International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 /doi:10.14445/22315381/IJETT-V69I3P233
Volume 69 Issue 3, 219-228, March 2021

SRSPIA: Security Based Recommender System Against Profile Injection Attack

Anjani Kumar Verma[#1], Veer Sain Dixit[*2]

[#1] *Department of Computer Science, University of Delhi, India*

[*2] *Department of Computer Science, ARSD College, University of Delhi, India*

[1]anjaniverma29@gmail.com, [2] veersaindixit@rediffmail.com

Abstract *— Information produced by users in terms of rating movies is a major ingredient on the web. Information filtering is the most popular way to obtain precise information that could be used for different e-commerce applications for branding and popularity. Today, security is the most prominent aspect of the safety of data available online. In this work, a novel Security-based Recommender System against Profile Injection attacks (SRSPIA) has been proposed for movies that consist of three phases. In the first phase, the collection of the data is crawled based on user rating behaviors on the movie(s) without the security integration in the system. In the second phase, the collection of the data is crawled after the security integration in the system. In the third phase, a model is designed which works in two modes. In the first mode, the different machine learning supervised classifiers are applied with the two datasets obtained from the above two phases individually. In the second mode, ensemble methods are applied to these datasets. It is observed that the accuracy has been improved with the ensemble approach on various performance evaluation metrics that have been used for analyzing the system. However, the performance is evaluated on this model and found that the results in the case of the proposed system dataset phase 2 show better accuracy in comparison to another phase and dataset.*

Keywords *— Recommender Systems, Profile injection attack, Security, Classifiers, MAE*

I. INTRODUCTION

The growth of Internet Technology rapidly changes; lots of information is available on the Internet. Organizing information, recommendations have become useful for better profiling.

Profile filtering plays the most important part in a recommendation system that classifies the data of several users. But nowadays, much suspicious content is also used to inject into the users' profiles available on the Internet. Therefore, critical questions arise; should we use an online recommender system for placing our data or not? Whether we should give partial information on different e-commerce sites? It is called feature selection in machine learning. When data is transformed into a matrix, rating on an item is regarded as a feature. It is important for the classification of different data available on the Internet. The high dimensional data is always hard to handle; using feature selection, the dimensionality of data could be reduced with the elimination of redundancy and irrelevant features to improvise the performance. The profile misuse is termed as a shilling attack in this account. Profile injection attack is most popular these days, as there is a huge availability of data on different portals. Using which a malicious person (or attacker) can enroll in the login session of some active user and molds the information by injecting the false information.

There are different challenges in profile injection attacks; a supervised classification method is used in the identification of attack profiles from the rating database as per the detection attributes [1-3] that are computed for the profiles present in the rating database. To discriminate genuine profiles from bogus ones is a difficult process. For instance, the attackers may change the rating style and destroy the sense of the items. As a result, the detection of attacks is difficult in this situation and suffers from a bad performance detection rate. Improvement under the classification of profiles is required new features, which are not dependent upon any particular rating pattern by the users [27].

The term "shilling" has coined [3] that is based on different attack models like Random and Average, where the attackers inject profile into the system [50]. They explored the factors for finding the attacks effectively, through which algorithm has been used, whether the system is generating predictions or recommendations, whether the attacks are detectable or not, and the characteristics of the item being attacked. For generating the recommendations, CF-based automated systems are used. Prediction shift and MAE are the evaluation metrics that have been used to measure the effectiveness detection. These evaluation metrics are most suitable for the systems that produce predictions for the items. The effectiveness of the attack is also influenced by the rating distribution of the target item. They hypothesized that the characteristics of the rating distribution of an item influence the impact of an attack on that item. Features of an item include popularity, entropy, and likability. The ratings of unpopular items and items with a high spread of ratings can be manipulated easily.

Hence, the RS is the best way to analyze the rating behavior of the users and predict future recommendations, which are secure from any type of hindrance in a bad manner.

II. PROBLEM DEFINITION

By this assumption that everyone has lots of accounts on different media like social sites, e-commerce sites, movie reviews, and rating sites. The users are keener to always log in to their web or app on their workstations such as laptop, mobile, etc. If we handover our laptop or mobile intentionally to someone for use, or if someone else has access to our laptop or mobile with malicious intentions, then there is a possibility of misuse by making changes in the logged account, that malicious person might create a problem with the stealing of information and change the meaning as well. There is a need to analyze the impact of the changes made by other users in the corresponding account.

Keeping this in mind, this paper proposed a rating system for review and rating the movie on the web where anyone can submit ratings for a movie by creating an account on it. A logger can like or dislike a movie and give it the highest to lowest ratings on a scale of 1 to 5. But if the user forgets to logout from the account and meanwhile, another malicious person has access to that account, then it could be the possibility that different views or opinions can be preserved about any movie in the session of the logged user. A disliked movie could be liked or vice versa and rated like an opposite to the logged user by decreasing or increasing the rating for a movie. It may also give an impact on the recommendation system that could recommend a movie to the logged user or active user, which is different from their interest. This paper proposes a secure-based recommender system for movies, which is designed to prevent the misuse of accounts and information's to make accurate recommendations further.

The major highlights of this paper are as follows:

- A Secure recommender system (SRSPIA) is proposed.
- An ensemble model is used with various machine learning supervised classifiers.
- Experimental performance evaluated on different metrics to get the best performance on different datasets.
- This research presents data from misuse and degraded the reputation of e-commerce businesses in the market through better security integration in RS. This is the best practice for every user to safely visit the sites and prevent them from profile attack.

