A framework for securing Recommender Systems through Correlation techniques

[1]Manish Jaiswal, [2]Tanveer J. Siddiqui

[1]Department of Electronics & Communication
[1]University of Allahabad, Allahabad, India
Email: manish.jk111@gmail.com[1]

Abstract: Recommender system is information filtering tools for handling the problem of information overwhelming. On the other side, security and protection is the prime concern for recommender system in order to provide genuine recommendation to the end users. Many e-commerce companies face attacks from malicious users who produce wrong recommendations to their buyers. Here, we presented the recommender system approaches, basic principles and overview of the different category of attacks discussed. In this work, we proposed a framework for genuine recommendation to users by detection of fake profile using correlation techniques based on correlation. Applicability of the proposed work in others domain with new techniques has been discussed in conclusion and future scope.

Keywords: Recommender system, attack model, collaborative filtering, clustering

[1] INTRODUCTION

In the era of digital revolution and increasing use of internet, problem of information overloading, this offers a huge volume of information in terms of services and products to users. In the same time, it's very difficult for users to decide which service or item is to be avail as per their choice. In such scenario, Recommender system(RS) have played a vital role in the age of information overload where there is huge volume of data creates confusion in user's mind in order to decision making for purchasing any products or services provided by any companies or service provider. RS proved its applicability in numerous application areas such as e-commerce, education, social media, entertainment industry, tourism, medical and so on. There are various e-commerce and entertainment industries are using RS for increasing their revenue such as Amazon, Jester, Movielens, Netflix, Yahoo Movies, OYO, spotify etc.

RECOMMENDER SYSTEM APPROACHES

There are mainly three filtering approaches of recommender system- Collaborative filtering (CF), Content based recommendation (CBR) and Hybrid filtering [2, 3]. Collaborative filtering (CF) approach of recommender systems based on basically similarity of user's taste in terms of their likes and dislikes. Suppose user A and user B are friends and they have similar taste and they both purchased product Y. Now User A buys any products X then the same product X will also be recommended to the user B due to assumption of their similar interest. CF is the most popular and widely used filtering approach of recommendation [1]. The recommendation in CF for a given user is based on behavior and evolution of other users. Collaborative filtering is divided in two classes namely memory-based and model-based collaborative filtering. Memory-based collaborative class is again divided into User-based (User-to-User) collaborative filtering and Item-based (Item-to-Item) collaborative filtering. Content based approach recommends items based on similarity comparison between the concept of item's and user's profile. The features of items are mapped with features of users in order to obtain user-item similarity. The top matched pairs are given as recommendations. The concept of collaborative filtering and content based recommendation can be view in fig. 1.

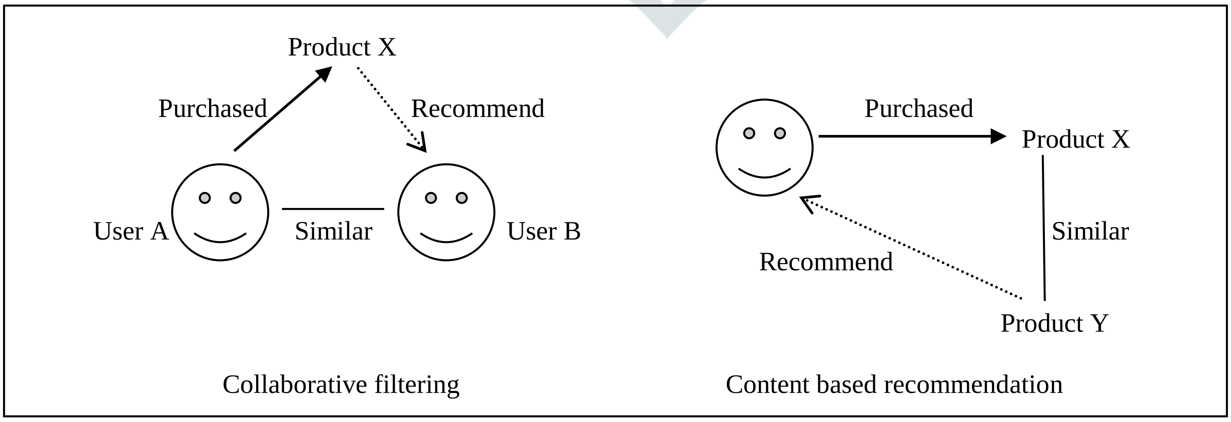

Figure 1.Recommender System Approaches

Both CF and CBR have their positive and negative features. In case of collaborative filtering, basic domain knowledge is sufficient. It is easy to implement and very simple but the main issues with CF are Sparsity, cold start problem for new users and new item. Scalability is another big issue with CF. Content based recommendation doesn't need any data about other users which makes it easier to scale corresponding to increasing number of users. The main disadvantage of CBR is limited content analysis which means that this model can only recommends items based on existing user's interests. A hybrid recommendation system is more enhanced category of RS which provide more accurate and better recommendation to users. Classification of recommender systems are depicted in fig. 2

