2018 IEEE International Conference on Systems, Man, and Cybernetics

An Effective Social Network Sentiment Mining Model for Healthcare Product Sales Analysis

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

Yo-Ping Huang
Department of Electrical Engineering
National Taipei University of Technology
Taipei, 10608 Taiwan
Email: yphuang@ntut.edu.tw

***Abstract*—Social network websites have become important marketing platforms for studying business models. When impressive advertisements are posted on the platforms, social network users like to comment or post their experiences as part of feedback about the products. Those interesting opinions will exert influences on other users' buying decisions but some may remain commented. This invokes people's interests to dig out interesting relationship between sentiment of social network users and the volume of product sales. Then, they can apply the discovered patterns of sentiment and sales to predict users' buying behaviors. This study proposes a framework based on data mining method to find interesting patterns of sentiment and sales. The proposed model starts by defining sentiment topics with their corresponding terms and then follows by a fuzzy model to infer the sentiment scores for user opinions. Each transaction in the database is transformed to attach with public sentiment scores, influential users' sentiment scores and volume of product sales. To better obtain the relationship among public sentiment, users' sentiment and volume of product sales, a mining method of inter-transaction association rules is considered to extract the interesting patterns of sentiment and sales. Two case studies are given to verify the effectiveness of the proposed method.**

***Keywords*— inter-transactions; sentiment analysis; opinion aanlysis, fuzzy logic; sales prediction**

I. Introduction

Social networks have become new influential propagation media [1, 14]. The influence propagation problem has been investigated for years [2-3, 5-6, 9, 16, 18-19]. People like to post opinions, comments, political issues, breaking news, essays, reports, or stories to share with others [9]. Some opinions may trigger thousands of followers' discussions and exert certain degrees of influence on user reaction. A typical example is that many people trust some political messages in social networks and they initiate civic movement to express their collective consciousness [4]. In fact, one can see that the social network users not only consult the opinions to make decisions but also forward opinions to other users that expand the influence.

Influence propagation makes the social networks become a powerful marketing channel [15, 17]. Many businesses set up their fan pages or blogs targeted on promoting their products. Advertisements are not the most efficient marketing tool in social networks. By contrast, the most important issue is about fans' opinions. Many social network users like to post their opinions to acknowledge their experiences on the products. Many users browsed the related opinions in the social networks before they decided to buy the products. These users want to make sure that they buy the right products. Hence, negative feedbacks on the products may deter users from shopping. By contrast, positive reviews may shorten the user's decisions to buy the products. This is apparent about the healthcare products, especially some people may exaggerate the effectiveness of the products. As long as certain users continue to forward the influential users' opinions to others, this definitely will play sort of influences on other users' decisions on whether to buy the products.

Among the influence propagation research, several studies have been presented to identify who the influential users of social networks are [1, 7, 12]. Some investigated on what types of social media advertisement were easier to have users' attention and biased them to forward products information to other users [22]. Social network users' sentiment analysis is an important research issue and most research evaluated users' sentiment through some specific terms in messages. Understanding users' sentiment can help the businesses to analyze how much their customers love a product [21].

This paper is focused on predicting users' buying intention based on sentiment information in the database. So far, only a few studies have investigated to relate users' sentiment and their buying behavior. If the relations between users' sentiment and volume of product sales can be found, these relational patterns of sentiment and sales may help businesses to explore their social network users' purchase behavior. This invokes us to start by proposing a framework to quantify users' sentiment scores, followed by exploiting a mining method of intertransaction association rules to discover interesting patterns between sentiment and sales. The users' sentiment scores are derived from the terms in their posted opinions. One single opinion or message can be seen as several predefined terms composition. Each term can be transformed into fuzzy linguistic terms, and after the procedure of defuzzication is taken, a crisp sentiment value can be derived for the term. An

2577-1655/18/$31.00 ©2018 IEEE
DOI 10.1109/SMC.2018.00370

opinion's sentiment score is the summarization of all predefined terms' sentiment value.

