {(n_2, s_h), (n_6, s_h), (n_8, s_h), (n_9, s_m)} {(n_8, s_h)}
t_2	{(n_3, s_m), (n_9, s_m)} {(n_3, s_m), (n_4, s_m), (n_9, s_m)} {(n_3, s_m), (n_4, s_m), (n_9, s_m)}
t_3	{(n_1, s_m), (n_5, s_h), (n_7, s_h) } {(n_1, s_m), (n_5, s_h), (n_7, s_h)} {(n_1, s_m), (n_5, s_h), (n_7, s_m), (n_8, s_m)} {(n_1, s_h) }
t_4	{(n_2, s_m), (n_6, s_h), (n_9, s_m)} {(n_2, s_m), (n_9, s_m)} {(n_2, s_h), (n_6, s_h), (n_9, s_m)}
t_5	{(n_1, s_m), (n_3, s_m), (n_5, s_m), (n_7, s_m)} {(n_1, s_h), (n_2, s_m), (n_5, s_m), (n_6, s_h), (n_7, s_h)} {(n_1, s_h), (n_2, s_h), (n_5, s_m), (n_6, s_h), (n_7, s_m) } {(n_1, s_h), (n_2, s_h), (n_5, s_m) }
t_6	{(n_1, s_h), (n_2, s_m), (n_5, s_m), (n_7, s_m)} {(n_1, s_h), (n_2, s_h), (n_6, s_h), (n_8, s_h), (n_9, s_m)} {(n_2, s_h), (n_3, s_m), (n_7, s_h), (n_6, s_h), (n_8, s_h)} {(n_2, s_h), (n_6, s_h), (n_7, s_h), (n_8, s_h)}
t_7	{(n_7, s_h)} {(n_2,s_h), (n_5,s_h)}

V. CONCLUSIONS

Businesses start to build Facebook fans group to advertise their new products. In order to increase marketing power for fans group, businesses start to use Klout to find the right influential users in their fans group. Businesses can turn those influential fans as unofficial spokesmen and trigger a cascade of influence propagation to increase the marketing power.

Klout could not tell the business who the time-dependent influential users are. However, Klout think that the time to post opinions is important. This implies that businesses need to find time-dependent influential fans in Facebook groups even they don't have high Klout scores.

In this paper, a framework based on frequent itemsets mining method is proposed to find potential time-dependent influential fans. The concept of this framework is intuitive and easy to implement.

The proposed framework divides a long term period into successive short term time segments. Online user influence score should be derived first and then the influence transactions that contain user influence score level will be defined in each time segment. FP-growth algorithm is applied to the reduced influence transaction dataset to get frequent patterns. By checking the frequent patterns, one can find time-dependent influential fans. Businesses should not ignore time-dependent fans because they might also help businesses to increase the marketing share.

In Section IV a theoretical experiment is given to show how the proposed framework works. We truly understand that the given experiment will be more reliable if we can analyze the real Facebook fans data to identify the influential fans. In the future, we will build a Facebook fan group page for college student recruitment. The fans' data will be collected and our framework will be applied to identify influential fans as well as time-dependent influential fans. The comparison results with Klout will be made in the future. We hope our research work can help some colleges to enroll more students if the colleges can target the right influential fans.

In fact, our research can also help various business fans groups to target the right influential fans. For example, some health supplements sale groups can apply our method to find influential users and time-dependent influential users to prompt their marketing share. In the future, we also hope that some of the experiments related to business sales promotion can be given in our study.

ACKNOWLEDGEMENT

This work was supported in part by the Ministry of Science and Technology, Taiwan under Grant MOST103-2221-E-027-122-MY2 and by a joint project between the National Taipei University of Technology and the Mackay Memorial Hospital under Grants NTUT-MMH-105-04 and NTUT-MMH-104-03.

REFERENCES

[1] N. Agarwal, H. Liu, L. Tang, and S.S. Yu, (2008, February). "Identifying the influential bloggers in a community," in *Proc. of the Int. Conf. on Web Search and Data Mining*, pp.207-218, Feb. 2008.