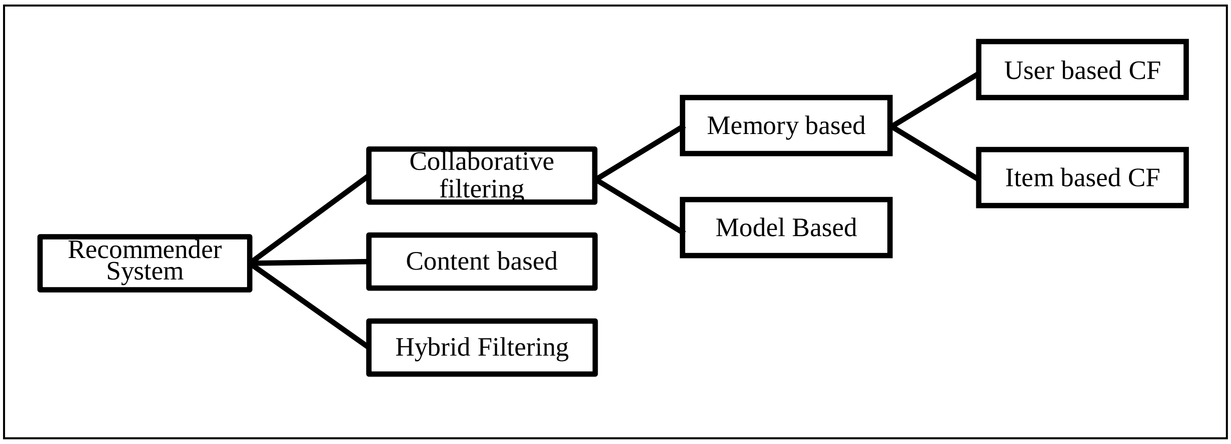

Figure 2.Classification of Recommender System

Hybridization can be done by various methods such as cascading based, meta-level, weighted hybrid, and based on feature combination. Spotify is very popular music and video services provider application that use the concept of hybrid recommender systems and provides more personalized recommendations to users from their database.

[2] RELATED WORK

This section presents state of art and various literatures related to attack detection and security of recommender system by various researchers. [5-7,10,12]. Some of previous research work is discussed here. [8] proposed a system for fake profile detection using concept of discovering the abnormal pattern online social media networking. Classifier detects the malicious account by analyzing either graph-level structures or user-level. Precision metric used to check the efficiency and accuracy of the system. Author used dataset in his work was collected from kaggle.

In [9], author demonstrates an identity detection mechanism on social networking websites. The approach was based on user behavior detection and physical characteristics of users. Accuracy of the system was measured by identifying false profile accurately. [11] demonstrated the system for clone attack detection on social media dataset. Clone attack referred the process of sending the malicious data several times by attackers at the receiver ends due to which there were problem of congestion in network. The proposed system was based was mining technique and very useful in handling and detection of spammers. Classification accuracy were used to test the efficiency of the proposed system

[13] implemented fake profile detection mechanism based on clustering technique. Main drawback of the proposed system was in handling missing values in the dataset which leads low accuracy of the system in the implemented work. In [14], researcher detects the identify of malicious patterns by using pattern recognition techniques. One shortcoming of this work was that authors did not pre-processed dataset which affected the performance of the system in terms of accuracy. Twitter dataset was used for the implementation of proposed work. [15] presented graph-based approach for false profile detection. By using path traversal techniques in graph, next node was founded. If any path was traversed many times then it was detected as false profile. Classification accuracy was used for system testing in the work.