This study is targeted to find the relational patterns between sentiment and sales such as "if we have the high positive public sentiment today, we will get more volumes of product sales three days later." Previously, association rule mining methods were applied to find causal relationship rules; however, they cannot find aforementioned patterns which reveal the associations among items in different transactions. In order to tackle this problem, we apply a mining algorithm of intertransaction association rules to dig out patterns having relations between users' sentiment and sales volume among transactions.

This paper is organized as follows. Section II introduces related works about social networks mining; sentiment analysis and mining algorithm of intertransaction association rules. Section III explains the sentiment evaluation method and procedures to find the interesting patterns between sentiment and sales. Section IV presents two case studies to illustrate how to derive the patterns between sentiment sales and how to apply them to predict the sales volume of products. Section V concludes the paper.

II. Related Works

A. Social Networks Data Mining

The research on social network data mining has been investigated for years and many research topics have been developed in social network data mining community such as identification of influential users, evaluation of users' sentiment, improvement of recommendation systems, and analysis of social networks. However, applying traditional data mining methods to directly analyze social network data was not easy due to the unstructured data types and the huge amount of data [9]. In order to tackle the problems, many novel methods have been proposed to engage in social network data mining. For example, directed graphs are often used to describe the social network in influence propagation models. The social network data can then be transformed to a format that is easier to be manipulated by data mining methods.

B. Social Network Sentiment Analysis

Social network sentiment analysis relied on natural language processing tools to analyze people's feeling behind what they post on social network. Bollen *et al.* [24] proposed a method to evaluate public sentiment in Twitter to understand public emotions. Their work assumes that public emotions will affect individual behavior and one can use public sentiment information to predict stock market. In fact, some methods have been proposed to predict product sales, book sales or movie sales using online sentiment information [24].

However, most previous research only relied on all the messages in social network to analyze public sentiment [21]. They did not consider users' influence. Since influential users' words may become the mainstream of public emotions, Luneva *et al.* thought the analysis of social network user's sentiment should take the users' influence into account in order to improve the accuracy of social network users' sentiment analysis [21].

Today, most studies focused their research on evaluating sentiment behind users' words. There are few studies to explore the relationship among public sentiment, influential users' sentiment and sales volume. In this paper, we intend to propose a framework to find the sentiment-sales patterns that can reveal the relationship among public sentiment, influential users' sentiment and product sales volume.

C. Inter-transaction Association Rule

The association rule that is found by traditional mining methods can only show the items relation within the same transaction. For example, the rule "if X bought diapers, then X will buy beers," only tells the relationship between the two items, diaper and beer, within the same transaction.

However, there are chances that people need to know the relationships among the items from the different transactions. For example, in this study we can find the rule like "if public sentiment is positive today, the sales volume will increase two days later." This rule was discovered from associations among items in different transactions. An association rule extracted from items within the same transaction was termed as an intra-transaction association, whereas a rule obtained from different transactions was called an inter-transaction association rule.

The discovered intertransaction association rules help people to explore the relational evolution among items in spatial or temporal domain. For example, if the intertransaction is "the positive sentiment will cause the increase of sales volume a couple of days later," one can realize that the positive sentiment will not boost the sales volume immediately.

A brief introduction of intertransaction association rules mining is given as follows [23].

First, assuming that $I=\{i_1, i_2, \ldots, i_k\}$ is a set of items. We denote D as a dimensional attribute and $Dom(D)$ is the domain of D. A 1-dimensional database is a database containing records in the form (d, I_j), where $d \in Dom(D)$ and $I_j \subseteq I$.

The dimensional attribute is used to describe the temporal or spatial information for the records. Fig. 1 showed an example of 1-dimensional database. In this example, we assume the dimensional attribute is one single day, and the items association may exist in any *p-day* intervals. Theoretically, inter-transaction association rules that span any arbitrary intervals can be found. But it is time-consuming to find all rules that span all intervals. To tackle the problem of mining all rules that span all intervals, the sliding window is introduced to form inter-transactions. That is, the concept of sliding window leads the algorithm of inter-transaction association rules mining to find the rules within the inter-transactions formed by sliding window, and one can adjust the sliding window to find different set of inter-transaction association rules.

