View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Available online at www.sciencedirect.com
ScienceDirect
Procedia Technology 10 (2013) 963 – 969
Procedia Technology

*International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) 2013

Detection of Profile-injection attacks in Recommender Systems using Outlier Analysis

Parthasarathi Chakraborty[a*], Sunil Karforma[b]

[a]Assistant Professor, University Institute of Technology, University of Burdwan, Burdwan, 713104, India
[b]Associate Professor, Department of Comp. Science, University of Burdwan, Burdwan, 713104, India

Abstract

E-Commerce recommender systems are vulnerable to different types of profile-injection attacks where a number of fake user profiles are inserted into the system to influence the recommendations made to the users. In this paper, we have proposed three strategies of detecting such attacks with the help of outlier analysis. In all these strategies, the attack-profiles are considered as outliers in the user rating dataset. Firstly, we have used Partition around Medoid (PAM) clustering algorithm in detecting the attack-profiles. An incremental version of the PAM algorithm has been applied and tested for evaluating the performance of the system in identifying the attack profiles when they come into the system. Experiments show that though PAM is able to detect attack profiles with larger number of filler items very well, a percentage of attack profiles with smaller number of filler items is not included in outlier clusters-they are included in large clusters. Secondly, we have applied a PAM-based outlier detection algorithm to find these attack profiles in large clusters. Finally, an angle based outlier detection strategy is used for finding attack profiles in the database under attack.

© 2013 The Authors. Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
Selection and peer-review under responsibility of the University of Kalyani, Department of Computer Science & Engineering

Keywords-PAM; Outlier-detection; Recommender System; profile-injection attack; Attack-profile; angel-based outlier detection;

1. Introduction

Many web sites attempt to help users by incorporating a recommender system that provides users with a list of items and/or web pages that are likely to interest them. Content-based filtering and collaborative filtering are usually

[*] Corresponding author. Tel.: +919432132772.
E-mail address: psc755@gmail.com.

2212-0173 © 2013 The Authors. Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
Selection and peer-review under responsibility of the University of Kalyani, Department of Computer Science & Engineering
doi:10.1016/j.protcy.2013.12.444

applied to predict these recommendations. Among these two, Collaborative filtering is the most common approach for designing e-commerce recommender systems. It works by building a database of items with users' opinions on them. Then a specific user is matched against this database in order to find her neighbors, those with whom he or she shares similar tastes. As the system is open to user input, chance of attack on it is always there. The researchers have discussed different types of attacks. The ultimate target of all type of profile injection attacks is either to push a product (or a group of products) or to nuke a product (or a group of products). In case of Random Attack [1] a pre-specified rating is assigned to the target item and random ratings are assigned to the filler items whereas in average attacks [1], rating of each filler item corresponds to the mean rating for that item. Some additional attack types have been specified by Burke et. al.[2] namely Bandwagon Attack, Segment attack, Reverse Bandwagon Attack and Love/Hate Attack. The last one is a very simple attack and requires no system knowledge where the attack profile consists of minimum/ maximum rating value for target items and maximum/ minimum rating value for filler items for nuke/push attack.

On the other hand, in literature, the researchers have proposed several outlier detection techniques. They can be broadly categorized into different groups namely distance based approach, density based approach, clustering based approach and depth based approach. In clustering based approach, the clusters having small number of members are considered as the clusters consisting of outliers assuming that outliers are a small percentage of the total data. The main advantage of this approach over the other approaches is that the outlier detection is totally unsupervised.

In clustering-based anomaly detection techniques it is assumed that the percentage of data, which is inserted during attack event, is very small compared to the total data. At the same time, the nature of the attack data also differs from the data without attack. Based on these two assumptions, the data members of clusters with small sizes are considered as outliers, which in turn, correspond to the attack data. As rightly mentioned in [3], the attack profiles are highly correlated and at the same time the number of attack profiles is very small compared to total number of genuine user profiles.

Keeping these two points in view, we have considered the problem of profile-injection attack detection as a problem of outlier detection in the user rating dataset and applied PAM clustering algorithm in detecting the outliers or injected attack profiles.

In our paper, after evaluating the performance of PAM clustering algorithm in detecting attack-profiles with different size of filler items, an incremental version of the PAM algorithm has been applied to test whether a new user profile which is going to be inserted into the system is an attack profile or not. Then we have applied a PAM-based outlier detection algorithm to find attack profiles in large clusters that are not identified by the PAM algorithm. Finally, an angle-based outlier detection strategy [16] is used for finding attack profiles in the attacked database.

