iJRASET
International Journal For Research in Applied Science and Engineering Technology

Volume: 11 Issue: V Month of publication: May 2023

DOI: https://doi.org/10.22214/ijraset.2023.51951

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Using Machine Learning Algorithms to Alleviate the Shilling Attack in a Recommendation System

Anurag Singh[1], Dr Subhadra Shaw[2]
[1]Research schooler, [2]Associate Professor of Computer Science & Application, AKS University Satna(M.P), India,

***Abstract:** Every e-commerce site must have a referral system. Shilling attacks, One of the biggest issues with recommendation systems is those that involve the creation of fake profiles in the system and biased ratings of things. They reduce the accuracy and improve the performance of the recommender system when making suggestions to users. The goal of attackers is to change the order of materials or objects that match their interests. Shilling attacks threaten the reliability of RS. Therefore, to maintain their validity and fairness, recommender systems must be able to detect shilling attacks. So far, suitable algorithms and methods have been presented for the detection of shilling attacks. Some of these approaches, though, either use low-order interactions or higher-order interactions, or they examine the rating matrix from a single point of view. This study aimed to propose a mechanism using users' rating matrix, rating time, and social network analysis output of users' profiles by Gaussian-Rough neural network to simultaneously use low-order and high-order interactions to detect shilling attacks. Finally, several experiments were conducted with three models: CNN, DNN, and RNN compared with the Proposed Model. The results indicated that the proposed method is more effective than the comparison methods regarding attack detection and overall detection, which proves the effectiveness of the proposed model.*

***Keywords:** Recommender systems, Collaborative filtering, Rating, Defense systems, Shilling Attack*

I. INTRODUCTION

Recommender systems (RSs) are common in e-commerce, where recommendations of items may be helpful to a customer in finding the item of his/her interest. RS predicts the ratings that would be given to an item by a user. Recent research has examined the robustness and vulnerabilities of various collaborative filtering (CF) techniques for recommendations. A collaborative filtering recommender system (CFRS) is vulnerable to “profile injection” or “Shilling Attacks(SA)” [1]. Anonymous users, who cannot be easily distinguished from genuine users, insert a huge number of pseudonymous profiles into the system with the intent of manipulating its recommendation concerning a target item. User-based CF and item-based CF are the two groups of CF recommendation techniques, respectively [2][5]. Being open makes CFRS susceptible to shilling attacks. The user-based algorithm at CF compiles user profiles that represent the tastes of various people and offers product recommendations and forecasts based on the viewpoints of other profiles with similar interests. If the database of a system contains data that is biased, then the attackers' profile may become peer to original users and may produce results in favour of attackers. In the case of item-based filtering, the items similar to the target item are considered, and predict the user's ratings on these similar items. From the attacker's perspective, two types of efforts are there to mount an attack. First, the amount of knowledge needed for mounting an attack. In high-knowledge attacks, an attacker must know the distribution of ratings in a system [3]. A low knowledge attack does not require details of the system. The second aspect is the effort required to add the number of profiles and ratings in the system's database to make the attack effective. However, the ratings have less importance as automated software agents can be used for inserting the ratings. Sites may employ policies that limit the speed of profile multiplication. Therefore, an attack that needs a large number of injected profiles in this system is less practical than needs a smaller number of injected profiles [4].

II. ATTACK MODEL

The Recommendation system helps the user find items or products according to their interest and preferences. When a user enters the system and starts injecting it into the system, it is known as a shilling attack. The attacks include profiles that contain biased information about malicious users. Attackers are successful even if they have little or no knowledge of the system.

Segment-based attacks, contrary to collaborative filtering, guarantee that the item the attacker is attacking is most likely to be suggested to the target users. Furthermore, some common attacks are Average, Bandwagon, random, Segment, Reverse Bandwagon, and Love/hate attack [5]. Average attack desires knowledge about the system because it considers separate average ratings for each item rather than the universal rating system.

Attackers choose items randomly and ratio those using ordinary scattering with mean rating and standard deviation [5]. A random attack is a low-knowledge-degree attack. The items are decided on at random and are rated using regular scattering based totally on standard deviation (SD) and a median valuation of the device. items set are empty or null. The target object set is rated in step with the shape of the attack like a nuke or push attack [18]. The most trending and liked by way of lots of customers items are popular. therefore, the possibilities are excessive that a high attacker will become much like the actual user. So, it's miles difficult to pick out the attacker from many of the real users. The malicious consumer makes a collection of segmented items that have a better chance of being desired by way of goal users, who belong to him/ her in a sure section. This creates a huge effect of the attack in which minimum scores are given to the items which are inside the filler set [18]. The reverse Bandwagon attack is also a type of nuke attack in that favoured things are like disliked matters, and the maximum rating values are given through the user. the love/hate attack is an efficient nuke attack in which the non-goal items randomly select the filler items. The score of all filler items is maximal in place of the target item [6].