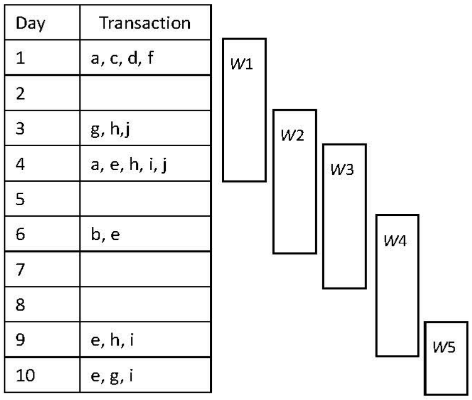


Fig. 1. A one-dimensional dataset example with the sliding window being set to 4-day.

In Fig. 1, a sliding window is set to 4-day interval and each sliding window W forms an intertransaction. We can apply association rules mining algorithm to obtain intertransaction rules from those transactions as shown in Table I [8, 10-11, 13, 20].

D. Fuzzy Logic and Linguistic Term

Fuzzy logic has long been applied to approximate human reasoning. The idea behind fuzzy logic is to define a fuzzy term set to describe the real world objective term. For example, a person's weight can be described by the fuzzy terms, *heavy*, *medium*, and *light*. There are two men; person A's weight is 50kg and B's weight is 52kg. We usually say that both of them are not heavy or they have no difference in weight. But if we use fuzzy logic to describe their weight, the expressions may look different. One can define membership functions to transform A's weight to the fuzzy terms as {heavy, 0.2}, {medium, 0.5} and {light, 0.3}. That is, the crisp value 50kg is transformed to three fuzzy terms with their corresponding membership degrees. B's weight is transformed to the fuzzy linguistic terms as {heavy, 0.25}, {medium, 0.5} and {light, 0.25}. So, they have different descriptions in terms of fuzzy weight levels.

Fuzzy logic can be applied to the task of semantic analysis [21]. A term can be represented as membership functions of the linguistic terms. For example, the term "good" can be represented as {(positive, 0.8), (objective, 0.2), (negative, 0)}. In order to get a sentiment value for the term good, the defuzzification procedure should be taken to transform the fuzzy membership degrees into crisp values [21].

III. Sentiment Scores Evaluation

In this section, we will propose a framework to get user opinion's sentiment degrees and user sentiment scores. Because influential users' sentiment may become the mainstream of public opinion, we are not only to evaluate public sentiment but also to evaluate influential uses' sentiment. Please note that our research goal is to find the relationships among public sentiments, influential users' sentiments and products sales volume. This section is to introduce the method that can derive the sentiment degree. Based on the work from [21], our proposed method will also predefine a term set that is related to the topic we are interesting in. Each user's opinion will be transformed to a set of predefined terms.

TABLE I. An Inter-transaction Database. Inter-transactions are Transformed from Fig. 1.

ID	Inter-Transaction
M_1	{a(0), c(0), d(0), f(0), g(2), h(2), j(2)}
M_2	{g(0), h(0), j(0), a(1), e(1), h(1), i(1), j(1)}
M_3	{a(0), e(0), h(0), i(0), j(0), b(2),e(2}
M_4	{b(0), e(0), e(3), h(3), i(3) }
M_5	{e(0), h(0), i(0),e(1), g(1), i(1)}

The following gives definitions to introduce the derivation of an opinion's sentiment degree and public sentiment scores from the predefined terms.

Definition 1. Let $K = \{k_i | i = 1,2,\ldots,n\}$ be a set of subject characteristics and n is the total number of characteristics. k_i is a subject characteristic. Each characteristic can be represented as a set of terms. Let $T_i = \{t_{ij} | i = 1,2,\ldots,n, j = 1,2,\ldots,m\}$ be a set of predefined terms and m is the number of terms related to the *ith* characteristic.