The remainder of this paper is organized as follows: In section 2, the problem definition is stated; in section 3, research background is explained, in section 4, material and methods is described, in section 5, description about the datasets used, in section 6, results are analyzed that are obtained from the experiments performed. Finally, the Conclusions are carefully placed in section 7.

III. RESEARCH BACKGROUND

Earlier several kinds of research have been conducted on profile injection attack detection. Unsupervised and supervised are the methods that are being used for detecting attack profiles [27]. There are several types of research on the problem of shilling attacks or profile injection attacks [5] that have been done which are summarized in table 1.

Table 1. Summary of research related to profile injection attack

Sources	Related Research
[30]	There are several metrics used for finding out the attacker's rating patterns, but it is unsuccessful for detection under small size attack.
[31]	They have developed an algorithm, which is detecting the groups of similar attacks. The computational complexity is high in this case.
[32]	A detection algorithm, UnRAP is proposed here which is used to identify attack profiles that give good on average and random attack model but bad in case of bandwagon attack (i.e. in case of filler size is small).
[33]	A hybrid method is proposed for detection purposes; multidimensional scaling methods with clustering-based methods [27] are used. In the case of Random and average attack, it effectively detects the attack.
[34,35]	Proposed PCA-Var-Select used for filtering. It takes advanced knowledge to effectively detect the various attacks [27].
[36–38]	Used supervised classifiers for the detection but suffer from low accuracy since too many genuine profiles are misclassified as attack profiles.
[39]	Uses a rough set theory for the detection; it detects most of the attack profiles but with low precision [27].

There are various shilling attack types proposed so far in the literature studies for attacking rating-based CF schemes. These attack types are present with the strategies, which are based on preferences given by the users, such as numeric and binary. To estimate the prediction, numeric ratings are more useful rather than binary. The online vendors or customers have keener towards an item with their choices of liked or disliked, and sometimes it is like how much the item is liked or disliked. Generally, we categorized these into binary and numeric ratings, respectively. Therefore, the numeric rating CF scheme and binary rating based CF scheme are more popular in the recommendations schemes [6].

Several types of research are available on shilling attacks, which is focused on statistical approaches is used for the detection of unusual patterns in the ratings. Tackling the problem profile injection attacks, machine-learning techniques can handle it easily. Different problems occur in shilling attacks: reduce the inclusion of malicious profiles from the dataset by introducing the

security aspect with "captcha" component or email authentication, which is under system defense. Another problem is the detection of profile attacks and removes them by using statistical methods with the CF approach, which is under the CF defense. But there are some limitations with this, such as all the malicious profiles cannot be filtered out; genuine profiles are wrongly filtered; using statistical information from the CF machine learning methods [8].

There is various Recommendation System operation based on filtering such like [9]: Content-based [10], where similarity of the contents can be used for further recommendations; Demographic-based [11], where recommendations are done through the consumption of items by users demographically similar; Social-based [12], where users graphs can be seen with the network users of same liking patterns; Context-aware [13], Collaborative Filtering (CF) [14] recommends, targeting user by collecting similar rating of the unknown users on an item.

Currently, the three filtering approaches such as Content, collaborative, and their hybrid [15], are used for commercial purpose. Hacking is one type of attack that is content-based and not vulnerable to profile injection attacks. Conversely, the more vulnerable is the CF approach; users are creating and updating their implicit or explicit ratings.

The shilling attack detection is categorized in the following manner [16][8]: Detection based on Classification, Detection based on Supervised learning, and Detection based on unsupervised learning. In detection-based classification there are statistical measures used such as RDMA [1], WDMA [18], DegSim [2], and Entropy [19]. The statistical attributes help the quality of variations in the detection of attacked RS [1]. There is the model-based approach used, which is the binding of profile and item-based algorithms that can improve the shilling attack detection rate [20].

In detection-based supervised learning, it is transformed into a binary problem classified as a genuine and fake rating. Naive Bayes [21] was extensively used in this case. [22], the author has used the generic statistical attributes to identify fake profiles from genuine ones; after obtaining values, it gets to train a model for the detection of malicious profiles.

Detection-based unsupervised algorithm [35], there is common detectors such as Principal Component Analysis (PCA) method [24], which is based on matrix computations. When it is combined with data complexity [25], it will refine the profiles; choose the authentic ones from the malicious profiles. It is also used to discover groups of a shilling, just like in Amazon [24]. A two-phase detector that computes the dissimilarity matrix between users is used with the Multidimensional Scaling (MDS) algorithm [26]. The hidden factors from the matrix are getting through the Support Vector Machine (SVM) [27]. Usually, fake ratings were added in a short period that can be done under a time-based detection scheme [28]. Clustering [29] is used to grouping the classes for ease to detect malicious groups.

A. Types of Attacking Models

The model is a strategy for attackers to inject bogus information in the rating database based on the knowledge of the RS. The structure of an attack profile can be seen in table 2 [40]. A profile that is used for an attack contains a rating vector on items in the RS are divided into four sets: I_S, I_F, I_N and, I_T.

I_S: It is the random selection of items that have a relationship with the target items [50]. Their ratings are generated by the function δ (I_S).k

I_F: random selection of filler items where ratings are done by the function β (I_F).m

I_N: Set of not rated items.

I_T: Rating is assigned to a maximum for the target item, i.e., for push γ (I_T)=r_{max} or minimum, i.e., for nuke γ (I_T)=r_{min}.