[3] ATTACK MODELS ON RECOMMENDER SYSTEMS

Collaborative filtering of recommendation is based on item rating and user profile. Due to this reason the recommender system is highly vulnerable to attacks. The attackers may provide more recommendations to promote their favorites product and fewer recommendations to competitor product. Initially, the main goal of attacker to either promote or demote the product by increasing or decreasing the ratings of targeted items that will receive either possible highest possible rating (increase the predicted value for target item) called "push attack" or the lowest possible rating (decrease the prediction value of target item) called "nuke attack". To breach the security, the attacker inserts multiple fake profiles in the system to creating attack.. These attacking profiles will include a rating for target item and some number of other ratings. The rating database is not publically available; more effective attacks can be designed by estimating distribution of values. In collaborative filtering system, attackers profile contains the items which could be set of selected items (filler items), unselected, and target items. Some common types of attacks [4] on recommender system are Random attack, Average attack, Bandwagon attack, Segment attack. In random attack the attack profiles randomly choose the target item to rate their rating except the target item. The injected profiles are filled with the random values. In the average attack, filler items are rated based on the average rating for each item across ratings for all users. Bandwagon attack is also termed as popular attack because an attacker can know the popularity of an item independent of recommender system. In such attack the attacker associates the attacked item with a handful of well known popular items. In segment attack, the attacker focuses basically on a particular set of users and increases the recommendations of target item for this set of users.

Principle of Attack

The recommender system provides recommendation based on user input rating. The recommender contains a known cognizance on which the receiver can keep faith. The attacker can enter multiple profiles and fake identities into the recommendation engine. The user-based k-nearest-neighbor algorithm is the most common algorithm for the user profiles which uses user input rating profiles. There are set of items for each user based on the user likes for that item. The system forms a neighborhood of peer users with similar tastes. It then extrapolates the user's predicted rating for a target item from the ratings of his or her peer users. The recommender systems are subjected to manipulation because of being dependent on user's profile. For example, in the given table 1, consider a hypothetical movie recommender system, there are 7 users (U1, U2, U3, U4, U5, U6, U7) and 6 movies for which rating are provided by users. There is no rating given by user for some movie. The rating is in range of scale 1-5. Where 1 showing the lowest and 5 showing the highest ratings. User John is a genuine user, has to get the new recommendation and rating for movie6. The given table 1 shows the profile of John with the other users and attacker's profile. Miranda is another user act as attackers who inserted three profiles in the system and provided the higher rating for movies 6. Without the attacker's profiles, user U6 is the most similar user to John, having highest correlation value. The predicted rating value associated with user U6 is 2. Which states the John will dislike the movie 6. After the insertion of attack, the most similar profile to John is attack profile 1, which is generating a predicted rating value 5 to the movie 6 and now the system will recommend movie 6 to John. The attacker can also use this technique to decrease the rating of a target item also.

Table 1- Example favoring the target movie 6

	Movie 1	Movie 2	Movie 3	Movie 4	Movie 5	Movie 6	Correlation with John
John	5	2	3	3		?	
User U1	2		4		4	1	–1.00
User U2	3	1	3		1	2	0.76
User U3	4	2	3	1		1	0.72
User U4	3	3	2	1	3	1	0.21
User U5		3		1	2		–1.00
User U6	4	3		3	3	2	0.94
User U7		5		1	5	1	–1.00
Attack profile 1	**5**		**3**		**2**	**5**	**1.00**
Attack profile 2	**5**	**1**	**4**		**2**	**5**	**0.89**
Attack profile 3	**5**	**2**	**2**	**2**		**5**	**0.93**
Correlation with Movie 6	0.85	–0.55	0.00	0.48	–0.59		

[4] Proposed METHODOLOGY

The main objective of proposed work is to provide security to the system by detection of attacks, which enhance the accuracy of the collaborative recommender systems. For our experiment we have chosen IMDB data set for the movie ratings. The data set contain 72 instances of movie ratings from users for 9 attributes. We have used WEKA tools for implementation of proposed work. We divided the Movie Rating data set into training and testing data. Then we inserted fake profiles into data set and analyzed using WEKA. It was clearly seen that the fake profile data was deviating far from the actual rating of the movie.

Clustering used for detection of attack

Generally, the attribute of fake user is very different from that of the genuine user. The attribute of the fake user lies much above that of genuine user since the deviation and the variance of that of fake user is very different from that of real user. We have used the combination of some of the mathematical formulae to calculate the deviation of the each user rating from the aggregate rating. Similarly we have computed the deviation of the fake user from that of the real user and it is found its deviation from that of clustered data of the genuine user. These three criteria are being use to compute for every user of its rating and their data are being fed into the database for further filtering criteria of finding the fake user from the real user.

The proposed framework contains mainly three parts which are user agent (UA), database and recommender agent (RA). The user inserted rating is received by UA, the database contains the record of ratings provided by users and finally RA provides recommendation based on rating values using recommendation algorithm. Before going to RA for recommendation process, the user inserted rating value will face the calculation. The calculation will be done on the basis of the stored ratings and the inserted rating value. The proposed framework includes three steps for detection of fake profiles for providing genuine recommendations.

1. Deviation from aggregate rating
2. Deviation of ratings from standard deviation.
3. Computation of covariance from that of pre-existed ratings.

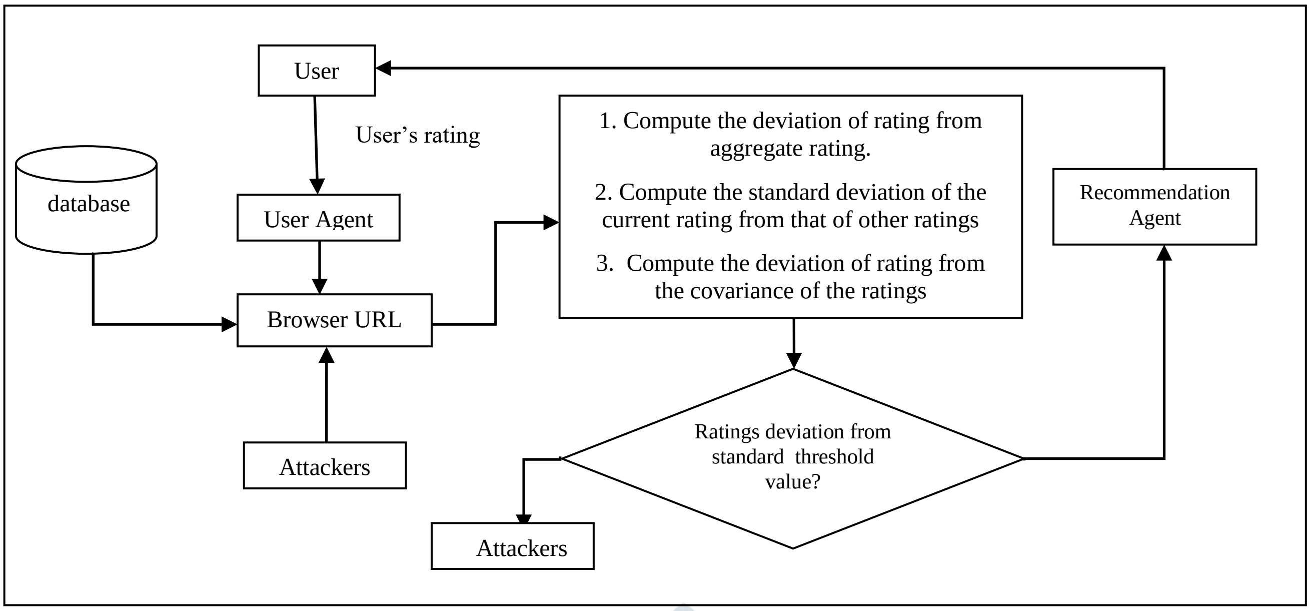
User
User's rating
database
User Agent
Browser URL
Attackers
1. Compute the deviation of rating from aggregate rating.
2. Compute the standard deviation of the current rating from that of other ratings
3. Compute the deviation of rating from the covariance of the ratings
Recommendation Agent
Ratings deviation from standard threshold value?
Attackers

Figure 3.Flowchart of proposed framework

Deviation of the rating from the aggregate rating: In this step, calculation of deviation of the user's input rating from that of aggregate rating of the that particular item (in this case it is movie rating). This deviation should not be more than a particular set of defined standard threshold which depends on the user's history of reviewing the movies.

$$R(user) = \frac{\sum_{i=0}^{u} Uu,i - Ai}{\frac{NUi}{Nu}}$$

Deviation of the rating from standard deviation: In this step, we calculated the deviation of the fake user's rating from that of standard deviation of the rating of that particular item (in this case it is movie's rating).

$$D(user) = \frac{\sum_{i=0}^{u} Du,i - SDi}{\frac{NDi}{Nu}}$$

Covariance computaion: In this final step, covariance is calculated of that particular rating of the input provided by the user. Now we'll check the deviation of the covariance with respect to the rating variance of the existing user that have given the overall rating This deviation should not be more or less than a particular defined standard which depends on the users history of reviewing the movies.

$$Cov(user) = \frac{\sum_{i=0}^{u} COu,i - COi}{\frac{NCOi}{Nu}}$$

[5] RESULTS AND DISCUSSION

In our experiment done with IMDB dataset, the inserted attackers profile is detected as fake profile if computed covariance is deviated from standard threshold rating otherwise rating of user will get inserted in RA for further recommendation to new user.