A user's opinion should be converted to a set of predefined terms according to its subject characteristics. For example, Taiwan pearl milk tea lovers are talking about the tea drink on fan page. We can define the characteristic set as {"*the taste quality*", "*the price deal*", "*the quality consistency*"}. The term set for the characteristic "*the taste quality*" can be defined as {"*delicious*", "*fair*", "*awful*"}. The opinion, I like the pearl milk tea; it tastes delicious, can be converted to the term set, {"*delicious*"}.

According to the work in [21], each term should be transformed to trapezoidal membership functions of the linguistic term.

Definition 2. We define a linguistic term set. Let $U = \{p, o, n\}$ be a linguistic term set, where p, o, and n represent positive, objective, and negative, respectively.

Definition 3. The normalization trapezoidal membership function is defined as $\mu_p + \mu_o + \mu_n = 1$, where μ_p, μ_o, μ_n are the membership function degrees for p, o, n, respectively.

Similar to the work in [21], we use SentiWordNet 3.0 to assign the degree of positivity, objectivity, and negativity for a term. The opinion in the above example is converted to the term set {"*delicious*"}, and it can be represented as {p=0.6, o=0.25, n=0.15}. The next step is to derive sentiment value for term delicious. One can perform the defuzzification process to obtain a crisp sentiment value v_{ij}.

Definition 4. Assume that T is the set of all terms we defined. A user's opinion can be transformed to a subset $ST \subseteq T$. The opinion's sentiment degree *deg* can be calculated as follows:

$$deg = \frac{1}{|ST|} \sum_{t_{ij}} v_{ij} \cdot (-1)^{\alpha} \quad (1)$$

where v_{ij} is the value for t_{ij}. $\alpha = 1$, if the v_{ij} is positive else $\alpha = -1$. $|ST|$ is the cardinality number of ST. deg $\in [-1,1]$. The term with objective value will be excluded from the sentiment degree derivation process.

Definition 5. The sentiment score, $SCORE(period_k, n_j)$, for a specific user n_j is the sum of all his opinions' sentiment degrees in the period $period_k$. For example, user n_3 posts his opinions about a product from April 1st to April 7th, 2018. His sentiment score on April 1st (day 1) is obtained from the summation of all his opinions' sentiment degrees on April 1st and is denoted as $SCORE(day_1, n_3)$.

Definition 6. The summation of all the fan page or blog users' sentiment in a specific time period is called as public sentiment in that specific time period and is denoted as $T\text{-}SCORE(period_k)$. For example, the fan page public sentiment on April 1st is denoted as $T\text{-}SCORE(day_1)$.

IV. Experiment Results

Two experiments are given in this section to illustrate the procedures in evaluating users' sentiment and the mining of relational patterns between sentiment and sales.

(1) The first experimental data were obtained from the Facebook beverages shop fan page. There are 79 fans in this group, and 3 fan page users who have top 3 Klout scores are chosen for influential user sentiment evaluation. All the posted opinions that were related to pearl milk tea were evaluated.

(2) We set two subject characteristics as {"*the taste quality*", "*the price deal*"}.

(3) Two term sets need to be defined in this experiment. The first term set T_1 for "*the taste quality*" is {"*delicious*", "*fair*", "*awful*"}. The second set T_2 for "*the price deal*" is {"*expensive*", "*reasonable*", "*cheap*"}.

(4) The linguistic term set and trapezoidal membership functions were defined in the experiment according to the Definitions 2 and 3.

(5) The experiment period was set to 7 days. The users' sentiment scores were recorded day by day.