Table 2. Attack profile Structure

I_S	I_F	I_N	I_T
I_{S1} … I_{Sk}	I_{F1} … I_{Fm}	I_{N1}… I_{Nq}	I_{T1} … I_{Tn}
$\delta(I_S)_1 \ldots \delta(I_S)_k$	$\beta(I_F)_1 \ldots \beta(I_F)_m$	Null…null	$\gamma(I_T)_1 \ldots \gamma(I_T)_n$

There are different types of attack models (see table 3). The push attack models that are used for mounting nuke attack. This can be achieved by assigning a minimum rating, i.e., $r_{min,}$ to the target item instead of r_{max}. There are two attack models for nuking items that have low knowledge about the systems.

Table 3. Summary of different attack models

Push Attack Models	
Random Attack	● Low knowledge attack, the random selection of filler items (IF) ● A normal distribution is used for rating and SD and means rating of the system [40]. ● Item is assigned either minimum or maximum rating on a scale of 1 to 5. ● Not much effective.
Average Attack	● Previous knowledge of the system is required; an average of ratings is performed [40]. Random selection of filler items. ● The normal distribution is used for rating and SD and means rating of the system [40]. ● Highly effective.
Bandwagon Attack	● Generates the biased profiles using Zipf's law distribution that contain the most popular items [40]. ● High possibility of similarity between attackers and the actual users. ● Low knowledge attack [50].
Segment Attack	● Requires less knowledge about the system [50]. ● Segment the rating to maximum, i.e., IS. To maximize the attack's impact, items in the filler set, IF, are assigned ratings to the minimum, i.e., r_{min}=1, thus maximize the variations between the item similarities.

Nuke Attack Models	
Reverse Bandwagon Attack	● In this attack, minimum ratings are assigned to the items in a target set and the selected items. ● It increases the possibility that the system would generate low predicted ratings for those items. ● This attack has less impact on user-based systems. But, it is a very effective attack against item-based RS [50].
Love-Hate Attack	● Very simple to mount because it requires no knowledge. ● The maximum rating is assigned to filler items, i.e., rmax. Minimum ratings are provided to the target items. It can be used as a push attack by switching rmax and rmin. ● Very little knowledge is required to mount an attack. ● This attack is not very effective when used as a push attack, but it is one of the most effective nuke attacks in the user-based CF systems [50].

IV. MATERIALS & METHODS

In this section, the web recommender system "SRSPIA" architecture is described component-wise. Each component is explained carefully, which is first collecting the data of user's ratings and then the collection of all those biased ratings after adding the security component in it.

Here, we have chosen "movie" ratings just because our focus is on entertainment item and movie is the best items on the digital platform to interact with so many users of different age groups where we can easily find the user opinions or preference(s) more effectively.

A. SRSPIA

On the Home page of the rating system SRSPIA shown in fig. 1, the movies were added differently, and all the movies are placed row-wise, having two different columns in front of it. The first column named 'Rating' will show the overall rating ratings given to the particular movie, and the other column named 'Your Rating' will show the rating which is given to the movie, and if the users haven't login to their account, then it will show the message as 'Please login to see your ratings' and then clicking on that 'login' button, the page will be redirected to the User Login Page.

Fig. 1 Recommender System: SRSPIA Home page

In the next component, User, Sign-up/Login Page, has been created. The user has to provide his First name, Last name, Email address for verification purposes, Password for creating this new account, Security question with an answer; after submitting these details, the page will be forwarded to the verification page for further verification.

In the Verification Page, the user will receive an email after submitting their sign-up details containing a One-Time Password, which is to be submitted on this Verification Page for creating an account.

In the Login Page, the user who already has an account can log in just by providing their email address and password of the account. After this process, this page will be redirected to the main page of the SRSPIA where all the movies are given, and then the user has to click on the movie icon or movie name for rating and reviewing the particular movie and then the rating and reviewing page of the same will be opened.

On the Rating and Reviewing Page, the trailer of the particular movie is provided, and the Review section is given where all the reviews of all the users are shown. Below this Review Section, there is the Rate scale in which user can rate the movie on a scale of 1-5, and after this, they can write their review in the Review section; after doing all this, user can submit the details, and all the Rating and SRSPIA given by user will be saved in a database.

If a user enters an incorrect answer for a security question, then it will show a pop-up and automatically logout from SRSPIA, which is shown in fig. 2 below.

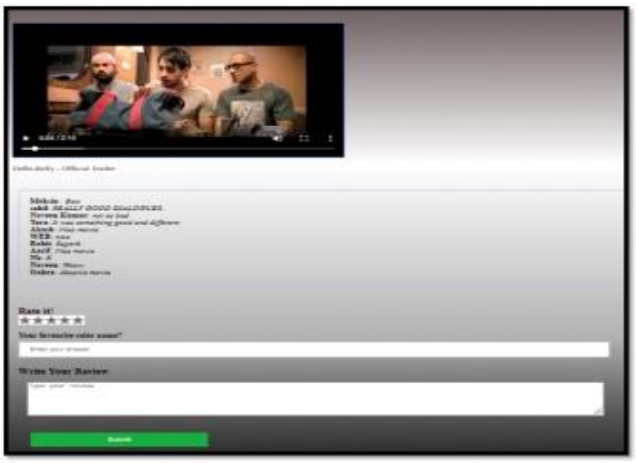

Fig. 2 Verification process for security credentials

After the study of various attack types in the previous section, here average attack is taken because of its efficient detection. Generally, In Genuine-fake attacks, the ratings are considered as high to low and low to high such as 1 is changed to 5 and vice versa in the rating scale of 1 to 5 to any item. In average attack type, the major component is the mean of the ratings available on the system is produced for the target item.

