Spring 4-10-2019

Personalized Recommendations on Twitter based on Explicit User Relationship Modelling

Abdullah Alshammari
University of Brighton, A.Alshammari1@uni.brighton.ac.uk

Nikolaos Polatidis
University of Brighton, n.polatidis@brighton.ac.uk

Stelios Kapetanakis
University of Brighton, S.Kapetanakis@brighton.ac.uk

Roger Evans
University of Brighton, r.p.evans@brighton.ac.uk

Gharbi Alshammari
University of Brighton, A.Alshammari1@uni.brighton.ac.uk

Personalized Recommendations on Twitter based on Explicit User Relationship Modelling

Abstract

Information overload is a recent phenomenon caused by a regular use of social media platforms among millions of users. Websites such as Twitter seem to be getting increasingly popular, providing a perfect platform for sharing information which can help in the process of modelling users and recommender system research. This research studies information overload and uses twitter user modelling through making use of explicit relationships amongst various users. This paper presents a novel personal profile mechanism that helps in the provision of more accurate recommendations by filtering overloaded information as it gathered from Twitter data. The presented method takes advantage of user explicit relationships on Twitter based on influence rule in order to gain information which is vital in the building of the personal profile of the user. In order to validate this proposed method's usefulness a simple tweet recommendation service was implemented by using content-based recommender system. This has also been evaluated using an offline evaluation process. Our proposed user profiles are compared against other profiles such as the baseline in order to have the proposed method's effectiveness checked. The experiment is implemented based on an experimental number of users.

Keywords: Recommender systems, User Modelling, User Profiling, Explicit Relationships, Twitter, Influence Score.

Abdullah Alshammari

1.0 Introduction

Real time web seems to be growing as an innovation or technology through which users communicate and send messages via various means such as Twitter. Twitter is a platform being used by millions of people all around the world. Through this social media platform, users are able to exchange and post messages that are short (up to 280 characters) which are referred to as tweets (Vosoughi, 2015). Twitter has been around for a number of years and it has been proven very effective when it comes to sharing casual information as well as breaking news.

Twitter is possible to consider it as a distinctive form of social media websites that present relationships based on a following strategy, something that makes it different from other classic social networking platforms which is based on reciprocal network like Facebook. Relationships that exist between users of Twitter can be informational or social or both. This is due to the fact that users are always following other users for the primary purpose of getting information that takes active part in a network of both interactions and relationships (Abel, Gao, Houben and Tao, 2011; Vosoughi, 2015).

As explained by Abel et al. (2011), there are studies which show Twitter as a vital resource for lots of approaches similar to recommender systems (RSs). RSs have been considered an integral part of many mobile and web applications, having as a goal to ensure the provision of context-aware, real-time and personalized information. This is to help in the increase in sales and user satisfaction. Many studies have used twitter in modelling users and building user profiles in order to have accurate recommendations delivered. This paper is focusing on a Twitter-user model profile via exploiting their relationships in order to improve the performance of recommender systems based on short-text (tweets) profiles within short-term, recent tweets (within last 2 weeks for instance). The following contributions are delivered:

- We propose a recommendation method that builds user profiles from tweets of user's friends with the influence rule redefined by our model
- The proposed method has been experimentally evaluated using a real dataset and well-known metrics with the recommendations delivered being more accurate when compared to alternative methods used as baselines.

The rest of the paper is organized as follows: Section 2 delivers the related work, section 3 presents the proposed method, section 4 explains the experimental evaluation and section 5 contains the conclusions.

2.0 Related Work

A method was proposed by Lee, Oh, Lim and Choi (2014) which was aimed at ensuring that there is improvement in the recommended news articles' accuracy that are from tweets of Twitter. The user profile got built via the extraction of nouns in tweets as well as retweets of users. It was discovered that recommendations based on Twitter are more accurate than the random or normal recommendations. The TRUPI system was proposed by Elmongui et al. (2015). This combined both social features and also the history of tweets from users. It can also capture the level of dynamism that users demonstrate towards various subject matters. This helped to measure how the interest of users changes over the course of time. Temporal dynamics were analysed by Abel et al. (2011) in Twitter profiles. This is majorly for recommendations that are personalized in the social web. Two different forms of profiles were built. These were based on entities and hashtags (for example celebrities and places). Some variables were taken into account such as user's activity, enrichments (making use of external resources like Wikipedia) and time sensitivity. The result revealed that the profile which is entitybased that has been built within a very short period of time; and enrichment outperformed other forms of profiles in a system that is news recommended based on the activities of Twitter. Also, lots of users have the problem of not being able to create a profile that is reliable. This is primarily due to insufficient data about their recent activities. Piao and Breslin (2016) were able to demonstrate that making use of a decay function in the case of long-term profiles which tends to give more weight to current topics of interest as compared to older topics (of interest) indicated much better performance in getting recommendations delivered as compared to long-term profiles when there is no decay function. Apart from that, it has been proven before by Abel, Gao, Houben and Tao (2013) that short-term profiles are better than complete profiles. A major solution has been the enriching of user profile through the use of other data. In the work of Abel et al. (2013), authors were able to model user profiles (in Twitter) through the use of various dimensions. They were able to carry out comparisons on each of them. One of such dimensions happens to be enrichment. Results have demonstrated that making use of external resources like news articles is much better than reliance on Twitter.