(6) It is not easy for association rules mining algorithm to deal with quantitative data. Hence, influential user's sentiments and public sentiment quantitative scores must be mapped into a specific interval in order to allow the successive intertransaction association rules mining work can be done in an easier way. In this experiment, the quantitative sentiment scores were mapped to one of the following intervals: {*hn* (high negative), *mn* (medium negative), *ln* (low negative), *lp* (low positive), *mp* (medium positive), *hp* (high positive)}.

(7) In this experiment, the quantitative sales volume was mapped to one of the following intervals: {*lv* (low volume), *mv* (medium volume), *hv* (high volume)}.

Fig. 2 shows the seven original records from day 1 to day 7. Each record kept the information of three influential users' sentiment scores, public sentiment score, and sales volume. For example, in day 1, the transaction is recorded as {*hpU1* (user1 high positive), *hpU2* (user2 high positive), *lpU3* (user3 low positive), *lnPub* (public low negative), *mv* (medium volume)}. Five intertransactions were obtained from setting the sliding window to three days. Table II shows all the inter-transactions that transformed from Fig. 1. Now, we can apply the association rules mining algorithm to find sentiment-sales patterns from inter-transaction dataset. Assuming the minimum support to be 3, Table III shows the 1-frequent itemsets, 2-frequent itemsets and 3-frequent itemsets. From Table III, one can find several interesting patterns. For example, if user 2 had high-positive sentiment in two successive days, the sales volume will be high in the third day. Another interesting finding is that the public sentiment does not show up in any frequent item set. That is, the public sentiment has no associations with the sales volume in this case. Normally, we think the positive public sentiment will lead high sales volume but the fan page users do not follow the rule. By contrast, the influential users play important role in marketing sales.

Fan pages or blogs that set up by different businesses may have different findings on the relations between sentiments and sales volume. The proposed method can be easily applied to find business patterns to track customers' purchase behavior. To further verify the applicability of the presented work another case study is given. Taiwan has become an aged society starting from March 2018. Many people like to talk about healthcare foods that induced companies to set up Facebook fan pages to advertise their products. We can also do the sentiment analysis task and then predict customers' purchase behavior for a specific dietary supplement food.

(1) The second experimental data were obtained from the Facebook health food shop fan page. There are 122 fans in this group, and 5 influential users' opinions are used for influential user sentiment evaluation. All opinions related to a specific vitamin product were evaluated.

(2) The set of subject characteristics, K, will be {"*the nutritional quality*", "*the vitamin effectiveness*"}.

(3) Two term sets need to be defined in this experiment. The first term set T_1 for "*the nutritional quality*" is {"*rich*", "*fair*", "*poor*"}. The second set T_2 for "*the vitamin effectiveness*" is {"*useful*", "*fair*", "*bad*"}.

(4) The linguistic term set and membership functions were defined in the experiment according to the Definitions 2 and 3.

(5) The experiment period was set to 7 days. The users' sentiment scores were recorded day by day. In this experiment, the public sentiment was separated to male public sentiment and female public sentiment.

(6) The quantitative sentiment scores were mapped to one of the following intervals: {*hn* (high negative), *mn* (medium negative), *ln* (low negative), *lp* (low positive), *mp* (medium positive), *hp* (high positive)}.

(7) The quantitative sales volume was mapped to one of the following intervals: {*lv* (low volume), *mv* (medium volume), *hv* (high volume)}.

Fig. 3 shows the original records. Table IV shows all the inter-transactions that were transformed from Fig. 3. Assuming the minimum support to be 4, Table V shows the 1-frequent itemsets, 2-frequent itemsets and 3-frequent itemsets.

The experimental results showed that the opinions from influential user 4 could lead the medium volume sales. Another finding was that even the male public sentiment was low positive; it still led the medium volume sales. It seems that men talk about health food related issues and like to buy health food.

V. Conclusions

Sentiment analysis is an important research issue related to opinion analysis. The fans' sentiment data can help users to capture the stream on fans' thinking. This paper started by investigating a method based on fuzzy logic to infer fans sentiment scores, followed by applying mining algorithm of intertransaction association rules to discover interesting patterns between sentiment and sales. Unlike previous research, not only public sentiment can be found but also influential users' sentiment can be identified. As a result, this study can extract the association rules among influential users' sentiments, public sentiments and sales volume.