By focusing on the online movie watchers, it is looking towards people having similar opinion upon the movie or product, so that similarity between others of the same like

or dislike could make some good pattern for future recommendations in a better way. As far as certain conditions that should be met, like, User(s) at least have watched the movie or product before giving their ratings and also on at least rated on an online portal system before. An algorithm is explained as follow:

Algorithm. Attack profile detection

Input: Ratings from **U**

Output: Secure ratings for recommendations

1. *For* all users' ($\mathbf{U_i}$) ratings on Movie items ($\mathbf{M_j}$) on SRSPIA in Phase 1 and Phase 2;
2. *if* there are no ratings of the user (**U**) on any movie items (**M**)

then the user is placed automatically in the dataset "**D1**" as a new user.

3. *else if* there is already ratings of the user are present in the SRSPIA

then the user (U) is categorized as an actual user or a genuine user.

4. *else if* there is already ratings present of the user and there are differences in the opinion in their rating pattern on the same items

then the user called a malicious or fake user and the profile termed as an attack profile stored in SRSPIA dataset "**D2**", attack patterns can be fall under these cases:

Case 1: Users rated **R** {1,2} -> {5}: fake

Case 2: Users rated **R** {3,4,5} -> {1}: fake

5. *else* store the ratings in the dataset "**D1**".
6. end *For*
7. Return rating **R** of **U**.

As some of the users are intentionally using the session of others, if the user is by mistake open their account, then other users might change the data, which can change the recommendation patterns and destroyed the meaning of other preferences. After collecting the data in phase one, which is the pre-experiment phase, collection of the data is done without security version then applying these data to classification techniques taken in proposed model and collecting the results obtained from this phase, based on that a comparison has performed with the results obtained through the generated dataset in dual mode. In phase two, which is the post-experiment phase, collection of the data is done through adding security components as secret questions in the system, and then the classification algorithms are applied to it, which were taken in the proposed model. After obtaining the results, a conclusion is made by comparing different datasets. Finally, the results have been used for better similarity predictions between different users, looking into the error rate. To collect the data, we have developed a movie rating system [4] and then send the link of the rating system portal to the different users and ask them to rate and review the movies on the system. Here, to better comparisons with other sets of data, we have taken [41] dataset of around 2375 ratings named as "D3" and evaluate the performance with the same measures. The phases are described in the next subsections.

B. Phase 1: Data collection through rating system without security version through SRSPIA

Here on the rating system portal, the discrimination cannot be seen between the genuine user and fake user, i.e., whether these ratings are submitted by an actual user of that account or submitted by someone other by that account because there is no security while submitting the rating for a movie. The dataset collection consists of 2375 ratings as "D1".

C. Phase 2: Data collection through rating system with security version through SRSPIA

Here the rating system portal has been introduced by adding the security component. The component used here, as the user has to select a security question and an answer for it during the signup process. When submitting a rating for any movie, users have to answer this question each time. In this phase, a prompt is generated for the user to allow access to their account to another user without telling the answer to their security question. If a user submits a rating for a movie with the correct answer, it will be submitted successfully to the database in his/her account; otherwise, it will pop a message "Incorrect answer." If the user submits an incorrect answer, the data is submitted to another database. At this rating portal, it can be easily discriminate between the actual user and the fake user because the rating submitted with the correct and incorrect responses are stored in two different databases. There are 2700 ratings as "D2".

Data collection through wrong credentials with security version is taken from the rating system which contains the wrong answered credentials data which is submitted by the malicious raters and has been logged out by the session with their wrong response of rating for further used to find the pattern of such malicious raters on other rating systems.

D. Phase 3: Model

We have broken down data in pre-processing as per the following Attribute selection and feature selection. Attribute selection in ML selects the relevant features from the set for model construction. It gets significantly increases accuracy. Increasing classification accuracy and identifying relevant attributes is done through feature selection. Here proposed model is comprised of different supervised classifiers such as BayesNet, Naive Bayes, J48, Random Forest, and LibSVM categorized as Mode 1 (see table 4). The other variants are the ensemble of these classifiers together called Mode 2:

In Mode 1, the different classifiers are taken individually and applied with the different sets of datasets. Here, five classifiers are taken, which are described as (see table 4):

