ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 6, June 2017

A Survey Paper on "Detecting Suspected Users by Utilizing Specific Distance Metric in Collaborative Filtering Recommender Systems"

Priyanka Bhagwan Patil[1],Ganesh Dhanokar[2]

Student, Department of Computer Engineering, GHRIEM. Jalgaon, India, G H Raisoni Group of Institutions, Jalgaon North Maharashtra University, India[1]

Assistant Professor, Department of Computer Engineering, GHRIEM. Jalgaon, India,G H Raisoni Group of Institutions, Jalgaon North Maharashtra University, India [2]

ABSTRACT: Recommender system is an imperative part of the data and internet business biological system.Collaborative filtering (CF) is a vital and well known innovation for recommender system.Collaborative filtering recommender systems (CFRSs) are basic parts of existing well known online business sites to make personalized recommendations.Be that as it may, current CF strategies experience the ill effects of such issues as "shilling" attacks or "profile injection" attacks because of its openness.While an extensive variety of recognition systems have been utilized, some of them depended on calculating similarity between users (consists of attackers and genuine users) keeping in mind the end goal to segregate those attackers.In practice, it is hard to catch every concerned attacker by abusing the similarity of users, in spite of the fact that it can be useful to filtering out more genuine users.Calculating directly the similarity between users on a whole dataset consumed high computation time although they can be powerful to catch the concerned attackers in some degree. In this paper we propose an unsupervised method to detect such attacks. At the first stage, we filter out a part of genuine users in order to reduce the enumeration time. At the second stage, we mainly focus on the effective similarity metric to better differ between attackers and genuine users based on the remaining users of the first stage and modulates the traditional similarity metric and the linkage information between users to improve the accuracy of similarity of users.

KEYWORDS: Recommender system, Shilling attack, Attack detection, Collaborative Filtering(CF), Similarity metric.

I. INTRODUCTION

Personalization recommender systems (RSs) become more and more in demand in e-commerce websites to automatically make personalized suggestions of services or products to clients.Collaborative filtering (CF) is an essential and well known technology for recommender systems. There has been a lot of work done both inindustry and the scholarly community. These methods are classified into user-based CF and item-based CF. The basic idea of user-based CF method is to search out a set of users who have similar favor patterns to a given user (i.e., "neighbors" of the user) and recommend to the user those items that other users in the same set like, while the item-based CF method aims to confer a user with the recommendation on an item based on the other items with high correlations (i.e., "neighbors" of the item). In all collaborative filtering methods, it is a valuable step to find users' (or items') neighbors, that is, a set of similar users (or items). However, Collaborative filtering recommender systems (CFRSs) are prone to manipulation from attackers due to its openness, which carefully infuse chosen attack profiles into CFRSs to prejudice the recommendation results to their benefits or decrease the trustworthiness of recommendation. These events are often called "shilling" attacks or "profile injection" attacks. Therefore, constructing an effective detection method to detect the attackers and remove them from the CFRSs is vital. A number of detection methods have been proposed to make CFRSs obstruct to such attacks. Some of them were based on calculating similarity between users (consists of attackers and genuine users) in order to differentiate those attackers. In practice it is hard to catch all concerned attackers by exploiting the similarity of users, albeit it can be helpful to filtering out more genuine users and show high computation time, albeit they can be effective to capture the concerned attackers in some extent. To resolve all the issues in existing systems we propose a new detection method called 0 suspected users by exploiting specific distance metric in collaborative filtering recommender systems to make CFRSs resistant to such attacks, which exploits a novel metric for calculating similarity between users. To detract the computation time, we firstly filter out a part of genuine users in order to detract the computation time. Since the attackers will target one or more specific items with lowest or highest rating many times if they want to demote (called nuke attack) or promote (called push attack) the items to the recommendation list, we can find out all suspected target items by using an complete count threshold. At the second stage, we primarily concentrate on the effective similarity metric to better distinguish between attackers and genuine users based on the remaining users of the first stage.

II. LITERATURE SURVEY

No	Paper Name	Author Name	Proposed System	Referred Point	Existing Disadvantages	Paper Advantages
1.	A novel approach to filter out malicious rating profiles from recommender systems	C. Chung, P. Hsu, and S. Huang	This paper proposes novel detection method to make recommender systems resistant to the “shilling” attacks or “profile injection” attacks.	From this Paper, we have referred problem as finding a mapping model between rating behavior and item distribution by exploiting the least-squares approximate solution.	• Existing system prior works eliminated attacking profiles with classification or PCA based methods. • They either suffer from the lack of negative cases or cannot cope with sparse data.	• To avoid both issues, the work proposed a method based on beta distributions.
2.	HHT-SVM: An online method for detecting profile injection attacks in collaborative recommender systems	F. Zhang and Q. Zhou	This paper propose online method (called HHT–SVM) to detect profile injection attacks by combining Hilbert–Huang transform (HHT) and support vector machine (SVM), which can work incrementally.	From this Paper, we have referred the empirical mode decomposition (EMD) approach.	• The existing classification methods are not appropriate for situations where the attack profiles are injected over time. • They require examining and processing the entire rating database to detect the attacks. • Detection precision is low.	• HHT–SVM detects profile injection attacks by combining HHT and SVM. • This is the basis for HHT–SVM to operate online. • Depend on the novelty- and popularity-based rating series
3.	Detection of abnormal profiles on group attacks in recommender systems	W. Zhou, Y. S. Koh, J. H. Wen, S. Burki, and G. Dobbie	This paper proposes the FIM technique to generate candidate groups, and then present several features for modeling the abnormal behavior at the group level.	From this Paper, we have referred the FIM technique.	• Malicious profiles are a set of items that make the profile look normal and make them harder to detect.	• DegSim and RDMA are used for identifying group attack profiles.