Enrichment of user profile through the use of data has been done in various ways. These could be making use of textual external resources (like Wikipedia or articles) or exploiting of URLs in tweets. In order to get URLs of tweets exploited, there was the recommendation of a CatStream system in the work of Garcia Esparza, O'Mahony and Smyth (2013). This makes use of the traditional or normal classification method to profile the users of Twitter. It was based primarily on their tweets' URLs. However, the focus of this system was on the URLs of users' tweets. This made it unsuitable for users who do not have enough number of tweets that has URLs. Alonso et al. (2010) categorized some tweets as being uninteresting and interesting through the use of crowdsourcing. This method was able to demonstrate that a URL link's existence is a feature which can be used accurately in selecting tweets that are interesting. However, it has its own shortcoming that there is a possibility for uninteresting tweets which does not have useful content being categorized incorrectly (Karidi, Stavrakas and Vassiliou, 2016). Authors made use of external resources like articles and Wikipedia in the work of Abel et al. (2011) and Garcia Esparza et al., (2013). User profiles that were enriched through the use of external resources outperformed those profiles there were built solely on the activities of Twitter. Methods such as these are very useful when it comes to supplying more details or information to the user profile. This can help to improve the recommender system's accuracy. However, data which are gotten from external resources will not have any relevance to the interest of users. This may likely affect the recommender system's performance in an adverse way. Furthermore, most users do not usually provide adequate URL links during their tweets.

There is a field which is yet to be investigated. This is exploiting the network of relationships which exists amongst Twitter users with the aim of having a specific user characterized and improving the recommender system's performance. This will be based on activities which are of short-text. It is clear that any user who is into the generation of short-term data (retweets and tweets) can possibly get categorized. This will be based on his behaviour by having historical data collected (that is timeline) which the user has generated by himself. However, for sufficient details to be acquired for the purpose of profiling, a method such as this may need to dig into the past. The gathered information may not be current though. There is also the problem of many users not having adequate data as well as URLs in their most recent activities through which a reliable profile can be created. For this problem of inadequate data to be addressed, we suggest the use of explicit links (for instance, following links) amongst users. This will help in getting relevant recent activities expanded. This method has an advantage which is the fact that there will be more recent data through which profiles can be built from. Through this, the performance of recommender system which is based on short-text messages will improve.

Having following links exploited is going to be achieved via searching for influential users in friends list. Calling a user as an influencer, his actions needs to be effective on users within the same network. Precisely, influencers are users who able to spread information through a network (Morone and Makse, 2015). Majority of the researches were focusing on the popularity of users based on their number of followers (indegree) and friends, and also how they interact with various users (Anger and Kittl, 2011; Bakshy et al., 2011; Chen et al., 2014). TwitterRank was proposed by Weng et al. (2010) to measure the users influence which is an updated version of Page Rank. Other researchers analyzed other measures such as retweets and mentions alongside with indegree to find influential users within a network (Romero et al., 2011). Riquelme and González-Cantergiani (2016) were able to collect and classify various twitter influence measures. However, it is believed that there are variations in the influential rule amongst users in the friends' list. Also, they stated that there is not any agreement with regards to what an influential user should be. Thus, there should be the need for a method, which can have influence score generated from the perspective of users, to be created. This is in relation to the interactions and behaviour of users. Once the influence score has been identified, there is need for incoming tweets to be classified into various categories such as relevant and irrelevant. Techniques have been proposed by researchers for the prediction of tweets which will likely be retweeted. This will be dependent on those features that are content-based (Naveed et al., 2011), coordinate ascent (CA) algorithm (Uysal and Croft, 2011) and collaborative tweet ranking (CTR) (Chen et al., 2012).

3.0 Proposed Method

Generally, the recommender system consists of two stages: user profiling and item ranking. User profiles are going to be built in this work via the extraction of information from those tweets in the timeline of the user and tweets by his/her following list. User profile seems to be what recommendation items are being ranked upon. The whole process will be displayed by Figure 1. Through the use of Twitter API, user's information will be collected. These details or information will be processed in order to identify those keywords that are important which have been posted by friends of the user. The steps which are involved in both stages will be explained in the subsections below.