Both illustrated datasets were obtained from the Facebook fan pages and the results verified the effectiveness of the proposed framework. In the first experiment, initially, we may think that public sentiments are also an important factor to affect purchase volume. However, based on the mining results, we found that influential user 2 was the key person to affect customers' purchase volume. In the second experiment, we gave the idea of separating the public sentiment to male sentiment and female sentiment in order to explore the difference between male and female behavior in purchasing health food. All the mining results also showed that we can not only rely on public sentiment to predict social network users' behavior but also influential users' sentiment should be taken into the sentiment analysis.

The proposed method can be applied to other product sales analysis. The extracted sentiment-sales patterns help the businesses to plan the marketing plan according to their users' purchase behaviors. Although the mining results may vary from company to company, we believe that the proposed method can help enterprise to design a better plan to sale the products in their social network blogs.

ACKNOWLEDGEMENT

This work was supported in part by the Ministry of Science and Technology, Taiwan under Grant MOST106-2221-E-027-001-, in part by the joint project between the National Taipei University of Technology and Mackay Memorial Hospital under Grants NTUT-MMH-105-04 and NTUT-MMH-106-03, and in part by the joint project between the National Taipei University of Technology and Chang Gung Memorial Hospital under Grant NTUT-CGMH-106-05.

REFERENCES

[1] D.M. Romero, W. Galuba, S. Asur, & B.A. Huberman, "Influence and passivity in social media," in *Proc. of Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Berlin, Heidelberg, pp.18-33, September, 2011.

[2] E. Bakshy, J.M. Hofman, W.A. Mason, & D.J. Watts, "Everyone's an influencer: quantifying influence on twitter," in *Proc. of the fourth ACM international conference on Web search and data mining*, pp.65-74, ACM, February, 2011.

[3] D. Kempe, J. Kleinberg, & E. Tardos, "Maximizing the spread of influence through a social network," in *Proc. of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp.137-146, ACM, August, 2003.

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

[5] E.E. Luneva, P.I. Banokin, & A.A. Yefremov, "Evaluation of social network user sentiments based on fuzzy sets," in *Proc. of IOP Conference Series: Materials Science and Engineering*, vol. 93, no. 1, IOP Publishing, 2015.

[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] F. Bonchi, C. Castillo, A. Gionis, & A. Jaimes, "Social network analysis and mining for business applications," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp.22-34, 2011.

[8] C.F. Ahmed, S.K. Tanbeer, B.S. Jeong, & Y.K. Lee, "Efficient tree structures for high utility pattern mining in incremental databases," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 12, pp.1708-1721, 2009.

[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. Pei, B. Mortazavi-Asl, H. Pinto, Q. Chen, U. Dayal, & M. Hsu, "Prefixspan: Mining sequential patterns efficiently by prefix-projected pattern growth," in *Proc. of the 17th international conference on data engineering*, pp.215-224, April, 2001.

[11] G. Pyun, U. Yun, & K.H. Ryu, "Efficient frequent pattern mining based on linear prefix tree," *Knowledge-Based Systems*, vol. 55, pp.125-139, 2014.

[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] M. Martínez-Ballesteros, F. Martínez-Álvarez, A. Troncoso, & J.C. Riquelme, "An evolutionary algorithm to discover quantitative association rules in multidimensional time series," *Soft Computing*, vol. 15, no. 10, pp.2065-2086, 2011.

[14] J. Leskovec, L.A. Adamic, & B.A. Huberman, "The dynamics of viral marketing," *ACM Transactions on the Web*, vol. 1, no. 1, pp.5-51, 2007.

[15] A. Sorterfield, "5 new studies show Facebook a marketing powerhouse," http://www.socialmediaexaminer.com, Mar. 2010.

[16] A. Ozcan, & S.G. Oguducu, "Link prediction in evolving heterogeneous networks using the NARX neural networks," *Knowledge and Information Systems*, pp.1-28, 2017.

[17] A.J. Mary, & L. Arockiam, "ASFuL: Aspect based sentiment summarization using fuzzy logic," in *Proc. of the Int. Conf. on Algorithms, Methodology, Models and Applications in Emerging Technologies*, pp.1-5, February, 2017.

[18] M.R. Karimi, E. Tavakoli, M. Farajtabar, L. Song, & M. Rodriguez, "Smart broadcasting: Do you want to be seen?," in *Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.1635-1644, ACM, August 2016.

[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] H. Chen, R.H. Chiang, & V.C. Storey, "Business intelligence and analytics: from big data to big impact," *MIS quarterly*, pp.1165-1188, 2012.

[21] E.E. Luneva, P.I. Banokin, A.A. Yefremov, and T. Tiropanis, "Method of evaluation of social network user sentiments based on fuzzy logic", *Key Engineering Materials*, vol. 685, pp.847-851, Feb. 2016.

[22] I.P. Cvijikj and F. Michahelles, "Online engagement factors on Facebook brand pages," *Social Network Analysis and Mining*, vol. 3, no. 4, pp.843-861, 2013.

[23] A.K.H. Tung, H. Lu, J. Han and L. Feng, "Efficient mining of intertransaction association rules," *IEEE Trans. on Knowledge and Data Engineering*, vol. 15, no. 1, pp.43-56, Jan./Feb. 2003.

[24] J. Bollen, H. Mao, and X.-J. Zeng, "Twitter mood predicts the stock market," *Journal of computational science*, vol. 2, no. 1, pp.1-8, March 2011.

Day	Records
1	(hpU1, hpU2, lpU3, lnPub, mv)
2	(hpU2, lpPub, lv)
3	(lpU1, lpU2, mnPub, hv)
4	(lpU1, hpU2, mpU3, mpPub, hv)
5	(hpU2, hpPub, mv)
6	(mpU1, hpU2, mpU3, lpPub, hv)
7	(hpU3, lpPub, hv)

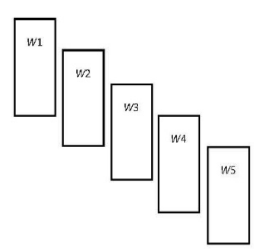


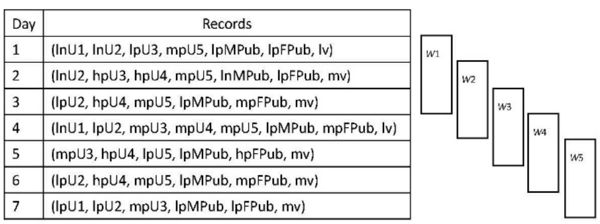
Fig. 2. The original record data set for beverages shop fan page.

TABLE II. THE INTER-TRANSACTION DATA SET. ALL THE TRANSACTIONS ARE TRANSFORMED FROM FIG. 2.

ID	Inter-Transaction
M_1	{hpU1(0), hpU2(0), lpU3(0), lnPub(0), mv(0), hpU2(1), lpPub(1), lv(1), lpU1(2), lpU2(2), mnPub(2), hv(2)}
M_2	{hpU2(0), lpPub(0), lv(0), lpU1(1), lpU2(1), mnPub(1), hv(1), lpU1(2), hpU2(2), mpU3(2), mpPub(2), hv(2)}
M_3	{lpU1(0), lpU2(0), mnPub(0), hv(0), lpU1(1), hpU2(1), mpU3(1), mpPub(1), hv(1), hpU2(2), hpPub(2), mv(2)}
M_4	{lpU1(0), hpU2(0), mpU3(0), mpPub(0), hv(0), hpU2(1), hpPub(1), mv(1), mpU1(2), hpU2(2), mpU3(2), lpPub(2), hv(2)}
M_5	{hpU2(0), hpPub(0), mv(0), mpU1(1), hpU2(1), mpU3(1), lpPub(1), hv(1), hpU3(2), lnPub(2), hv(2)}

TABLE III. THE FREQUENT ITEM SETS MINED FROM TABLE II. THE MINIMUM SUPPORT VALUE IS SET TO 3.