Table 4. Summary of different ML classifiers

Classifiers	Description
BayesNet	There are various search algorithms, and quality measures are used in Bayes Network. Here, the base class, which provides a Bayes Network classifier that includes the network structure, conditional probability distributions, etc., as a data structure and facilities the learning algorithms. This type of network has used the statistical that represents a set of variables with conditional dependencies with the help of a DAG. It also predicts the likelihood of an event occurred with several possible known causes with contributing factors. For instance, it could represent the probabilistic relationships between symptoms and diseases. Given symptoms, the network can be used to compute the probabilities of the presence of various diseases.
Naive Bayes	In machine learning, the Naïve Bayes classifier is based on the Bayes rule of conditional probability with independence assumptions. In a given dataset, the Naive Bayes calculates a set of probabilities by combinations of values, and also it has a fast decision-making process. The Naive Bayesian classifier [42] is used in supervised induction tasks that provide a simple approach, clear semantics & representation, and learning probabilistic knowledge. The main goal of this classifier is to accurately predict the class of test instances with the inclusion of training instances. We can say that it is a specialized form of BayesNet; it relies on two important assumptions. First, it assumes that the predictive attributes are conditionally independent given the class, and secondly, no hidden or latent attributes influence the prediction process.
J48	The Decision Tree is a predictive learning method that could decide the target value from several attribute values in the data to the new sample arrived in the set. The DT-J48 is used to generate a tree from the given dataset. The Tree Algorithm is used to find the way for the attributes-vector to behave on the number of instances. Based on training instances, the classes for the new instances are found [43]. It also predicts the target variable by making rules. This critical classification can be easily resolved. J48 is an extension of ID3, which is allowed to detect the missing values, pruning the tree, continuous attributes with the range and derivations of rule, etc.
Random Forest	Set creation of decision trees from a randomly selected subset of the training set is done using a random forest classifier. It uses the aggregates of votes from distinct selection trees and then decides the class of the test object. More than one algorithm could be combined using a random forest classifier. It is a promising classifier in [44] called RF, suggested by Breiman, which is used in many remote-sensing applications. This can operate on large databases efficiently and handles lots of variables simultaneously. RF generates the better estimation in classification; it generates generalized errors under the estimation. To compute the proximities between pairs of cases, RF can easily locate the outliers. We can say that it is robust to noise and computationally lighter than the ensemble methods.
LibSVM	SVM is a supervised learning algorithm that can be used for classification and regression problems. It has the kernel that transforms non-linear high-dimensional data to a low dimension with the optimal boundary for the classification. The Library of Support Vector Machines (SVMs) is popular for classification, regression, and other learning tasks. LibSVM [45] helps in two ways: firstly, training a data set to obtain a model and secondly, using the model to predict information of a testing data set. It provides the special settings for an unbalanced dataset.

In Mode 2, an Ensemble of classifiers is taken and applied with the different sets of datasets.

The collection of more than one classifier together becomes Ensemble learning algorithms (e.g., bagging and boosting). They are more accurate and robust to noise than single classifiers [46]. The main idea behind classifier ensembles is based on that a set of classifiers does perform better classifications than an individual classifier does. The framework of the proposed methodology is represented in fig. 3.

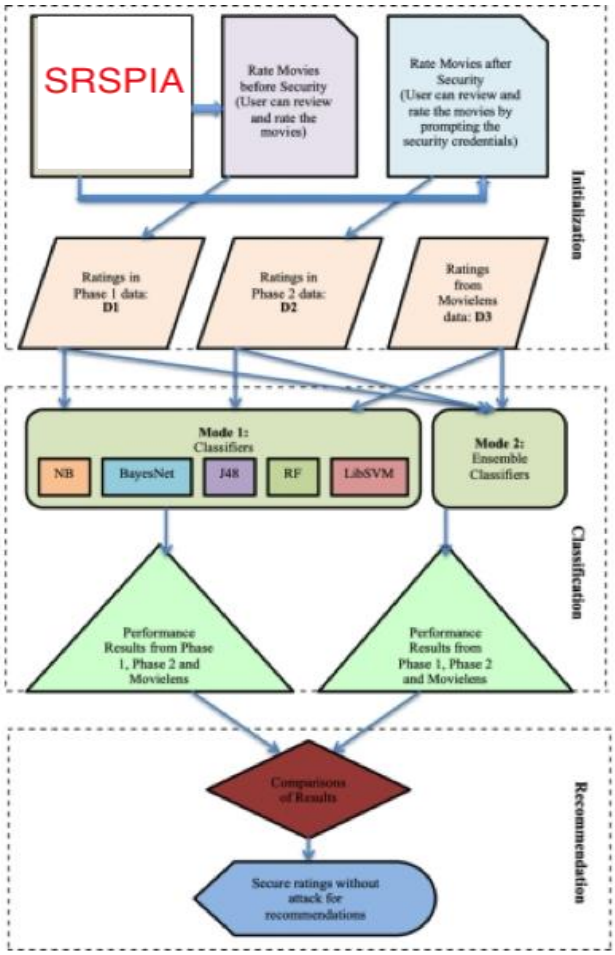

Fig. 3 Architecture of Proposed Methodology

V. DATASET

This section is designed for the experiments that are performed on the proposed model on different datasets and analyzed the results obtained from the system. This section discussed as follows:

Our dataset is the collection of Userid, Movie, and Rating matrix that has a collection of values of user opinions through ratings on different movies. See table 5, which represents the user as U and the movie as M with the ratings below.

Table 5. Dataset Structure

	M_1	M_2	M_3	M_4	M_N
U_1	1	2	4	1	5
U_2	4	1	5	2	3
U_3	5	4	1	1	2
U_N	1	3	2	5	5

The experiment has performed as follows: Hardware: CPU is Intel Core i7 processors, Windows 8 with 16 GB RAM. Software: All tests are performed on Weka 3.8. There are 200 users, and 20 movie items on the rating system are used, which are to be rated and reviewed by the users shown in table 6.

Table 6. Experimental Dataset Movie Rating: SRSPIA

Dataset	Total Ratings (Phase 1) D1	Total Ratings (Phase 2) D2	Total Ratings D3
No. of Rating Instances	2375	2375	2375
No. of Users	200	200	200
No. of Movie Items	20	20	20

For this experiment, an average attack is taken of at most 30% of attack data is the influence of the total actual data because the average attack used to predict the user's preference quite in an easy manner from their profiles and the previous research studies [5], it is always best to detect the attack within the range between 5% to 25% of filler size.