4.	Defending recommender systems by influence analysis	M. Morid and M. Shajari	This paper proposes an attack detection method based on user influence in recommender systems.	From this Paper, we have referred an attack detection method.	• Attack detection performance is low.	• Improve attack detection performance. • Better detection performance. • Improved the stability of a recommender System.
5.	Detection of shilling attacks in recommender systems via spectral clustering	Z. Zhang and S. R. Kulkarni	This paper defines the issue as finding a most extreme submatrix in the user-user similarity matrix.	From this Paper, We have referred a spectral clustering algorithm to find the min-cut solution to estimate the highly correlated group.	• Shilling attack detection precision is low.	• Search maximum submatrix by translating the matrix into a graph and apply a spectral clustering algorithm to improve detection precision.
6.	Shilling attacks against recommender systems: A comprehensive survey	I. Gunes, C. Kaleli, A. Bilge, and H. Polat	This paper proposes decoding shilling attack detection schemes in detail and robust algorithms proposed so far might open a lead to develop new detection schemes and increase such robust algorithms.	From this Paper, We have referred robust recommendation algorithms.	• Some shilling attack types are missing.	• Described the missing scopes. • Covered possible shilling attack types and explained them briefly. • Grouped shilling attacks according to some dimensions.
7.	Graph-based detection of shilling attacks in recommender systems	Z. Zhang and S. Kulkarni	This paper presents a method to make recommender systems obstructive to these attacks in the case that the attack profiles are highly correlated with each other.	From this Paper, We have referredhow find a maximum submatrix in the similarity matrix.	• Shilling attack detection precision is low.	• Improve shilling attack detection precision.
8.	A survey on shilling attack models and detection techniques for recommender systems	Z. A. Wu, Y. Q. Wang, and J. Cao	This paper reviews the states of art and the main problems of existing works related to shilling attack models and detection techniques, and attempts to sketch an extensive and evident outline for this new and active research realm.	From this Paper, We have referred the shilling attack models and detection techniques.	• All detection techniques are not given into in single survey.	• Present a survey of existing research on the shilling model, algorithm dependence, attack detection, and attack evaluation metrics.

9.	Attacking item-based recommender systems with power items	C. E. Seminario and D. C. Wilson	This paper shows that the Power Item Attack (PIA) is ready to affect not only user-based and SVD-based recommenders but also the heretofore highly robust item-based approach, using a novel multi-target attack vector.	From this Paper, We have referred the robust item-based approach.	• Item-based systems remained robust to the PUA (Power User Attack)	• Investigate a new, complementary attack model, the Power Item Attack (PIA). • PIA is able to impact not only user-based and SVD-based recommenders but also the heretofore highly robust item-based approach, using a novel multi-target attack vector.
10.	Pair wised specific distance learning from physical linkages	J. Hu, D. C. Zhan, X. Wu, Y. Jiang, and Z. H. Zhou	This approach exploits the structures of physical linkages and in particular captures the key observations that nonmetric and clique linkages imply the appearance of different or unique semantics, respectively.	From this Paper, We have referred PSD for multi-class learning and further extend it to multi-label learning.	• Existing approaches aim to construct a global distance metric that is applicable to all data points. • However, different data points may have different properties and may require different distance metrics. • Data points in real tasks are often connected by physical links but these links information has not been exploited in distance metric learning.	• Develop the pair wised specific distance (PSD) approach that exploits the structure of physical linkages. • PSD will generate different distances for different pairs of data points • PSD is useful especially in the scenarios where there are very limited labeled training data points and no explicit constraints are given.

III. EXISTING SYSTEM APPROACH

Constructing an effective detection method to detect the attackers and remove them from the CFRSs is crucial. In the existing systems a number of detection methods have been proposed to make CFRSs resistant to such “shilling” attacks or “profile injection” attacks. Some of them were based on calculating similarity between users (consists of attackers and genuine users) in order to distinguish those attackers. Practically speaking it is hard to catch all concerned attackers by exploiting the similarity of users, although it can be helpful to filtering out more genuine users and attackers and show high computation time, although these existing systems can be effective to capture the concerned attackers in some extent.

IV. PROPOSED SYSTEM APPROACH

In this Paper, we propose a new detection method to make CFRSs resistant to such attacks, which exploits a novel metric for calculating similarity between users. To lower the computation time, firstly we filter out more genuine users as far as possible determined by using mistrusted target items. Since the attackers will target one or more specific items with lowest or highest rating many times if they want to demote (called nuke attack) or promote (called push attack) the items to the recommendation list, we can find out all mistrusted target items by using an complete count threshold. Based on the remaining users, we employ a new similarity metric inspired from the pair wised specific distance, directing to measure impressively the similarity between users.