III. RELATED WORKS

Many researchers attempt statistical machine-learning strategies to identify shilling attacks. The publicity of shilling attacks in collaborative RS got the eye of numerous researchers in the latest past. various disclosure techniques were advised as much as the previous time that can be divided into semi-supervised techniques, unsupervised methods, and supervised methods.

Within the league of supervised detection techniques, the authors have provided level technique via exploiting a support vector system (SVM) situated classifier and target object evaluation. This approach will perceive the attacks together with a huge filler length, and the success of detecting the attack is very terrible when the detection of an attack with a small filler and small attack sizes [6].

The authors also provided a number of statistical indicators for attack profiles with high and low densities [7]. In a different paper, the authors introduced a fine-grained recommendation system (RS) for a social ecosystem designed to recommend users' friends. The main purpose is to find a consistent way to obtain information representation of overlapping interests in different subcategories [8].

Furthermore, it is advised that RS data used in social networking websites be supported by an ontology, and this approach is supported by a shared ontology model useful for both the contented advertisement and user profiles[9].

Additionally, the RS fits the category of unsupervised detection model perfectly for assisting consumers with the already-present features. Several algorithms, such as the K-mean algorithm, mini-batch K-mean, mean-shift algorithm, and clustering algorithm are used by recommender systems to access groupings[10].

likewise, researchers have included the "SA" in some attack models already in use, attack detection algorithms, and cost-benefit evaluations in other works. this is in addition to our knowledge of the most up-to-date and complete survey of the "SA"[11]. Some of the security attacks are determined and solved using machine learning methods, but ignored shilling attacks [12][13].

Undoubtedly, security threats can also be raised due to false information and to avoid it, experimental analysis is conducted in this paper. This paper is on the shilling attacks on the RS application Entertainment category where various machine learning algorithms are applied to check the accuracy of the identification of shilling attacks on trending YouTube videos.

To defend against shilling attacks, some solutions use both preventive and reactive mechanisms to mitigate the impact of shilling attacks in the victim network, the intermediate network, and the source network [14].

A topic-level recommendation algorithm built on trust was found by Zhang [15] to be more secure. When subjected to mean attack, he adds a topic-based trust model to CF algorithms and comes to the conclusion that this strategy is more stable than the conventional kNN method. To increase accuracy, Donovan and Smyth introduced confidence-based models in CF [16].

Later study addressed this issue by changing the trust-building procedure, which reduces prediction shift by 75% when compared to the traditional CF method. Due to the scant data in the ratings matrix, it is challenging to calculate the similarity between users; therefore, a trust metric is necessary to resolve this issue. A reliable CF method built on the "web of trust" metric was introduced by Avesani and Massa [17].

IV. METHODOLOGY

The methodology used in this paper is observed in Fig. 1.

Data Set Selection: Based on the literature review, we identified and selected the most suitable dataset for the validation of machine learning algorithms in shilling attacks. The collective dataset is the Amazon dataset from Kaggle.

- Data pre-processing: In this phase, we pre-processed the dataset using cleaning or filling in the missing values with the correlation. For this, data regarding service domain sites need to be considered. Therefore, we analyzed and categorized a data set as per the requirements of the study.

After pre-processing the dataset. We modified the rating in the dataset that modified rating is called a shilling attack by an attacker. Split the dataset in Train and test, Train the model with BRNN, and test the prediction. And result accuracy is 98.7%.