VI. RESULTS

The performance can be assessed through different measures; Here, MAE, precision, recall, and f-measure are used.

A. MAE

It measures the model error and basic criteria for classification. It is defined as in Eq.(1):

$$\text{MAE} = \sum_{i=1}^{n} \frac{predicted\ values_i - observed\ values}{n} \quad (1)$$

where n is the total number of observations.

B. Precision

The amount of relevancy found in the set of recommendations is defined as in Eq.(2):

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

C. Recall

The amount of relevancy found successfully is defined as in Eq.(3):

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

D. F-measure

The harmonic mean of precision and recall is done under f-measure that is defined as in Eq.(4):

$$\text{F} - \text{Measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

E. Case example

See fig. 4, and a graph is showing the ratings of both actual and fake users by the same logged account on the movies, and using this, differences are predicted between the rating of actual and fake users. The dataset instances for these movies have been taken from SRSPIA for this example.

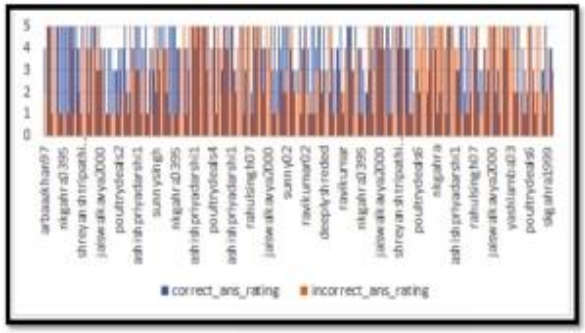

Fig. 4 Rating Submitted by Actual user vs. Fake user

For instance, a movie sample is taken for "X" in which the users rating pattern is detected as actual and fake that is represented in fig. 5 below:

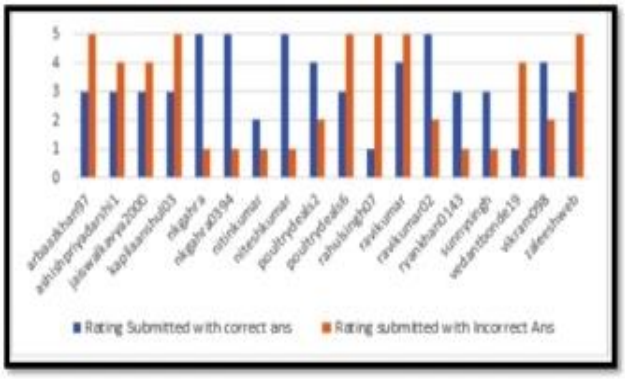

Fig. 5 Actual user vs. fake user rating for a movie: "X."

From the figure above, it can be seen that there is a possibility that a fake user can submit the same rating as submitted by the actual use for a movie. Detecting such types of profiles, these classification algorithms are used, and the highest accuracy is obtained. The Precision and Mean Absolute Error is calculated and compare to the results to obtain the classifier with the highest accuracy. The accuracy, precision, F-measure and mean absolute error for all classifiers are represented in table 7.

Table 7. Experimental Results on the different dataset with proposed models

Classifier		MAE	Precision	Recall	F-measure
LibSVM	D1	0.157	0.844	0.842	0.839
	D2	0.129	0.874	0.871	0.866
	D3	0.157	0.844	0.842	0.839
BayesNet	D1	0.215	0.868	0.867	0.864
	D2	0.216	0.882	0.880	0.877
	D3	0.215	0.868	0.867	0.864
J48	D1	0.243	0.860	0.854	0.850
	D2	0.224	0.874	0.871	0.866
	D3	0.243	0.860	0.854	0.850
Naïve Bayes	D1	0.213	0.840	0.841	0.839
	D2	0.207	0.846	0.847	0.846
	D3	0.213	0.840	0.841	0.839
Random Forest	D1	0.356	0.819	0.821	0.820
	D2	0.332	0.826	0.828	0.826
	D3	0.356	0.819	0.821	0.820
Ensemble Classifier	D1	0.144	0.858	0.856	0.853
	D2	**0.128**	**0.872**	**0.871**	**0.867**
	D3	0.144	0.858	0.856	0.853

From the table above, it has been noticed that the classifiers ensemble gives better results with secure data version (D2) on various performance measures than the without security version (D1). It can be easily seen in fig. 6 that the Ensemble of classifiers has the lowest MAE 0.128 than the without secure version as 0.144. Therefore, any bad profile in the RS database can be filtered-out by this method before the recommendation.