V. PROPOSED ARCHITECTURE

Our proposed approach consists of two stages: the stage of filteringout genuine users by misusing suspected target items andthe stage of separating attackers by employing a newsimilarity metric. At the first stage, we filter out a part ofgenuine users in order to decrease the computation time. To decide the target items, wedesign an absolute count threshold ε which pushes or nukesthe same items with the highest or lowest at least ε times. Ifthe count for an item is greater than ε, then the item *susi*isregarded as suspected target item. Users (consist of attackersand genuine users) who rated the *susi*with the highest *rmax*or lowest *rmin*are considered as attackers. It is noticeable that there will be more false positives or false negative if $\varepsilon$is too small or large. Atthe second stage, we mainly focus on the effective similarity metric to better differentiate between attackers and genuine users based on the remaining users of the first stage.

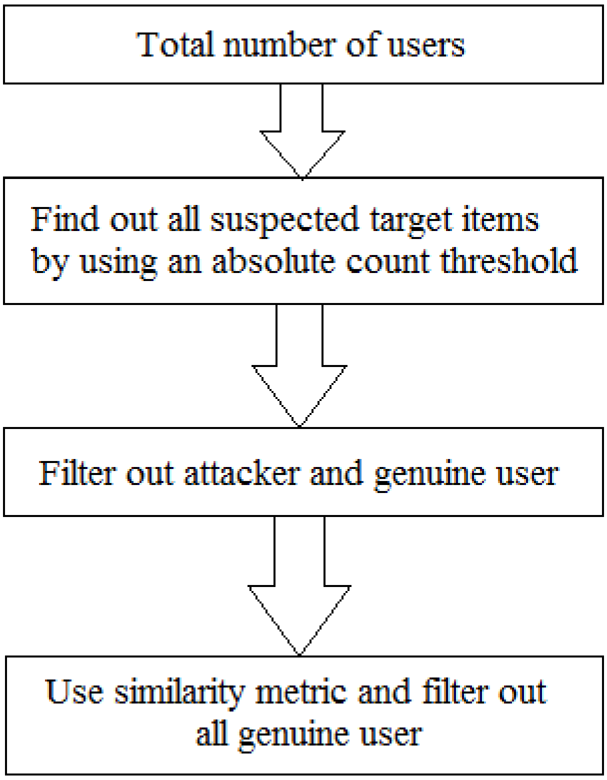
Total number of users
Find out all suspected target items by using an absolute count threshold
Filter out attacker and genuine user
Use similarity metric and filter out all genuine user

Figure 1: System Architechture

VI. CONCLUSION

In this paper, we have exhibited an unsupervised detection method called defending suspected users by exploiting specific distance metric in collaborative filtering recommender systems for detecting attacks like “shilling” attacks or “profile injection” attacks, which exploits the pair wised specific distance to generate a similarity metric. Firstly, we filter our more genuine users as far as possible determined by using mistrusted target items. And then, based on the remained users, we continue to filter out more genuine users used by an empirical threshold of the new similarity metric.

REFERENCES

[1] C. Chung, P. Hsu, and S. Huang, “A novel approach to filter out malicious rating profiles from recommender systems,” Journal of Decision Support Systems, pp. 314–325, 2013.

[2] F. Zhang and Q. Zhou, “HHT-SVM: An online method for detecting profile injection attacks in collaborative recommendersystems,” Knowledge-Based Systems, 2014.

[3] W. Zhou, Y. S. Koh, J. H. Wen, S. Burki, and G. Dobbie, “Detection of abnormal profiles on group attacks in recommender systems,” Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval, pp. 955–958, 2014.

[4] M. Morid and M. Shajari, “Defending recommender systems by influence analysis,” Information Retrieval, pp. 137–152, 2014.

[5] Z. Zhang and S. R. Kulkarni, “Detection of shilling attacks in recommender systems via spectral clustering,” International Conference on Information Fusion, pp. 1–8, 2014.

[6] I. Gunes, C. Kaleli, A. Bilge, and H. Polat, “Shilling attacks against recommender systems: A comprehensive survey,” Artificial Intelligence Review, pp. 1–33, 2012.

[7] Z. Zhang and S. Kulkarni, “Graph-based detection of shilling attacks in recommender systems,” IEEE International Workshop on Machine Learning for Signal Processing, pp. 1–6, 2013.

[8] Z. A. Wu, Y. Q. Wang, and J. Cao, “A survey on shilling attack models and detection techniques for recommender systems,” Science China, vol. 59, no. 7, pp. 551–560, 2014.

[9] C. E. Seminario and D. C. Wilson., “Attacking item-based recommender systems with power items,” ACM Conference on Recommender Systems, pp. 57–64, 2014.

[10]J. Hu, D. C. Zhan, X. Wu, Y. Jiang, and Z. H. Zhou, “Pair wised specific distance learning from physical linkages,” ACM Trans. Knowl. Discov.Data, 2014.