2. Related Work

Detection of profile-injection attacks on recommender systems have been studied by many researchers. Supervised classification techniques have been used in [2] in order to distinguish attack profiles from genuine user profiles.

In their paper [4], authors used hierarchical clustering technique in detecting outliers. They have compared the performance of several hierarchical clustering algorithms in this regard. The authors of paper [5] have proposed a two-stage approach of outlier detection by using the concept of minimum spanning tree along with clustering. A Fuzzy-based clustering algorithm has been proposed by authors of paper [6] in detecting outliers in data. In both of the papers [7] and [8], PAM has been used as clustering algorithm. The authors of the first paper [7] have used a separation technique after applying PAM algorithm. In the second paper [8], the members of the small clusters generated by the clustering algorithm are identified as outliers.

The unsupervised methods have been used by many researchers in the field of network intrusion detection [9], [10]. In the area of recommender system also, a few works [11], [12] have also been reported in the literature where unsupervised methods have been used as a tool of attack detection. In their paper [11], authors proposed a Principal Component Analysis (PCA) based clustering algorithm for detecting attack profiles based on the assumption that attack profiles are very highly correlated with each other.

3. Our Approach

In our approach of attack detection, the user profiles, which have been injected into the system during attack event, are considered as outliers, which are, in turn, detected by clustering algorithms and then by a angle-based outlier detection strategy [16]. In this paper, for the purpose of outlier detection, we have applied PAM (Partition Around median) algorithm [14], one of the first k-medoid algorithms introduced in the literature.

3.1. PAM

In PAM algorithm, the most centrally located object (called medoid) in each cluster is considered as cluster center. As medoids are less influenced by outliers than means, PAM is more robust than k-means[13] in presence of outliers in the dataset. In PAM algorithm, an initial set of medoids is selected first. Each one of the selected medoids are then replaced iteratively by one of the non-selected medoids until sum of the distances of the data objects to their closest medoids is improved.

The algorithm is as follows-

1. Arbitrarily select k objects as medoid points out of n data points (n>k).
2. Repeat
3. Associate each remaining data object in the given data set to most similar medoid.
4. Randomly select a non-medoid object, $O_{random.}$
5. Compute the total cost, S of swapping medoid object O_j with $O_{random.}$
6. If $S < 0$ then swap O_j with O_{random} to form the new set of k-medoid objects
7. until no change.

We have taken Euclidean distance measure to define similarity between two data objects. After applying the PAM algorithm on the user rating profiles, we identify those user profiles as attack profiles that belong to the small clusters. Following the definition of small cluster given in [4] we identify outliers as the data objects that belong to a cluster having size lesser than half the average number of points in the k clusters.

3.2. Cluster Updation

Whenever a new rating data comes into the system we need to reconstruct the existing clusters which is very costly. We apply a simple cluster-updating algorithm proposed by [15] based on the previous algorithm to dynamically update the clusters. After partitioning the existing n data points using PAM algorithm, assume, the total distance which is the summation of all distances from each data point to its closest medoid is T. This information is also stored in the database and used by our cluster updating algorithm which is as follows-

1. For the new rating data object D′, find the closest medoid, say M′ with distance Dist′.
2. Compute total distance T′ if M′ is replaced with D′.
3. if $T' < T + Dist'$ then swap M′ with D′ to form the new set of medoids.

3.3. Identifying Attack-profiles in Larger Clusters

From our experiment we observe that most of the attackers are detected in outlier clusters when the percentage of filler items of the attack profiles is high. When the filler percentage is low, a large fraction of attackers are not detected and belongs to the large clusters. To identify those attackers who belongs to large clusters we have used a PAM-based outlier detection algorithm developed by the authors of paper [8].

The algorithm proposed in paper [8] first computes absolute difference between each data point of a large cluster and the medoid of that cluster. If the difference crosses a threshold limit for a data point then that data point is detected as outlier. The threshold value is computed as 1.5 times of the average of differences between each data point of a large cluster and the medoid of that cluster (ADMP as they name). The algorithm can be described as-

```
1. Apply PAM to detect outlier clusters.
2. For each cluster not defined as outlier cluster-
    a. For each data point in that cluster-
        i. Calculate ADMP, the absolute difference between the data point and the medoid of that
           cluster.
        ii. If it crosses Threshold T, declare that data point as outlier.
```

3.4. Identifying Attack-profiles using angle-based outlier detection strategies

The algorithm angle-based outlier detection (ABOD) [16] strategy considers the direction of distance vectors along with the distance between points in a vector space. It calculates angle-based outlier factor, ABOF for each point in the dataset and based on the value of ABOF, the outlier ness of a point is determined. The idea is as follows.