[6] CONCLUSION AND FUTURE SCOPE

Security issue is the major concern of the recommender system. In this work, we have analyzed different types of attacks relevant to recommendation and we proposed a novel approach and framework for detection of attackers profile by using correlation technique in machine learning. By capturing and removing the fake profile, efficiency and performance of proposed systems enhanced in terms of accurate recommendation to users without any bias. The proposed techniques can be extended by using various machines learning classifier for attack detection and to provide genuine recommendation for users. Natural learning processing and sentiment analysis can also be explored for existing work to achieve more secure and robustness.

[7] ACKNOWLEDGMENT

We are thankful to the Department of Electronics and Communication, University of Allahabad, for providing necessary facilities for the execution of present work.

REFERENCES

[1] R. Burke. 2002. "Hybrid recommender systems: Survey and experiments."User Modeling and User-Adapted Interaction, 12(4), pp. 331–370.

[2] G. Adomavicius and A. Zuzhilin. 2005. “Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions”, IEEE Transaction on knowledge and Data Engineering, vol. 17(6), pp. 734-749.

[3] A. Salah, N. Rogovschi, M, Nadif. 2015. “A dynamic collaborative filltering system via a weighted clustering approach”. Neurocomputing 175(206), doi:10.1016/j.neucom.2015.10.050.

[4] P. Chirita, W. Nejdl, and C. Zamfir. 2005. ”Preventing shilling attacks in online recommender systems”. WIDM: Proceedings of the seventh annual ACM international workshop on Web information and data management. ACM Press, New York, pp. 67–74.

[5] B. Mehta, W. Nejdl 2008. “Attack resistant collaborative filtering”.Proceedings of International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 75-82.

[6] R. Burke, B. Mobasher, C. Williams, R. Bhaumik. 2006. “Detecting profile injection attacks in collaborative recommender systems”. International Conference on Enterprise Computing, E-Commerce, and E-Services.IEEE, pp. 23-23.

[7] R. Burke, B. Mobasher, C. Williams, R. Bhaumik. 2005. “Segment-based injection attacks against collaborative filtering recommender systems”. IEEE International Conference on Data Mining.

[8] Y. Boshmaf et al. 2016. “Íntegro: Leveraging victim prediction for robust fake account detection in large scale OSNs,” Computer Security, vol. 61, pp. 142–168.

[9] M. Tsikerdekis and S. Zeadally. 2014. “Multiple Account Identity Deception Detection in Social Media Using Nonverbal Behavior” IEEE Transaction Information Forensics Security., vol. 9(8), pp. 1311–1321.

[10] Z. Cheng, N. Hurley. 2009. “Effective diverse and obfuscated attacks on model-based recommender systems. Proceedings of 3rd ACM Conference on Recommender systems, pp. 141-148.

[11] L. Jin, H. Takabi, and J. B. D. Joshi. 2011. “Towards active detection of identity clone attacks on online social networks,” Proceedings of the first ACM Conference on Data and Application Security and Privacy, pp. 27–38.

[12] W. Carrer-Neto, ML. Hernández-Alcaraz, R. Valencia-Garcia, F. Garcia-Sanchez. 2012. “Social knowledge-based recommender system”. Expert Systems with Applications, doi: 10.1016/j.eswa.2012.03.025, pp. 10990-11000.

[13] C. Xiao, D. M. Freeman and T. Hwa. 2015. “Detecting Clusters of Fake Accounts in Online Social Networks Categories and Subject Descriptors,” ACM, doi: http://dx.doi.org/10.1145/2808769.2808779.

[14] S. Gurajala, J. S. White, B. Hudson, and J. N. Matthews. 2015. “Fake twitter accounts: Profile characteristics obtained using an activity-based pattern detection approach,” Proceeding: ACM International Conference, vol. 2015, doi: 10.1145/2789187.2789206.

[15] M. Mohammadrezaei, M. E. Shiri, and A. M. Rahmani. 2018. “Identifying fake accounts on social networks based on graph analysis and classification algorithms,” Security Communication. Networks, doi: 10.1155/2018/5923156.