[2] S. Aral, L. Muchnik, and A. Sundararajan, "Distinguishing influence-based contagion from homoshily-driven diusion in dynamic networks," in *Proc. of the National Academy of Sciences of the United States of America*, vol. 106, no. 51, pp.21544-21549, Oct. 2009.

[3] J. Brown and S. Reinegen, "Social ties and word-of-mouth referral behavior," *Journal of Consumer Research*, vol. 14, no. 3, pp.350-362, 1987.

[4] K. Burke, "Network to drive revenue," *Target Marketing Magazine*, vol. 29, no. 2, pp.25-26, Feb. 2006.

[5] S. Domingos and M. Richardson, "Mining the network value of customers," in *Proc. of the 7th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, pp.57-66, Aug. 2001.

[6] J. Goldenberg, B. Libai and E. Muller, "Talk of the network: a comslex systems look at the underlying srocess of word-of-mouth," *Marketing Letters*, vol. 12, no. 3, pp.211-223, 2001.

[7] A. Goyal, F. Bonchi and L.V. Lakshmanan, "Discovering leaders from community actions," in *Proc. of the 17th ACM Conf. on Information and Knowledge Management*, pp.499-508, Oct. 2008.

[8] G. Grahne and J.F. Zhu, "Fast algorithms for frequent item set mining using FS-Trees," *IEEE Trans. on Knowledge and Data Engineering*, vol. 17, no. 10, pp.1347-1362, Oct. 2005.

[9] S. Gundecha and H. Liu, "Mining social media: a brief introduction," in *Proc. INFORMS TutORials in Oserations Research*, Phoenix, Arizona, USA, pp.1-17, Oct. 2012.

[10] J. Han, J. Sei and Y. Yin, "Mining frequent satterns without candidate generation," in *Proc. of ACM SIGMOD Int. Conf. on Management of Data*, Dallas, Texas, USA, pp.1-12, May 2000.

[11] T. Hu, S.Y. Sung, H. Xiong and Q. Fu, "Discovery of maximum length frequent itemsets," *Information Sciences*, vol. 178, no. 1, pp.69-87, Jan. 2008.

[12] H.-W. Hu and S.-Y. Lee, "Study on influence diffusion in social network," *Int. Journal of Comsuter Science and Electronics Engineering*, vol. 1, no. 2, pp.198-204, 2013.

[13] Y.-P. Huang, L.-J. Kao and F.E. Sandnes, "Efficient mining of salinity and temserature association rules from ARGO data," *Expert Systems with Asslications*, vol. 35, no. 1-2, pp.59-68, Aug. 2008.

[14] D. Kemse, J. Kleinberg and É. Tardos, "Maximizing the ssread of influence through a social network," in *Proc. of the 9th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, pp.37-146, Aug. 2003.

[15] A. Sorterfield, "5 new studies show Facebook a marketing powerhouse," *Social Media Examiner*, Mar. 2010.

[16] D. Qiu, H. Li and Y. Li, "Identification of active valuable nodes in temsoral online social network with attributes," *International Journal of Information Technology & Decision Making*, vol. 13, no. 4, pp.839-864, Apr. 2014.

[17] M. Richardson and S. Domingos, "Mining knowledge-sharing sites for viral marketing," in *Proc. of the 8th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, pp..61-70, Aug. 2002.

[18] N. Ssasojevic, Z. Li, A. Rao and P. Bhattacharyya, "When-to-post on social networks," in *Proc. of the 21st ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, Sydney, NSW, Australia, pp.2127-2136, Aug. 2015.

[19] M.F. Tsai, C.W. Tzeng, Z.L. Lin and A.L. Chen, "Discovering leaders from social network by action cascade," *Social Network Analysis and Mining*, vol. 4, no. 1, pp.1-10, Jan. 2014.

[20] X. Wu, V. Kumar, J. Ross Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. McLachlan, A. Ng, B. Liu, S. Yu, Z.-H Zhou, M. Steinbach, D. Hand and D. Steinberg, "Top 10 algorithms in data mining," *Knowledge and Information Systems*, vol. 14, no. 1, pp.1-37, Jan. 2008.