For any point in a data cluster, if the angle between the difference vectors of any other two points (of that cluster) is measured, it can be seen that the angle varies widely. The variance of these angles becomes smaller for data points which lies in the border of the cluster and it will be even more smaller for the true outliers lying far apart from the data cluster. The angle-based outlier factor, ABOF of data point A is calculated as

$$ABOF(\vec{A}) = VAR_{\vec{B},\vec{C}\in D}\left(\frac{\left(\overline{AB},\overline{AC}\right)}{\left\|\overline{AB}\right\|^{2}\bullet\left\|\overline{AC}\right\|^{2}}\right)$$

where D is the database and $\overline{AB}$ is the difference vectors between $\vec{B}$ and $\vec{A}$ and $\overline{AC}$ is the difference vectors between C and $\vec{A}$ respectively. The scalar product of the difference vectors AB and AC is normalized by the quadratic product of the length of the difference vectors, AB and AC , i.e. the angle is weighted less if the corresponding points are far from the query point. By this weighting factor, the distance influences the value after all, but only to a minor part.

4. Result

4.1. Dataset

In our experiment we have used MovieLens dataset (movielens.umn.edu). The data set used contained 100,000 ratings from 943 users and 1682 movies (items), with each user rated at least 20 items. The item sparsity is easily computed as 0.9369. The ratings in the MovieLens dataset are integers ranging from 1 to 5.

4.2. Experiments

In our experiment, the attack size (in terms of false user profiles) has been set to one percent of the total number of users in the dataset. The attacker profiles have been built following Random attack model, where the target item is given the maximum rating value and the filler items are given random rating values. In comparing performance of PAM algorithm in attack-profile detection, we have used different percentages of filler item in the attack profiles. As per definition of small cluster given in [4] clusters having less than ninety five members are considered as clusters having outlier or attack profiles for MovieLens dataset that we have used. In our experiment, True Positive is considered as number of attackers present in small clusters, False Positive is considered as number of non-attackers present in small clusters and False Negative is considered as number of attackers not present in small clusters.

Table 1 shows that when percentage of filler items is 70%, the performance of PAM algorithm in detecting the attack profiles is 100% i.e. all the attack-profiles belong to the outlier clusters. When percentage of filler items is 60%, 68% (average of five runs) attack-profiles belongs to outlier clusters whereas other 40% attack-profiles belongs to the other large clusters. In case of attack-profiles with 40% percent of filler items, the average number of attack-profiles detected correctly reduces to 16%.

At the time of evaluation of the performance of the incremental version of the PAM algorithm [15], we first incorporate 0.5 percent attack profiles (i.e. five fake user profiles in our case) into the system and perform PAM algorithm. Then five other attack profiles (rest 0.5 percent of 1 percent attack profiles) are inserted into the system one at a time. We, then, check whether the algorithm is able to identify them as attack profile or not. When the size of the filler items in the attack profiles is 70%, the algorithm gives 100% accuracy. For filler items size 60%, the accuracy is 80% (the average of three runs) whereas for filler items size 60%, the accuracy is 33%.
The performance of the algorithm [8] for detecting attack-profiles in larger clusters has been shown in the last column of the Table 1. In all runs of the experiment for filler item percentage 40 and 60, all the attack-profiles in the large clusters are detected successfully and the number of genuine user-profiles detected as attack-profiles reduces significantly.

Table 1: Performance of PAM and PAM-based algorithm in detecting attack-profiles for different percentage of filler items in attack-profiles.

Percentage of Filler items in attack-profile	Run	% of attacker detected in outlier cluster by PAM	Size of clusters where other attackers belong	Number of False attacker detected in large clusters by PAM-based algorithm
70%	1	100	-	-
70%	2	100	-	-
70%	3	100	-	-
70%	4	100	-	-
70%	5	100	-	-
60%	1	100	-	-
60%	2	40	899	44
60%	3	30	809	26
60%	4	90	864	53
60%	5	80	861	51

40%	1	10	841	76
40%	2	0	915	96
40%	3	10	888	74
40%	4	50	877	62
40%	5	10	880	66

At the time of evaluation of the performance of the angle-based outlier detection strategy [16], we have taken a part of the MovieLens dataset (around 25%) and attack size was set to 4%. This is just due to reducing the running time of the algorithm. The results have been shown in Table 2. The ABOF value calculated for the attacker profiles are much lower than the ABOF value calculated for the non-attacker profiles. This indicates that the attacker profiles lies far apart from the other profiles in the high-dimensional space.