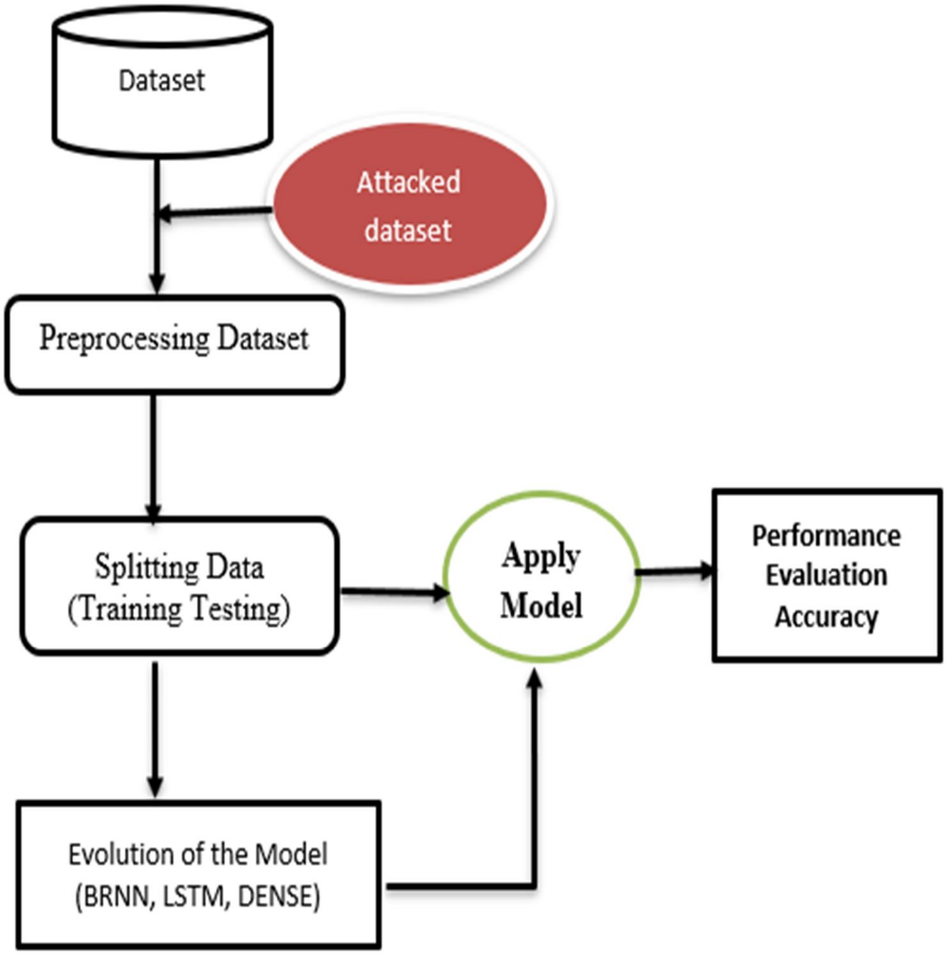

Dataset
Attacked dataset
Preprocessing Dataset
Splitting Data (Training Testing)
Apply Model
Performance Evaluation Accuracy
Evolution of the Model (BRNN, LSTM, DENSE)

Figure. 1. Proposed Methodology

Performance analysis: We conducted a performance analysis based on the results obtained from the performance analysis phase shown in Table 1.