1-item sets	{hpU2(0)}, {hpU2(1)}, {hpU2(2)}, {hv(1)}, {hv(2)}
2-item sets	{hpU2(0), hv(2)}, {hpU2(1), hv(2)}
3-item sets	{hpU2(0), hpU2(1), hv(2)}



Day	Records
1	(lnU1, lnU2, lpU3, mpU5, lpMPub, lpFPub, lv)
2	(lnU2, hpU3, hpU4, mpU5, lnMPub, lpFPub, mv)
3	(lpU2, hpU4, mpU5, lpMPub, mpFPub, mv)
4	(lnU1, lpU2, mpU3, mpU4, mpU5, lpMPub, mpFPub, lv)
5	(mpU3, hpU4, lpU5, lpMPub, hpFPub, mv)
6	(lpU2, hpU4, mpU5, lpMPub, mpFPub, mv)
7	(lpU1, lpU2, mpU3, lpMPub, lpFPub, mv)

Fig. 3. The original record data set for health food shop fan page.

TABLE IV. THE INTER-TRANSACTION DATA SET. ALL THE TRANSACTIONS ARE TRANSFORMED FROM FIG. 3.

ID	Inter-Transaction
M_1	{lnU1(0), lnU2(0), lpU3(0), mpU5(0), lpMPub(0), lpFPub(0), lv(0), lnU2(1), hpU3(1), hpU4(1), mpU5(1), lnMPub(1), lpFPub(1), mv(1), lpU2(2), hpU4(2), mpU5(2), lpMPub(2), mpFPub(2), mv(2)}
M_2	{lnU2(0), hpU3(0), hpU4(0), mpU5(0), lnMPub(0), lpFPub(0), mv(0), lpU2(1), hpU4(1), mpU5(1), lpMPub(1), mpFPub(1), mv(1), lnU1(2), lpU2(2), mpU3(2), mpU4(2), mpU5(2), lpMPub(2), mpFPub(2), lv(2)}
M_3	{lpU2(0), hpU4(0), mpU5(0), lpMPub(0), mpFPub(0), mv(0), lnU1(1), lpU2(1), mpU3(1), mpU4(1), mpU5(1), lpMPub(1), mpFPub(1), lv(1), mpU3(2), hpU4(2), lpU5(2), lpMPub(2), hpFPub(2), mv(2)}
M_4	{lnU1(0), lpU2(0), mpU3(0), mpU4(0), mpU5(0), lpMPub(0), mpFPub(0), lv(0), mpU3(1), hpU4(1), lpU5(1), lpMPub(1), hpFPub(1), mv(1), lpU2(2), hpU4(2), mpU5(2), lpMPub(2), mpFPub(2), mv(2)}
M_5	{mpU3(0), hpU4(0), lpU5(0), lpMPub(0), hpFPub(0), mv(0), lpU2(1), hpU4(1), mpU5(1), lpMPub(1), mpFPub(1), mv(1), lpU1(2), lpU2(2), mpU3(2), lpMPub(2), lpFPub(2), mv(2)}

TABLE V. THE FREQUENT ITEM SETS MINED FROM TABLE IV. THE MINIMUM SUPPORT VALUE IS SET TO 4.

1-item sets	{mpU5(0)}, {lpMPub(0)}, {hpU4(1)}, {mpU5(1)}, {lpMPub(1)}, {mv(1)}, {lpU2(2)}, {lpMPub(2)}, {mv(2)}
2-item sets	{lpMPub(0), mv(2)}, {hpU4(1), mv(1)}, {lpMPub(2), mv(2)}
3-item sets	N/A