Table 2: Average ABOD value for the attackers and non-attackers for different filler percentages.

Filler Percentage	Average ABOF value of the Attackers	Average ABOF value of the Non-Attackers
40%	2.06E-06	3.96E-04
60%	1.12E-06	3.98E-04
70%	1.00E-06	3.97E-04

5. Conclusion

In our work, we have used Partition around Medoid (PAM) clustering algorithm in detecting the attack-profiles. An incremental version of the PAM algorithm has been applied and tested for evaluating the performance of the system in identifying the attack profiles when they come into the system. Experiments show that though PAM is able to detect attack profiles with larger number of filler items very well, a percentage of attack profiles with smaller number of filler items is not included in outlier clusters-they are included in large clusters. We have applied a PAM-based outlier detection algorithm to find these attack profiles in large clusters. Finally, an angle-based outlier detection strategy has been used for the same purpose.

References

[1] Lam, S. And Riedl, J. Shilling recommender systems for fun and profit. In Proceedings of the 13th International WWW Conference (New York, NY)(2004).

[2] Burke, R.,Mobasher, B.,Williams, C., And Bhaumik, R. 2006b. Detecting profile injection attacks in collaborative recommender systems. In Proceedings of the IEEE Joint Conference on Ecommerce Technology and Enterprise Computing, E-Commerce and E-Services (CEC/EEE 2006, Palo Alto, CA)(2006).

[3] Mehta Bhaskar, 2007. Unsupervised Shilling Detection for Collaborative Filtering. Association for the Advancement of Artificial Intelligence (www.aai.org), 2007.

[4] Loureiro,A., L. Torgo and C. Soares, 2004. Outlier Detection using Clustering Methods: a Data Cleaning Application, in Proceedings of KDNet Symposium on Knowledge-based Systems for the Public Sector. Bonn, Germany.

[5] John Peter.S., Department of computer science and research center St.Xavier‟s College, Palayamkottai, An Efficient Algorithm for Local Outlier Detection Using Minimum Spanning Tree, International Journal of Research and Reviews in Computer Science (IJRRCS), March 2011.

[6] Cutsem, B and I. Gath, 1993. Detection of Outliers and Robust Estimation using Fuzzy Clustering, Computational

Statistics & Data Analyses 15, pp. 47-61.

[7] Acuna E. and Rodriguez C., (2004), A Meta Analysis Study of Outlier Detection Methods in Classification, Technical paper, Department of Mathematics, University of Puerto Rico at Mayaguez, available at academic.uprm.edu/~eacuna/paperout.pdf. In proceedings IPSI 2004, Venice.

[8] Al- Zoubi, M. B., An Effective Clustering-Based Approach for Outlier Detection, European Journal of Scientific Research, Vol. 28, No. 2, 2009, pp. 310-316.

[9] Eskin,E., Arnold, A., Prerau,M., Portnoy, L., Stolfo, S.: A Geometric Framework for Unsupervised Anomaly Detection: Detecting Intrusions in Unlabeled Data. In: Proceedings of the Seventeenth International Conference on Machine Learning. Morgan Kaufmann Publichsers Inc, pp. 255-262. (2000)

[10] Portnoy, L., Eskin, E., Stolfo,S.: Intrusion detection with unlabeled data using clustering. In: Proceeding ACM Workshop on Data Mining Applied to Security. (2001)

[11] Mehta Bhaskar, 2007. Unsupervised Shilling Detection for Collaborative Filtering. Association for the Advancement of Artificial Intelligence (www.aai.org), 2007.

[12] M. O. K. Bryan and P. Cunningham, “Unsupervised retrieval of attack profiles in collaborative recommender systems,” in Technical Report, University College Dublin, 2008.

[13] MacQueen, J.,1967. Some methods for classification and analysis of multivariate observations. Proc. 5th Berkeley Symp. Math. Stat. and Prob, pp. 281-97.

[14] Kaufman, L. and P. Rousseeuw, 1990. Finding Groups in Data: An Introduction to Cluster Analysis. John Wiley & Sons.

[15] Chakraborty, P.S., “A Scalable Collaborative Filtering based Recommender System using Incremental Clustering”,In: Proceeding of IEEE International Advance Computing Conference, 2009.

[16] H.-P. Kriegel, M. S. hubert, and A. Zimek. Angle-based outlier detection in highdimensional data. In KDD ’08: Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, pages 444–452, New York, NY, USA, 2008. ACM.