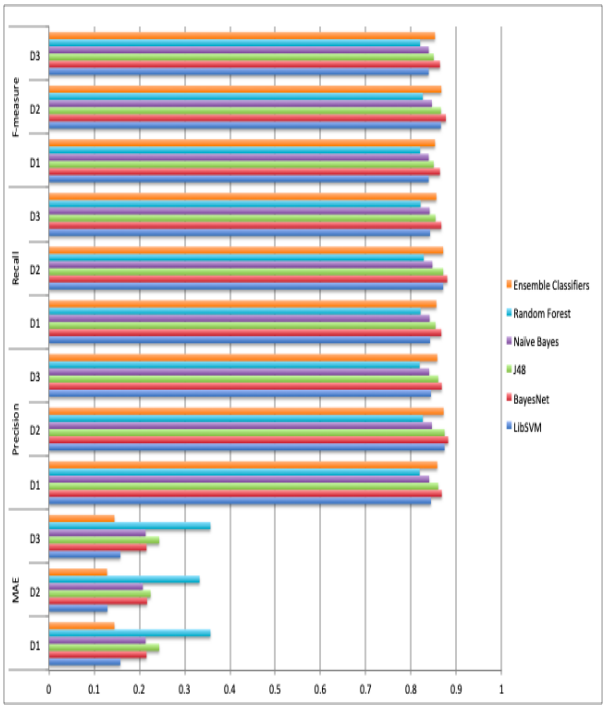

Fig. 6 Performance of Classifiers on different measures on different datasets

VII. CONCLUSIONS

It is concluded that building a secure system is helps to prevent the misuse of someone's account and information. SRSPIA is amended with security in such a way that if the user is giving the wrong answer to the security question which was asked during the time of sign up and was added in the previous phase, the user will be log out automatically from the session, which will simultaneously redirect it to the login page and also will send an alert email to the linked email account that maybe someone is trying to access their account. It has concluded that the integration of security is helpful in profiling (good or bad) gives minimum error than the without secure version on web recommender system.

In the future, we want to extend our work in a way to integrate second-tier security like the face recognition feature for posting the ratings on products or items on RS. This is further applied with deep learning for the better detection of shilling attacks. This may be further applied with other recommender systems such as Netflix and Amazon prime to finding out the performance accuracy.

ACKNOWLEDGMENT

I would like to acknowledge with warm thanks to my supervisor, who has guided me throughout the stages of writing this paper. His remarks are very helpful, valuable, and needful of writing this paper in changer in the field of secure recommendation system.

REFERENCES

[1] Burke, R., Mobasher, B., Williams, C., & Bhaumik, R. (2006, August). Classification features for attack detection in collaborative recommender systems. In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 542-547).

[2] Williams, C. A., Mobasher, B., Burke, R., & Bhaumik, R. (2006, August). Detecting profile injection attacks in collaborative filtering: a classification-based approach. In International Workshop on Knowledge Discovery on the Web (pp. 167-186). Springer, Berlin, Heidelberg.

[3] Lam, S. K., & Riedl, J. (2004, May). Shilling recommender systems for fun and profit. In Proceedings of the 13th international conference on World Wide Web (pp. 393-402).

[4] Secure system for movie review and rating: SRSPIA. (2019). https://projectatcic.online/ Accessed 12 October 2019.

[5] Verma, A. K., & Dixit, V. S. (2019). A Comparative Evaluation of Profile Injection Attacks. In Advances in Data and Information Sciences (pp. 43-52). Springer, Singapore.

[6] Miyahara, K., & Pazzani, M.J. (2002). Improvement of Collaborative Filtering with the Simple Bayesian Classifier 1.

[7] Kumari, T., & Bedi, P. (2017). A comprehensive study of shilling attacks in recommender systems. International Journal of Computer Science Issues (IJCSI), 14(4), 44.

[8] Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. Knowledge-based systems, 46, 109-132.

[9] Wu, M. L., Chang, C. H., & Liu, R. Z. (2014). Integrating content-based filtering with collaborative filtering using co-clustering with augmented matrices. Expert Systems with Applications, 41(6), 2754-2761.

[10] Al-Shamri, M. Y. H. (2016). User profiling approaches for demographic recommender systems. Knowledge-Based Systems, 100, 175-187.

[11] Yu, J., Gao, M., Rong, W., Song, Y., & Xiong, Q. (2017). A social recommender based on factorization and distance metric learning. IEEE Access, 5, 21557-21566.

[12] Yang, Z., Wu, B., Zheng, K., Wang, X., & Lei, L. (2016). A survey of collaborative filtering-based recommender systems for mobile internet applications. IEEE Access, 4, 3273-3287.

[13] Bobadilla, J., Hernando, A., Ortega, F., & Gutiérrez, A. (2012). Collaborative filtering based on significances. Information Sciences, 185(1), 1-17.

[14] Paradarami, T. K., Bastian, N. D., & Wightman, J. L. (2017). A hybrid recommender system using artificial neural networks. Expert Systems with Applications, 83, 300-313.

[15] Wang, Y., Qian, L., Li, F., & Zhang, L. (2018). A comparative study on shilling detection methods for trustworthy recommendations. Journal of Systems Science and Systems Engineering, 27(4), 458-478.

[16] Chirita, P. A., Nejdl, W., & Zamfir, C. (2005, November). Preventing shilling attacks in online recommender systems. In Proceedings of the 7th annual ACM international workshop on Web information and data management (pp. 67-74).

[17] Burke, R., Mobasher, B., Williams, C., & Bhaumik, R. (2006, August). Classification features for attack detection in collaborative recommender systems. In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 542-547).

[18] Zhang, S., Ouyang, Y., Ford, J., & Makedon, F. (2006, August). Analysis of a low-dimensional linear model under recommendation attacks. In Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval (pp. 517-524).

[19] Li, X., Gao, M., Rong, W., Xiong, Q., & Wen, J. (2016, June). Shilling attacks analysis in collaborative filtering based web service recommendation systems. In 2016 IEEE International Conference on Web Services (ICWS) (pp. 538-545). IEEE.

[20] Cao, J., Wu, Z., Mao, B., & Zhang, Y. (2013). Shilling attack detection utilizing the semi-supervised learning method for collaborative recommender system. World Wide Web, 16(5-6), 729-748.

[21] CSPIT, C. (2016). A Novel Supervised Approach to Detection of Shilling Attack in Collaborative Filtering Based Recommendation System. International Journal of Computer Science and Information Security (IJCSIS), 14(4).