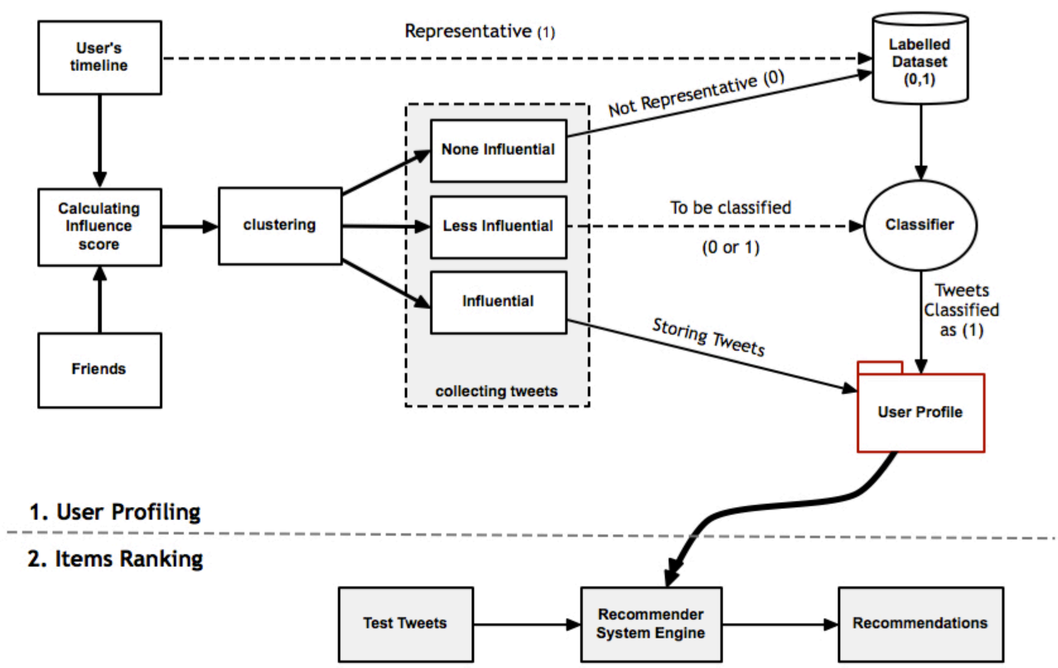

Figure 1. The general steps of the proposed method.

3.1 User Profiling Stage

This stage involves the development of user profiles which contain vital details regarding the user. These are profiles which are built from other users' tweets (friends) who are directly related to the Twitter user. Every profile is being built as keyword profile. Pre-processing steps will have to be implemented prior to the process of recommendation. This will be based on those steps which Micarelli and Sciarrone (2004) suggested. The aim of this is to ensure that tweets are filtered in order for important contents to be extracted. Tweets which the user generates on his own and also retweeted indicate his interest. However, tweets that are received (incoming tweets) from links which are explicit will need to be evaluated and classified. The following steps explain how this stage works:

Step1: Before the profiles are being built, data about the user will be collected from his Twitter timeline. This includes favourited tweets, timeline tweets and friends list.

Step2: This is when influence score can be computed between users and their friends in order to enable us to get such friends ranked based on their level of importance. We will also be able to collect contents that are appropriate. Influence score takes certain variables into consideration such as favouriting, replying, re-tweeting and following. It can be very useful in looking for friends that are important to the user. Influential friends are actually found through the use of the influential score from the perspective of the user. By applying equation 1, the influence score is computed.

$$\text{Influence Score}(\text{u1}, \text{u2}) = \left(\frac{\sum \text{RTs (u2)}}{\sum \text{RTs p(u1)}} + \frac{\sum \text{RTs (u2)}}{\sum \text{Ts p(u2)}} + \frac{\sum \text{MT(u2)}}{\sum \text{MT p(u1)}} + \frac{\sum \text{FV(u2)}}{\sum \text{FV p(u1)}}\right) \times \frac{1}{4} \qquad (1)$$

In equation 1, user1 (u1) represents the original user while user2 (u2) represents the followed friend. ∑RTs (u2) is the total number of tweets posted by user2 and re-tweeted by user1. ∑RTs p(u1) is the total number of re-tweets in the user1 profile. ∑Ts p(u2) represents the number of tweets in the user2 profile. The number of replies (mentions) that user1 posted to user2 represented by ∑MT(u2). ∑MT p(u1) is the total number of mentions in user1's profile. The total number of tweets from user2 that user1 has favourited is represented by ∑FV(u2). The total number of favourited tweets in user1's profile is presented by ∑FV p(u1). Finally, 1/4 is used to normalize the score between 0 and 1.

Step3: Friends list will be grouped into 3 parts based on influence score, which are non-influential friends, less influential friends and influential friends using K-means clustering algorithm. After that, tweets will be collected from all the three groups.