Table 1: Comparison of proposed shilling attack methods based on accuracy

Method	Accuracy (%)
CNN [94]	88.45
DNN [94]	92.75
RNN [94]	94.28
Proposed	98.7

A graphical comparison is shown below

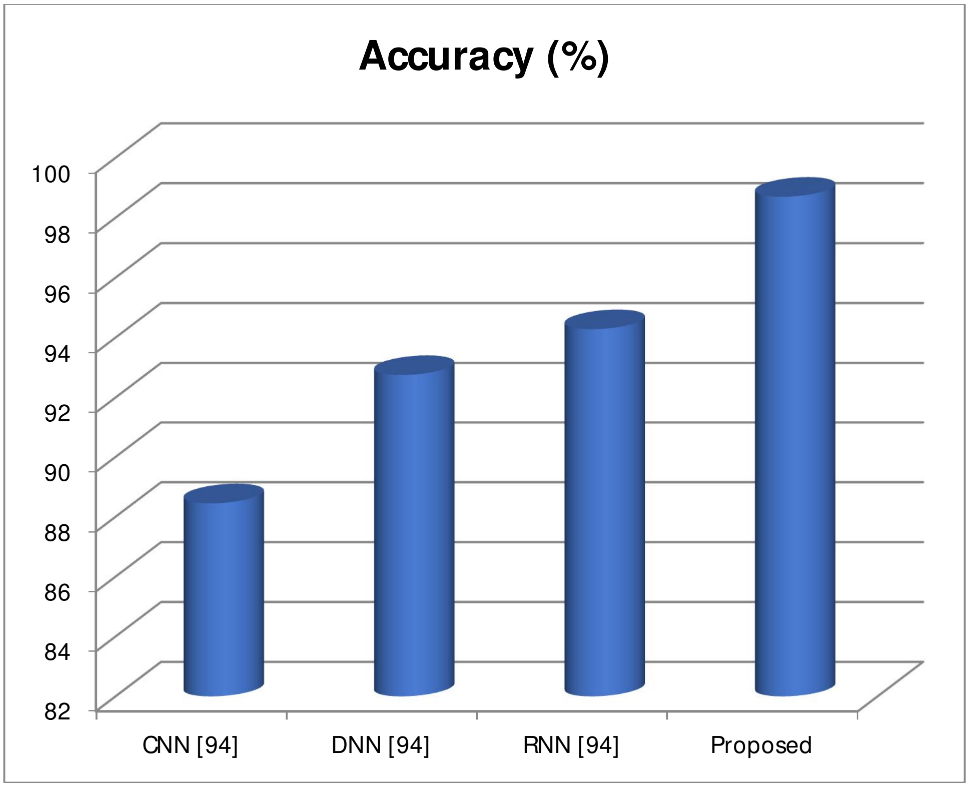


Figure. 2. A graphical comparison

V. CONCLUSION

We have assessed and examined the success of various machine-learning algorithms in this article. On information from the Amazon dataset, we conducted an experimental investigation. This dataset can be used to learn more about user preferences for specific historical genres or to forecast a product's success based on its rating performance, among other things. We examine and evaluate three different ML algorithms: CNN, DNN, and RNN. The algorithm is assessed based on its precision and accuracy. Table 1 presents the findings. After reducing attacks, the proposed solution increases the recommendation accuracy to 98.7%.

VI. FUTURE WORK

For future work, we will plan to increase the accuracy of this model and the diversity of this model and system settings to test our model's robustness in different environments. In future work, we will put forward a more versatile algorithm to identify various target items. Furthermore, further research work will also discover the hidden relationships between users and items in recommender systems, and present additional group features for detecting attack users.

REFERENCES

[1] Lam S. (2004). "Shilling Recommender Systems for Fun and Profit", In Proceedings of the 13th international conference on World Wide Web, ACM, pp. 393-402.

[2] Herlocker, J., Konstan, J., Borchers, A. and Riedl, J. (1999). "An algorithmic framework for performing collaborative filtering", In Proceedings of the 22nd ACM Conference on Research and development in Information Retrieval (SIGIR'99).

[3] Chichani , A ; Golwala , J; Gundecha, T; Gawande,K (2018). "Advancing recommender systems by mitigating shilling attacks", 9th ICCCNT 2018 July 10-12, 2018, IISC, IEEE – 43488, Bengaluru Bengaluru, India

[4] Tong C, Yin X, Li J, Zhu T, Lv R, Sun L, Rodrigues JJ (2018) A shilling attack detector based on convolutional neural network for collaborative recommender system in social aware network. Comput J 61(7):949–958

[5] Mobasher, B., Burke, R., Bhaumik, R. and Williams C. (2005). "Effective Attack Models for Shilling Item-Based Collaborative Filtering Systems", In Proc. of the 2005 Web KDD Workshop, Chicago, Illinois, 2005.

[6] Cai H, Zhang F (2019) Detecting shilling attacks in recommender systems based on analysis of user rating behavior. Knowl-Based Syst 177:22–43

[7] Zhou Q, Wu J, Duan L (2020) Recommendation attack detection based on deep learning. J Inf Secur Appl 52:102493

[8] Aivazoglou M, Roussos AO, Margaris D, Vassilakis C, Ioannidis S, Polakis J,Spiliotopoulos D (2020) A fine-grained social network recommender system. Soc Netw Anal Min 10(1):8

[9] Garcı´a-Sa´nchez F, Colomo-Palacios R, Valencia-Garcı´a R (2020) A social-semantic recommender system for advertisements. Inf Process Manage 57(2):102153

[10] Putri DCG, Leu JS, Seda P (2020) Design of an unsupervised machine learning-based movie recommender system. Symmetry 12(2):185

[11] Gunes I, Kaleli C, Bilge A, Polat H (2014) Shilling attacks against recommender systems: a comprehensive survey. Artif Intell Rev 42(4):767–799

[12] Kumar PV, Reddy VR (2014) A survey on recommender systems (RSS) and its applications. Int J Innov Res Comput Commun Eng 2(8):5254–5260

[13] Bland JA, Petty MD, Whitaker TS, Maxwell KP, Cantrell WA (2020) Machine learning cyberattack and defense strategies. Comput Secur 92:101738