[22] Hernando, A., Bobadilla, J., & Ortega, F. (2016). A non-negative matrix factorization for collaborative filtering recommender systems based on a Bayesian probabilistic model. Knowledge-Based Systems, 97, 188-202.

[23] Mehta, B., & Nejdl, W. (2009). Unsupervised strategies for shilling detection and robust collaborative filtering. User Modeling and User-Adapted Interaction, 19(1-2), 65-97.

[24] Zhang, F., Deng, Z. J., He, Z. M., Lin, X. C., & Sun, L. L. (2018, July). Detection of shilling attack in collaborative filtering recommender system by PCA and data complexity. In 2018 International Conference on Machine Learning and Cybernetics (ICMLC) (Vol. 2, pp. 673-678). IEEE.

[25] He, W., Xu, G., Wang, Y., Wu, Z., Bu, Z., Cao, J., & Yang, D. (2016). Discovering shilling groups in a real e-commerce platform. Online Information Review.

[26] Lee, J. S., & Zhu, D. (2012). Shilling attack detection—a new approach for a trustworthy recommender system. INFORMS Journal on Computing, 24(1), 117-131.

[27] Zhang, F., & Zhou, Q. (2014). HHT–SVM: An online method for detecting profile injection attacks in collaborative recommender systems. Knowledge-Based Systems, 65, 96-105.

[28] Hao, Y., & Zhang, F. (2018). Detecting shilling profiles in collaborative recommender systems via multidimensional profile temporal features. IET Information Security, 12(4), 362-374.

[29] Dhimmar, J. H., & Chauhan, R. (2015). An accuracy improvement of detection of profile-injection attacks in recommender systems using outlier analysis. International Journal of Computer Applications, 122(10).

[30] Chirita, P. A., Nejdl, W., & Zamfir, C. (2005, November). Preventing shilling attacks in online recommender systems. In Proceedings of the 7th annual ACM international workshop on Web information and data management (pp. 67-74).

[31] Su, X. F., Zeng, H. J., & Chen, Z. (2005, May). Finding group shilling in the recommendation system. In Special interest tracks and posters of the 14th international conference on World Wide Web (pp. 960-961).

[32] Bryan, K., O'Mahony, M., & Cunningham, P. (2008, October). Unsupervised retrieval of attack profiles in collaborative recommender systems. In Proceedings of the 2008 ACM conference on Recommender systems (pp. 155-162).

[33] Lee, J. S., & Zhu, D. (2012). Shilling attack detection—a new approach for a trustworthy recommender system. INFORMS Journal on Computing, 24(1), 117-131.

[34] Mehta, B., Hofmann, T., & Fankhauser, P. (2007, January). Lies and propaganda: detecting spam users in collaborative filtering. In Proceedings of the 12th international conference on Intelligent user interfaces (pp. 14-21).

[35] Mehta, B., & Nejdl, W. (2009). Unsupervised strategies for shilling detection and robust collaborative filtering. User Modeling and User-Adapted Interaction, 19(1-2), 65-97.

[36] Williams, C. A., Mobasher, B., & Burke, R. (2007). Defending recommender systems: detection of profile injection attacks. Service-Oriented Computing and Applications, 1(3), 157-170.

[37] Burke, R., Mobasher, B., Williams, C., & Bhaumik, R. (2006, August). Classification features for attack detection in collaborative recommender systems. In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 542-547).

[38] Williams, C. A., Mobasher, B., Burke, R., & Bhaumik, R. (2006, August). Detecting profile injection attacks in collaborative filtering: a classification-based approach. In International Workshop on Knowledge Discovery on the Web (pp. 167-186). Springer, Berlin, Heidelberg.

[39] He, F., Wang, X., & Liu, B. (2010, August). Attack detection by rough set theory in the recommendation system. In 2010 IEEE International Conference on Granular Computing (pp. 692-695). IEEE.

[40] Mobasher, B., Burke, R., Bhaumik, R., & Williams, C. (2005, August). Effective attack models for shilling item-based collaborative filtering systems. In Proceedings of the WebKDD Workshop (pp. 13-23). Citeseer.

[41] MovieLens Dataset 1M. (2019). https://grouplens.org/datasets/movielens/1m/ Accessed 21 December 2019.
[42] John, G. H., & Langley, P. (1995). Estimating Continuous Distributions in Bayesian Classifiers. In Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence.
[43] Buntine, W. L. (1994). Operations for learning with graphical models. Journal of artificial intelligence research, 2, 159-225.
[44] Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.
[45] Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. ACM transactions on intelligent systems and technology (TIST), 2(3), 1-27.
[46] Breiman, L. (1996). Bagging predictors. Machine learning, 24(2), 123-140.
[47] Weka 3: Machine Learning Software in Java. (2019). "https://www.cs.waikato.ac.nz/~ml/weka/" Accessed 11 March 2019.
[48] Chirita, P. A., Nejdl, W., & Zamfir, C. (2005, November). Preventing shilling attacks in online recommender systems. In Proceedings of the 7th annual ACM international workshop on Web information and data management (pp. 67-74).
[49] Alonso, S., Bobadilla, J., Ortega, F., & Moya, R. (2019). Robust model-based reliability approach to tackle shilling attacks in collaborative filtering recommender systems. IEEE Access, 7, 41782-41798.
[50] Kaur, P., & Goel, S. (2016, August). Shilling attack models in recommender system. In 2016 International Conference on Inventive Computation Technologies (ICICT) (Vol. 2, pp. 1-5). IEEE.