Step4: Every tweet which has its origin from the influential group will be added to the user profile while those tweets which have their origins from the non-influential group will not be added. Those tweets which have their origins from less influential friends will be classified as re-tweetable (representative) or not re-tweetable (not representative). This process made use of various classifiers: Neural Networks, K-Nearest Neighbor, Decision Tree, Support Vector Machine, Random Forest, and Naïve Bayes. Every classifier is trained through the use of a labelled data set which is from the timeline of the user and also non-influential users' tweets. Those tweets which have their origins from the user timeline will be tagged as re-tweetable (representatives) while those from non-influential users are going to be tagged as not re-tweetable (not representative).

Step5: The dataset is divided into two parts, which are testing and training sets. Testing sets is used in the computation of every classifier's accuracy. Any classifier that is the highest in terms of accuracy is going to be automatically chosen in order to get the tweets of less influential users classified. This step is to ensure that tweets are classified through the classifier which is the most accurate. After getting the tweets of less influential users classified, those tweets which have been classified as re-tweetable (representative) will be stored in the profile of the user alongside influencers' tweets.

3.2 Items Ranking Stage

During this stage, recommendation items are a group of tweets which the user is going to indicate interest in by the retweet action. Vector space model representation is used in this process and user profile as well as recommendation items will be considered as vectors. The angles that are between them will be computed. Once an item is very close to the user profile, such is a sign that it is relevant. The closer it is, the more relevant it will become. In order for the angles to be calculated, the cosine similarity is defined in equation 1 and applied. In cosine, the distance between to point A and B is calculated and the values range from -1, total dissimilarity and 1 total similarity.

$$similarity = \cos(\theta) = \frac{A \cdot B}{||A||||B||} = \frac{\sum_{i=1}^{n} A, B,}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}} \tag{2}$$

The user profiles have been built as explained in the last subsection. It will function as the basis through which the group of tweets inside the recommendation items are ranked. All the tweets inside the recommendation items are going to be evaluated or ranked based on how they are similar to the user profile. Before the cosine similarity applied, pre-processing steps are applied that are suggested by Micarelli and Sciarrone (2004). Furthermore, tweets are going to be excluded when the text that is remaining is less than 3 words. Lastly, there will be computation of the similarity score between user profile and tweet profile. Cosine similarity equation is what this calculation will be based on. Therefore, every recommendation item is ranked. Also, top-k tweets will be recommended to such user.

4.0 Experimental evaluation

In order for the advantage of this proposed method to be validated, there was the implementation of tweets recommender system. An evaluation was also performed offline using some set of users. Through the use of Twitter API which can be found inside the development section of the official website of Twitter, some randomly chosen users' (29) timelines were collected as well examined. In the case of recommendation items, those test tweets which users had indicated some levels of interest before through re-tweets are collected from timelines of users. The following subsection will be explaining the methods through which these tweets are collected.

4.1 Experiment Setup

Once the examined user's timeline has been collected and prior to getting the influence score computed, as well as getting users clustered into the 3 main groups, the dataset has been divided as shown in Figure 2. It has been broken down into 3-time frames.

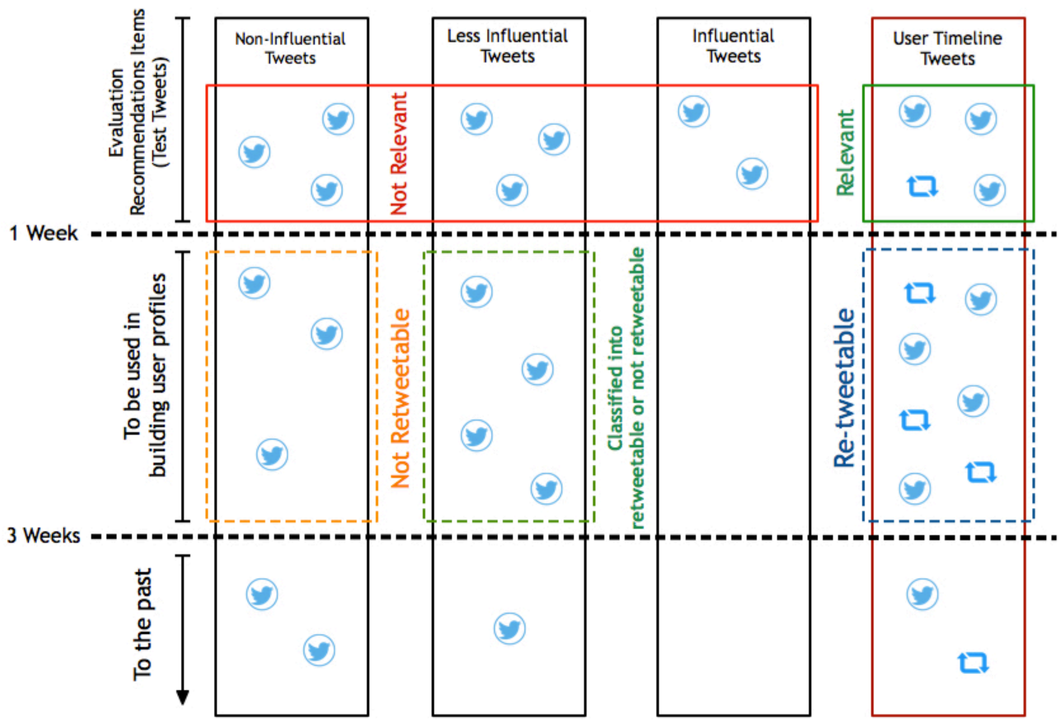

Figure 2. **Dividing the user timeline into three evaluation time frames.**

The first-time frame's tweets are made use of as test item. It is clear that they were retweeted simply because the user was interested in them. This is similar to traveling into the past in order to predict the future which is known already. This can help in the process of evaluation. The second time frame's tweets (between 1 week and 3 weeks ago) will be used in getting user profile built from various sources. The third time frame is going to be used alongside the second time frame in order for the influence score to be computed from the examined user's timeline. Furthermore, the timeline of the user will be used during machine learning classification.

Profiles: For each user of the 29 users, some profiles are made and then compared against each other and against the baseline. This is to ensure which profile performance is the best in order to compare it later against other metrics such as similarity, distance and indegree. The profiles are explained below:

1. **Baseline:** This includes all the tweets on the timeline of the user. It contains tweets that have been posted and also those which are re-tweeted.
2. **BLCinf:** This includes every tweet from the timeline of the user, short term tweets of friends that are influential and also those ones from less influential friends which have been classified as representative.

3. **STBLCinf:** This includes only those second time period's tweets (short term) and also those from the timeline of the user. It also includes every tweet by influential and less influential friends which are classified as representative.
4. **STBLinf:** These are only second period's tweets. They include tweets from the timeline of the user, by influential friends and also less influential friends. Classification is not given any consideration.
5. **BLinf:** This is every tweet from the timeline of the user, short term tweets made by less influential and influential friends. This profile does not consider less influential friends' tweets classification.

Test tweets: Test tweets are being used in evaluating the recommender system's accuracy. These are a group of tweets that are gotten from the first-time frame (week 1). The recommender system makes use of them as recommendation items. It is a collection which contains tweets that are relevant and irrelevant. From the user's timeline, those items that are considered to be relevant which he has retweeted will be known. Irrelevant items are tweets collected from friends and the user has not shown any action to them. Therefore, recommendation items are a combination of both items that are relevant and irrelevant. This will help to ensure that the accuracy is measured by the recommender system when it runs different user profiles. It will also enable the built profiles and the baseline profiles to be compared.

Evaluation metrics: This study made use of offline evaluation in measuring the recommender system's accuracy through the use of various user profiles (Uysal and Croft, 2011). Various user profiles were used in the recommender system and then they are compared. This research made use of the metrics in measuring the accuracy of the system's performance, which are average of precision @ k (P@k), average precision (AP) and mean average precision (MAP). P@k is the amount of correct recommendations in the top-n list of recommendations and is defined in equation 3. AP is defined in equation 4 where p(k) is the precision @ k and rel (k) is an indicator counting as one if the item is relevant or zero otherwise. Moreover, relevant not retrieved items receive a score of zero. Lastly, MAP which is defined in equation 5 with Q being a query and the equation returning the mean of the average precision scores for a set of queries.

$$P@k = \frac{relevant\ recommended\ items}{total\ recommender\ items} \tag{3}$$

$$AP = \frac{\sum_{k=1}^{n}(p(k)x\ rel(k))}{number\ of\ relevant\ recommendations} \tag{4}$$

$$MAP = \frac{\sum_{q=1}^{Q} AP(q)}{Q} \tag{5}$$

Through the AP, the system will be measured in terms of how good it is at retrieving top-k relevant items. The MAP measures the effectiveness of the system in getting all relevant items retrieved.

4.2 Results

This subsection explains the results obtained from the evaluation metrics. In the metric of the Average of Precision @top-k recommendations, the tested values of k are: 1, 3, 5, 10, 15 and 20. (See Figure 3).

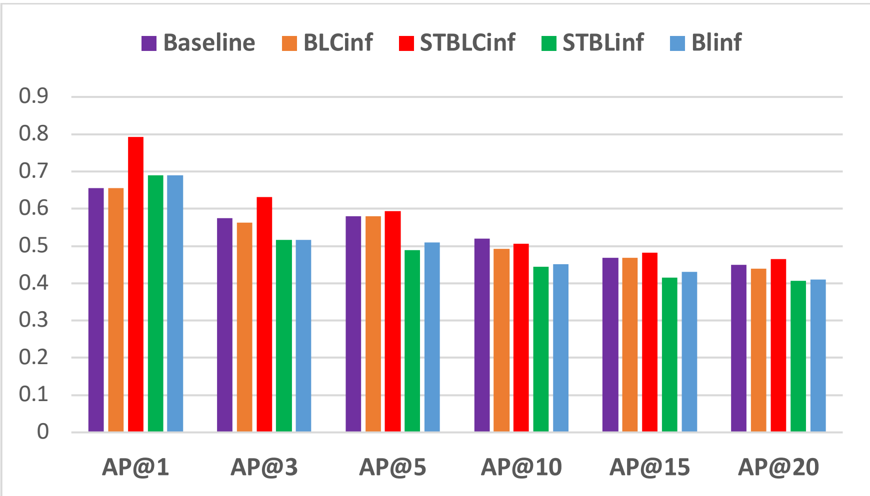

Figure 3. The average precision @1, 3, 5, 10, 15 and 20 of profiles.

When the top-k is set to 1, 3 and 5, results showed that the profile STBLCinf scored the highest average precision among all other strategies of building user profiles. It contains the short-term tweets of: user timeline posts, influential friend tweets and less influential friend tweets that are classified as relative. This strategy shows how powerful the profile is to give relative recommendations on top of the recommendation list to users. Otherwise, the baseline and BLCinf profiles scored the lowest average precision on top-1. The baseline outperformed all other profiles when the top-k is set to 10. The STBLCinf profile came in the second place whereas STBLinf achieved the lowest Average Precision. Continuously, Profile STBLCinf outperformed all other profiles in top-k = 15 and 20. However, again profile STBLinf achieved the lowest average precision and that might give a clear view that non-relevant tweets affected the accuracy of the recommender system. STBLCinf and STBLinf were built similarly but the only difference is that the latter included all tweets from less influential users without any consideration of classification. As a result, this small difference can affect the performance and make it achieve the lowest reliable profile.

In the Mean Average Precision (MAP), Figure 4 shows the results that the profile STBLCinf achieved the highest mean average precision against all profiles. Also, profile BLCinf achieved better performance than the baseline and that might mean enriching the baseline profiles with some related data can improve the performance of the recommender system. Also, profile BLinf was built similarly to BLCinf achieved less MAP than the latter and the baseline profile. This may clarify that enhancing profile with none related data (unclassified tweets in this case) can reduce the quality of delivering recommendations even worse than the baseline. Also building profiles based only on timeline (baseline) cannot deliver more relevant recommendations. On the other hand, profile STBLinf achieved the lowest MAP.

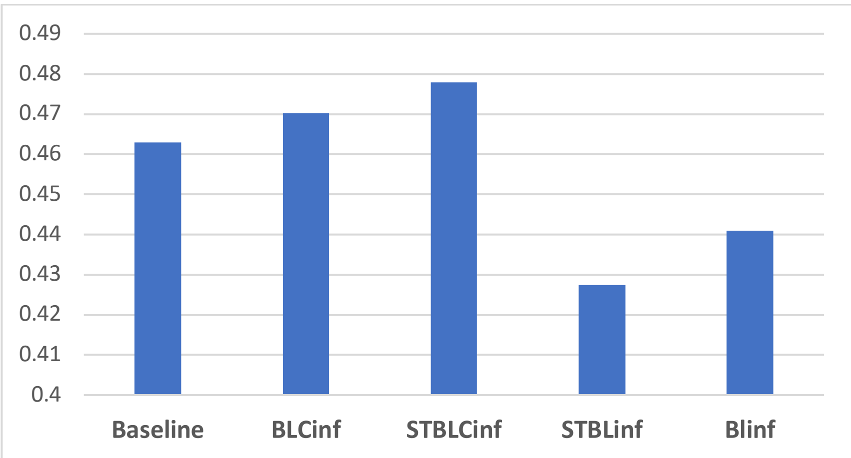

Figure 4. The Mean Average Precision (MAP) of profiles.

To validate how strong our proposed influence score that the profile STBLCinf built based on, we compare the profile against other 5 profiles, which are based on similarity, distance and followers count. The similarity metrics are: Cosine and Jaccard. Euclidean and Manhattan distances are also used. Finally, followers count metric is used to build a profile and the reason is that there are literature researches that have used this during experiments as a sign of influence (Cha et al., 2010; Bakshy et al., 2011; Razis and Anagnostopoulos, 2014; Riquelme and González-Cantergiani, 2016). We built the mentioned profiles in the same way of the profile STBLCinf. Additionally, the clustering was applied based on the mentioned metrics instead of the influence score. In the results, the profile STBLCinf outperformed all other profiles in Average precision @3, 5, 15 and @20 as in Figure 5. Whereas, it achieved the same average precision @1 with Jaccard profile. Euclidean profile outperformed all other profiles when the average precision is set to 10. Also, it came in the second place in AP@3, 5, 15 and 20.

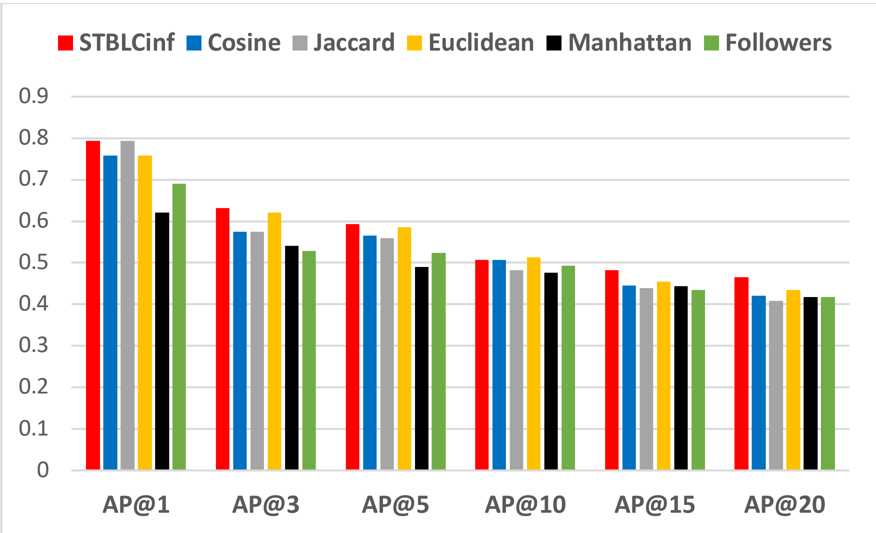

Figure 5. The average precision (AP) @1, 3, 5, 10, 15 and 20 of profiles.

Results in figure 6 also showed that in the Mean Average Precision, the profiles STBLCinf and Euclidean achieved the highest MAP among all other profiles. Surprisingly, the followers count profile achieved the lowest MAP and this might clarify that the number of followers is not a reliable metric to measure how influential a user is. Also, this prove our hypothesis about the influence rule has to be based on the user preferences and not on the influencer himself.

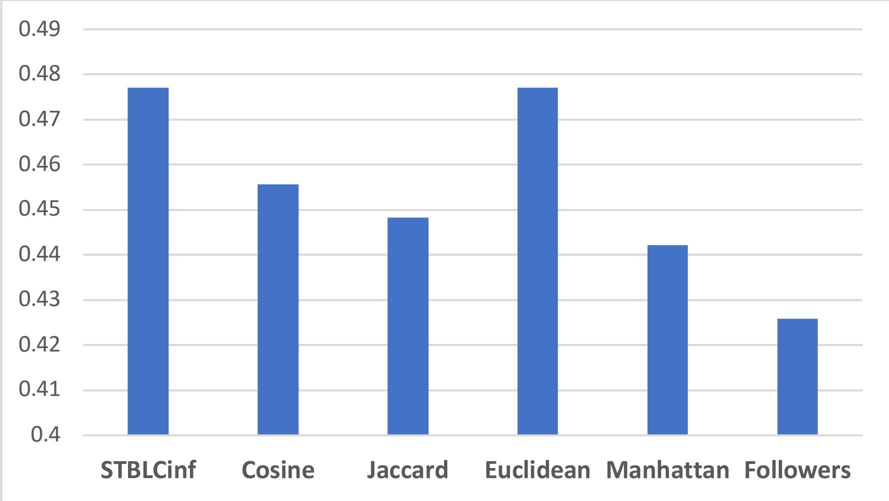

Figure 6. The Mean Average Precision (MAP) of different profiles.

5.0 Conclusion

In conclusion, a new method through which user's profiles can be built via exploiting Twitter explicit network structure. This is to ensure that the short-text-based recommender systems' performance is improved by a better way of modelling user. The new user profile takes advantage of following links between users as well as their friends in order to gather short-term tweets. Through these tweets, profiles were built. The Twitter's influence rule has also been redefined. This has helped us in having the following list clustered into 3 groups: non-influential, less influential and influential. Due to this, the tweets of influential users are stored to user profile. On the other hand, tweets of non-influential users are excluded. In order for the representative tweets to be stored into the use's profiles, less influential users' tweets have been classified by classifiers. This method's advantage has been validated through an evaluation which was carried out offline. A prototype tweets-recommender system is what this was based on. Our method's discriminative power is presented through testing as well as making comparison of our method against baseline and followers count profiles. Also, the proposed profile was compared against other similarity metrics such as Cosine, Jaccard, Euclidean and Manhattan. Various forms of relationships will be explored in future between users and their friends. Another thing that will need to be considered is the similarity existing between users and their followers. This will enable us to expand those group of tweets which indicate the interest of the user.

References

Atkinson, C. J. and Brooks, L. (2005) *In the Age of the Humanchine*, In Proceedings of International Conference on Information Systems (ICIS), AIS, Las Vegas, USA.
Abel, F., Gao, Q., Houben, G. J., & Tao, K. (2011, June). Analyzing temporal dynamics in twitter profiles for personalized recommendations in the social web. In Proceedings of the 3rd International Web Science Conference (p. 2). ACM.
Abel, F., Gao, Q., Houben, G. J., & Tao, K. (2013, August). Twitter-Based User Modeling for News Recommendations. In IJCAI (Vol. 13, pp. 2962-2966).
Alonso, O., Carson, C., Gerster, D., Ji, X., & Nabar, S. U. (2010, July). Detecting uninteresting content in text streams. In SIGIR Crowdsourcing for Search Evaluation Workshop.
Anger, I., & Kittl, C. (2011, September). Measuring influence on Twitter. In Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies (p. 31). ACM.
Bakshy, E., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011, February). Everyone's an influencer: quantifying influence on twitter. In Proceedings of the fourth ACM international conference on Web search and data mining (pp. 65-74). ACM.
Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, P. K. (2010). Measuring user influence in twitter: The million follower fallacy. Icwsm, 10(10-17), 30.
Chen, K., Chen, T., Zheng, G., Jin, O., Yao, E., & Yu, Y. (2012, August). Collaborative personalized tweet recommendation. In Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval (pp. 661-670). ACM.
Chen, C., Gao, D., Li, W., & Hou, Y. (2014, July). Inferring topic-dependent influence roles of Twitter users. In Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval (pp. 1203-1206). ACM.
Elmongui, H. G., Mansour, R., Morsy, H., Khater, S., El-Sharkasy, A., & Ibrahim, R. (2015, April). TRUPI: Twitter recommendation based on users' personal interests. In International Conference on Intelligent Text Processing and Computational Linguistics (pp. 272-284). Springer, Cham.
Garcia Esparza, S., O'Mahony, M. P., & Smyth, B. (2013, March). Catstream: categorising tweets for user profiling and stream filtering. In Proceedings of the 2013 international conference on Intelligent user interfaces (pp. 25-36). ACM.
Karidi, D. P., Stavrakas, Y., & Vassiliou, Y. (2016, July). A personalized Tweet recommendation approach based on concept graphs. In Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld), 2016 Intl IEEE Conferences (pp. 253-260). IEEE.
Lee, W. J., Oh, K. J., Lim, C. G., & Choi, H. J. (2014). User profile extraction from Twitter for personalized news recommendation. Proc. 16th Advanced Communication Technology, 779-783.
Micarelli, A., & Sciarrone, F. (2004). Anatomy and empirical evaluation of an adaptive web-based information filtering system. User Modeling and User-Adapted Interaction, 14(2-3), 159-200.
Morone, F., & Makse, H. A. (2015). Influence maximization in complex networks through optimal percolation. *Nature*, *524*(7563), 65.
Naveed, N., Gottron, T., Kunegis, J., & Alhadi, A. C. (2011, June). Bad news travel fast: A content-based analysis of interestingness on twitter. In Proceedings of the 3rd international web science conference (p. 8). ACM.

Piao, G., & Breslin, J. G. (2016, September). Exploring dynamics and semantics of user interests for user modeling on Twitter for link recommendations. In Proceedings of the 12th International Conference on Semantic Systems (pp. 81-88). ACM.

Razis, G., & Anagnostopoulos, I. (2014, September). InfluenceTracker: Rating the impact of a Twitter account. In IFIP International Conference on Artificial Intelligence Applications and Innovations (pp. 184-195). Springer, Berlin, Heidelberg.

Ricci, F., Rokach, L., & Shapira, B. (2015). Recommender systems: introduction and challenges. In Recommender systems handbook. Springer, Boston, MA.

Riquelme, F., & González-Cantergiani, P. (2016). Measuring user influence on Twitter: A survey. Information Processing & Management, 52(5), 949-975.

Uysal, I., & Croft, W. B. (2011, October). User oriented tweet ranking: a filtering approach to microblogs. In Proceedings of the 20th ACM international conference on Information and knowledge management (pp. 2261-2264). ACM.

Vosoughi, S. (2015). Automatic detection and verification of rumors on Twitter (Doctoral dissertation, Massachusetts Institute of Technology).

Weng, J., Lim, E. P., Jiang, J., & He, Q. (2010, February). Twitterrank: finding topic-sensitive influential twitterers. In Proceedings of the third ACM international conference on Web search and data mining (pp. 261-270). ACM.