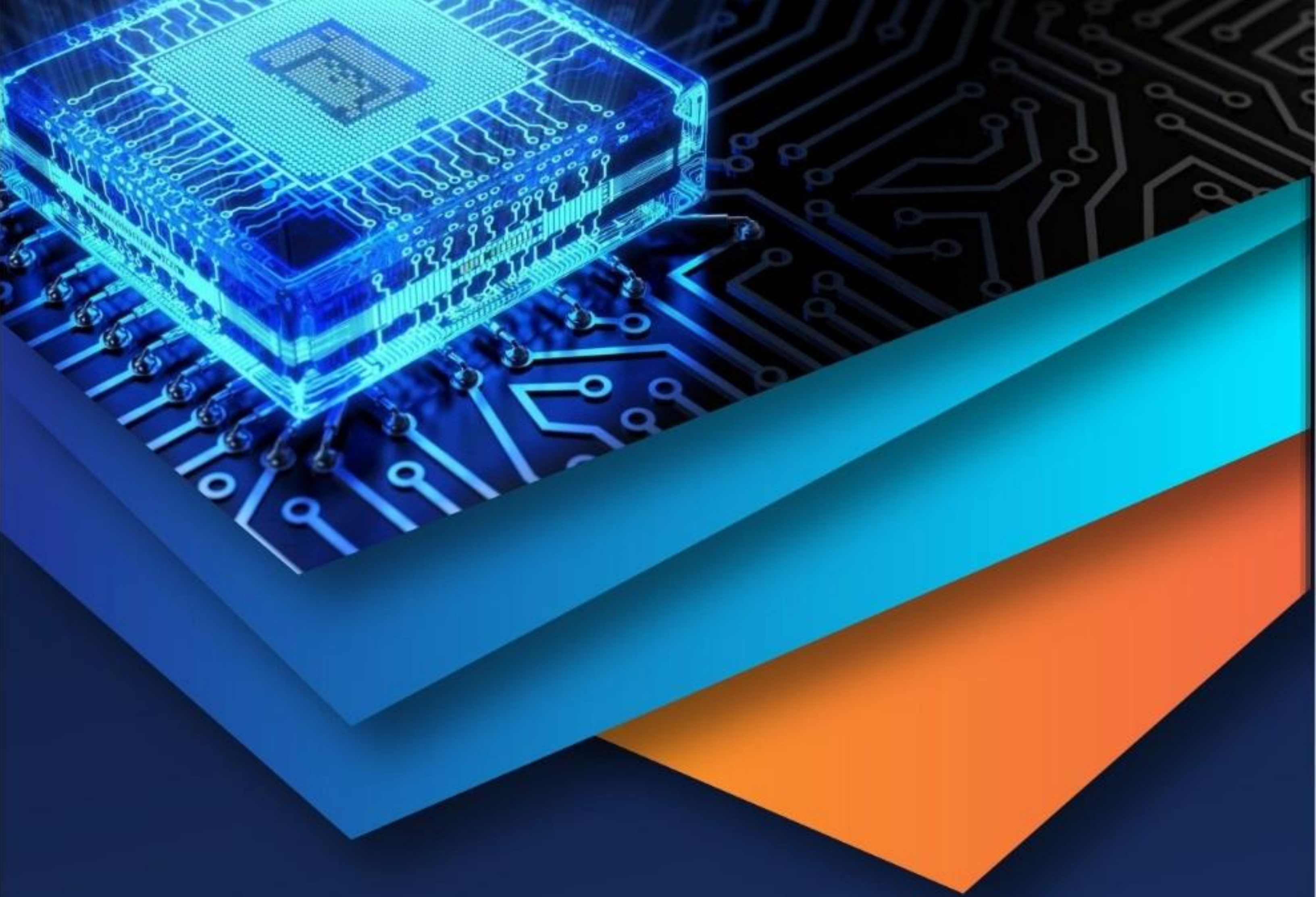
[14] Mirkovic J. and Reiher P.(2004). “A taxonomy of ddos attack and ddos defense mechanisms,” ACM SIGCOMM Computer Communication Review, vol. 34, no. 2, pp. 39–53,

[15] Zhang F. (2009). "Average Shilling Attack against Trust-Based Recommender Systems", International Conference on Information Management, Innovation Management and Industrial Engineering, pp. 588-591.

[16] O'Donovan , J. and Smyth , B.(2005). "Trust in Recommender Systems", IUI, Association for Computing Machinery, New York, NY, USA..

[17] Massa P. and Avesani P.(2007), "Trust-aware recommender systems", in Proceedings of the 1st ACM Conference on Recommender Systems (RecSys ’07), pp. 17-24.

[18] O’Mahony MP, Hurley NJ, & Silvestre GC (2005) Recommender systems: Attack types and strategies. In AAAI, pp. 334–339.



doi crossref

10.22214/IJRASET

INDEX COPERNICUS

45.98

ISRA JIF

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

ISI

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY