Necmettin Erbakan Üniversitesi
Fen ve Mühendislik Bilimleri Dergisi
Journal of Science and Engineering (NEJSE)
DOI: 10.47112/neufmbd.2023.22

Cilt: 5 Sayı:2 Yıl: 2023 | Araştırma Makelesi / Research Article | E-ISSN: 2667- 7989

Şilin Ataklarının Tek Sınıflı Destek Vektör Makinaları ile Tespiti

Halil İbrahim AYAZ [1] **Zehra KAMIŞLI ÖZTÜRK** [2]

[1] Necmettin Erbakan University, Faculty of Engineering, Department of Industrial Engineering, Konya, Türkiye, hiayaz@erbakan.edu.tr *(Corresponding Author)*
[2] Eskişehir Technical University, Faculty of Engineering, Department of Industrial Engineering, Eskişehir, Türkiye, zkamisli@eskisehir.edu.tr

Makale Bilgileri

Makale Geçmişi
Geliş: 12.07.2023
Kabul: 19.09.2023
Yayın: 31.12.2023

Anahtar Kelimeler:
Öneri sistemleri,
Şilin atakları,
Tek sınıflı destek vektör makinaları.

ÖZ

Öneri sistemleri çeşitli çevrimiçi platformlarda hayati bir rol oynar ve kullanıcıların tercihlerini göz önünde bulundurarak yeni ürünler, hizmetler ve içerikler keşfetmelerine yardımcı olur. Bununla birlikte, bu sistemler, kötü niyetli kullanıcıların derecelendirmeleri yapay olarak şişirdiği veya söndürdüğü ve önyargılı önerilere yol açtığı şilin saldırıları yoluyla manipülasyona karşı savunmasızdır. Bu saldırıları araştırmanın, anlamanın ve hafifletmenin önemini vurgulamak çok önemlidir. Bu tür saldırıları tespit etmek, tavsiye sistemlerinin bütünlüğünü ve etkinliğini korumak için çok önemlidir. Literatürde, şilin saldırılarını tespit etmek için birçok çalışma sunulmuştur. En iyi bilinen kümeleme yöntemleri farklı saldırı modelleri için uyarlanmıştır. Bu makalede, şilin saldırılarını tespit etmek için gürbüz bir teknik olarak Tek Sınıflı Destek Vektör Makineleri kullanımını araştırıyoruz. Tek Sınıflı Destek Vektör Makinaları, öncelikle anomali tespiti ve aykırılık tespiti görevleri için tasarlanmış geleneksel Destek Vektör Makinelerinin özel bir çeşididir. Önerilen yöntemi doğrulamak için MovieLens100K veri kümesi kullanılmıştır. Sonuç olarak, farklı boyut ve doluluk oranlı saldırılar için hassasiyet ve geri çağırma değerleri verilmiştir.

Shilling Attack Detection with One Class Support Vector Machines

Article Info

Article History
Received: 12.07.2023
Accepted: 19.09.2023
Published: 31.12.2023

Keywords:
Recommender systems,
Shilling attacks,
OCSVM.

ABSTRACT

Recommender systems play a vital role in various online platforms, assisting users in discovering new products, services, and content considering their preferences. However, these systems are vulnerable to manipulation through shilling attacks, where malicious users artificially inflate or deflate ratings, leading to biased recommendations. It is crucial to emphasize the importance of researching, understanding, and mitigating these attacks. Detecting such attacks is crucial to maintaining the integrity and effectiveness of recommender systems. In the literature, lots of studies are presented to detect shilling attacks. The most well-known clustering methods are adapted for different attack models. This paper explores using One-Class Support Vector Machines (OCSVM) as a robust technique for detecting shilling attacks. One-Class SVMs are a specialized variant of the traditional Support Vector Machines, primarily designed for anomaly and novelty detection tasks. MovieLens100K dataset is used to validate the proposed method. As a result, precision and recall values are given for different attack and filler sizes.

Atıf/Citation: Ayaz, H.İ. & Kamışlı Öztürk, Z. (2023). Shilling attack detection with one class support vector machines, *Necmettin Erbakan University Journal of Science and Engineering*, *5*(2), 246-256. https://doi.org/10.47112/neufmbd.2023.22

INTRODUCTION

Recommender systems have become indispensable in our digital lives, aiding us in discovering new products, services, and content. These systems leverage user preferences and behaviors to provide personalized recommendations. Recommender systems have broad application areas in the real world, such as e-commerce [1,2], healthcare [3,4], energy [5,6], e-learning [7,8], and social networks [9].

Despite their growing popularity, recommender systems face a persistent threat known as shilling attacks [10–12]. These attacks undermine the trustworthiness and integrity of these systems, potentially leading to biased recommendations and compromised user experiences. It is essential to shed light on this issue and explore potential solutions to safeguard recommender systems from such attacks. Classification [13,14], clustering [15–17] and statistical [18] based approaches have been proposed to detect these attacks.

Shilling attacks entail inserting false or biased data into the recommender system's algorithms in an effort to change the results produced by the system. Attackers may make several fictitious user identities, dishonestly review products, or offer cunning commentary to influence recommendations in their favor. Attackers try to sway the algorithms' knowledge of user preferences and skew the recommendations by taking advantage of the system's reliance on historical user data.

Shilling attacks not only undermine user trust but also jeopardize the integrity and fairness of recommender systems. Attackers can alter the underlying algorithms and prejudice the recommendation process by injecting false or manipulated data. This can have a cascading effect, maintaining certain things or preferences while marginalizing others. As a result, diverse and useful recommendations may be repressed, resulting in a limiting of user experiences and limited exposure to new information.

An attack profile is defined by Bhaumik et al. [18] is given in Table 1. An attack profile consists of four main parts. Selected items (I_S) that have specific characteristics decided by attackers to make the attack harder to detect. For example, some attacks select popular items [19,20] whereas some attacks select unpopular items [21]. Filler items (I_F) are randomly chosen items to score with target item. Filler items score for making attack profiles difficult to detect. For a genuine user, it is not possible to scores all items in the database. Unrated items (I_O) aren't rate in attack profiles as with genuine users. Target item (I_T) is the item that rates want to be increased or decreased depending on the type of attacks. Some attack models don't include all the parts of this example.

Table 1. *An example attack profile*

I_S	I_F	I_O	I_T
Selected Items	Filler Items	Unrated Items	Target Item

Different attack models are generated to decrease detect possibility. The basic attack models are random and average is the simplest attack models in the literature. These attack models don't have selected items parts. In random attack model, filler items are rated by general mean [22]. On the other hand, filler items are rated by items' rate means in average attack model [23]. Target item is rated maximum or minimum rate considering attack type. For push attacks, target item is rated with maximum rate. for pull attacks, target item is rated with minimum rate. Bandwagon attacks are the extension of the random and average attacks. There are two types of bandwagon attacks in the literature; random and average bandwagon attacks. In both bandwagon attack types, popular items are selected and maximum rates are given with target item. Filler items are rated similar way in random and average attack.

Segment attacks are more effective in item-based recommender systems. In segment attacks, users associated with a specific product is selected and recommendations are proposed to these users [24]. Core profiles that generated from user database are used in probe attacks [21]. Target items rate is manipulated in used profiles. Aforementioned attack models are generally push attack models. They provide increasing popularity of target item [25]. To decrease popularity of target item, love/hate, reverse bandwagon and perfect knowledge attack models are proposed in the literature [21,26].

The literature analyzes shilling attack detection methods in three main areas: unsupervised, semi-

supervised, and supervised. Unsupervised methods are preferred over other methods. With different clustering methods, attacks can be detected without requiring any class labels for existing data. Since the ratings made by attack profiles will exhibit an abnormal distribution, they will be excluded from the clusters formed by genuine profiles.

The labels of data aren't supported in clustering algorithms [27]. Different clustering methods are used for shilling attack detection in the literature. In this study, suspicious user ratings of users are extracted by Markov chaining. Then, suspicious users are detected by hierarchical clustering of the obtained suspiciousness degrees [28]. In another study, the scoring behavior of users in the database is examined by extracting four detection features. Then, using PCA and k-means methods, a set of suspicious users is created. the gathering behavior of suspicious users is evaluated to distinguish between attackers and real users. Thus, the aim is to increase the accuracy rate [29].

The authors investigate the advantages of integrating the soft co-clustering algorithm with the user propensity similarity method, and they provide a soft co-clustering with propensity similarity model for detecting shilling attacks [30]. Another study investigates the attributes a profile must have to be an attacker. The user-product score matrix, the user-connection matrix, and the similarity between users are used to determine whether there are anomalies between the ratings. The k-means method detects attacker users with the obtained information [31]. In the study, a new attribute is proposed for attack users. Then three alternative detection model is proposed for shilling attacks [32].

Supervised methods need labeled data for classification to further data. In recommender systems, support vector machines are preferable for detecting attack users. In [33] uses a modified SVM model for attack detection. Gaussian Mixture Model is used as a machine-learning model to obtain high detection rates. In some scenarios, there are a few labeled data for detection. In this case, semi-supervised methods help to detection of shilling attacks. The well-known classification method Naïve Base is used to classify shilling attacks [34]. This study uses the Expectation Maximization (EM) algorithm to develop the initial classifier iteratively. In another study [35], Co-Forest algorithm is used for attack detection semi-supervised.

SVM is a powerful method for classification problems. However, it needs to label data for classify data. The study uses one class SVM to detect shilling attacks in the data. The One-Class Support Vector Machines (SVMs) represent an intriguing and powerful tool for addressing various challenges in machine learning and data analysis. One-Class SVMs are a specialized variant of the traditional Support Vector Machines, primarily designed for anomaly detection and novelty detection tasks.

The application of One-Class SVMs can be particularly valuable in multiple domains, including computer science, engineering, social sciences, and even natural sciences. These algorithms allow researchers to tackle problems where the focus is on identifying rare or unusual observations distinct from the majority of the data points. By doing so, One-Class SVMs aid in understanding exceptional patterns, detecting outliers, and characterizing anomalies that deviate significantly from the norm.

One-Class SVMs operate on the principle of separating a high-dimensional feature space into two regions: the region that contains the majority of the data (normal instances) and the region that represents the outliers or anomalies. This separation is achieved by constructing an optimal hyperplane that maximizes the margin between the normal instances and the origin of the feature space. In this way, One-Class SVMs learn to encapsulate the characteristics of normal data instances and provide a measure of their separability.

One-Class SVMs can be employed in various scenarios. For instance, in computer science, these algorithms can help identify cybersecurity threats, detect network intrusions, or distinguish malicious software from benign programs [36]. In engineering, One-Class SVMs can aid in fault diagnosis, anomaly detection in industrial processes, or quality control in manufacturing [37,38]. Social sciences can benefit from One-Class SVMs for identifying abnormal behavior patterns in human activities, fraudulent transactions in finance, or anomalous trends in sociological studies [39,40]. Even in natural sciences, these algorithms can assist in detecting rare events, outliers in environmental data, or anomalies in biological systems [41].

One of the key advantages of One-Class SVMs is their ability to operate with limited training data. Given that anomaly detection problems typically involve sparse and imbalanced datasets, One-Class SVMs excel in capturing the underlying structure of normal instances while being robust to the presence of outliers. Moreover, these algorithms provide a measure of confidence or probability associated with the classification, aiding researchers in interpreting the significance of detected anomalies.

Despite their numerous advantages, One-Class SVMs also have certain limitations. The choice of appropriate kernel functions, which are responsible for mapping the data into high-dimensional spaces, can significantly impact the performance of these algorithms. Optimization aims to improve the performance of the algorithms [42]. Parameter optimization is important for optimization problems [43]. Additionally, selecting optimal parameters, such as the kernel width and the regularization parameter, often requires careful tuning and cross-validation [44–49]. There are two parameters for OCSVM, v and ρ. v takes values between 0 and 1 and control the level of relaxation, and ρ is the bias term. These parameters are explained in detail in the next section. Their level is highly effective on accuracy rate of method. In Figure , effect of different levels of parameter is given. In Figure 1.a, basic levels of parameters are given. For these levels, train error is 9.5% and novel regular error is 20% and novel abnormal error is 0%. As we can see from Figure 1.a, an overfitting is observed for this scenario. All outliers are classified correctly, but errors for regular item is high. If the ρ is decreased while v is same boundaries are so large then, abnormal items are in the normal region. When two parameter levels are decrease, normal items errors are low, but abnormal items error is still high. If parameter levels are selected as v=0.01, ρ=0.1, the best accuracy rate is obtained. However, this level is not optimal. Different methods are proposed for parameter selection, these methods will be given at the end of this section.

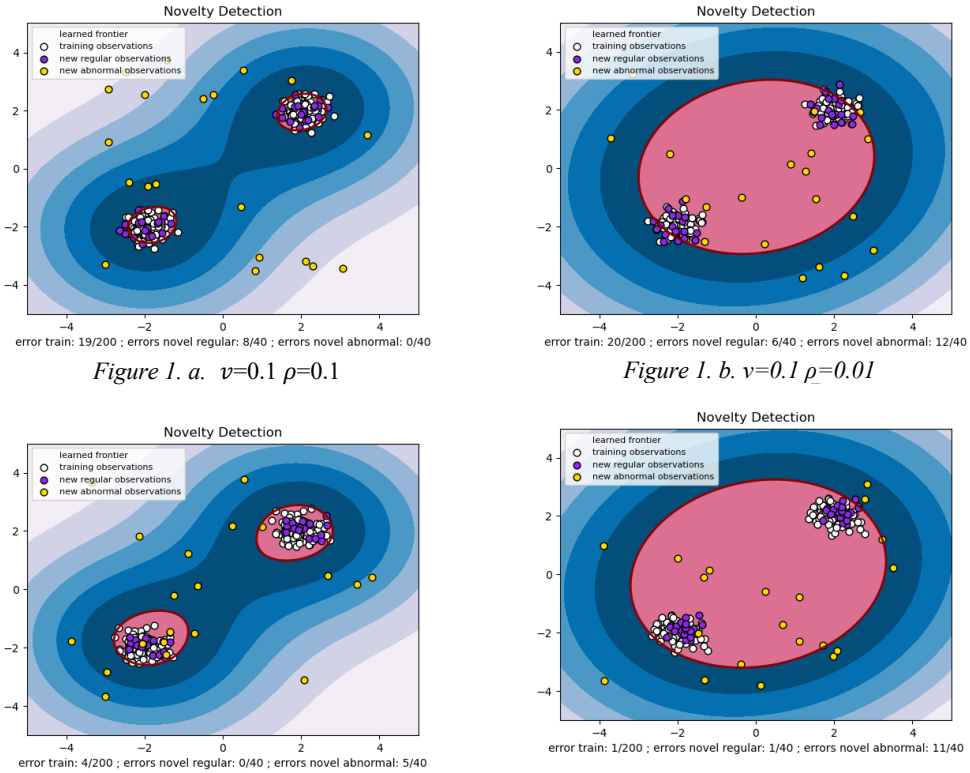

Figure 1. a. v=0.1 ρ=0.1 *Figure 1. b.* v=0.1 ρ=0.01

Figure 1. c. v=0.01 ρ=0.1 *Figure 1. d.* v=0.01 ρ=0.01

Figure 1. *Effects of different levels of OCSVM parameters*

In this study, the OCSVM method is adapted to the recommender system for shilling attack detection. Hyperparameters of OCSVM are selected by an unsupervised method.

Aforementioned studies show that there is a broad literature about shilling attack detection. In this study, unlike the literature, attack detection with one class support vector machines is presented. Thus, it is possible to detect different attacks with the model to be built with only the information of genuine users. Hyperparameters

of OCSVM is provided by Quick Model Selection method.

This study organized as follow: in introduction section, the problem and attack types are defined. A literature review is given about the problem, attack types, attack detection methods and used methods in the study. Furthermore, motivation of the study is given in the introduction section. In materials and methods, OCSVM algorithm is defined and pseudo codes of used algorithms are given. Followingly, in experimental evaluation section, an example is given to validate to proposed methodology. Discussion of results are given in this section. Lastly, in conclusions, results of study and future directions are given.

MATERIALS AND METHODS

OCSVM method is introduced in this section. In this study, a one-class classifier is needed to classify genuine users from attack profiles. Basically, single class classifiers can be grouped into four categories: intensity-based, distance-based, reconstruction-based and boundary-based. Since the obtained model will be used in the future, the boundary-based OCSVM method is preferred.

One-Class Support Vector Machines

The primal quadratic problem that defines the OCSVM classifier is

$$\begin{aligned} &\min_{\omega,\xi,\rho} \frac{1}{2}\left\|\omega^2\right\| + \frac{1}{vn}\sum_{i=1}^{n}\xi_i - \rho \\ &s.t.\left\langle \omega.\varphi\left(x_i\right)\right\rangle \geq \rho - \xi_i \\ &\xi_i \geq 0, i = 1,2,...,n \end{aligned} \tag{1}$$

where $\omega \in R^d$, $\xi = \left[\xi_1,...,\xi_n\right]$ are slack variables, ρ is a bias term, and $0 < v \leq 1$. Slack variables allow to relax the constraints imposed by the problem. The v parameter controls the level of relaxation. If v is set near 0, the penalty term $\sum_{i=1}^{n}\xi_i$ vanishes. In other words, it causes overfitting, the model aims to separate all points in the dataset. In contrast, if this parameter is very close to 1, the algorithm has more latitude to minimize (1). In this case underfitting is observed. Most of the points labeled as anomaly.

(1) is transformed into the following quadratic dual problem using Lagrange multipliers in order to avoid computing the explicit mapping:

$$\begin{aligned} &\min_{[a_i]} \frac{1}{2}\sum_{i,j} a_i a_j k\left(x_i, x_j\right) \\ &s.t. 0 \leq a_i \leq 1/vn \\ &\sum_{i=1}^{n} a_i = 1,\ i = 1,2,...,n \end{aligned} \tag{2}$$

Shilling attacks pose a significant threat to user trust in recommender systems. When users perceive recommendations as biased or manipulated, their confidence in the system diminishes. Users may question the authenticity of the recommended items and the system's ability to understand their preferences accurately. This erosion of trust can lead to decreased engagement, reduced satisfaction, and even abandonment of the recommender system altogether.

The pseudo code of proposed method is given in Algorithm 1. In the proposed study, firstly most-rated 50 items are selected from user item matrix. Then each item matrix is calculated. Followingly, train test and attack profiles are constructed. Model parameter optimization is provided by Quick Model Selection method for each train set. In this study 5-fold cross-validation is applied. For each test set model validation is provided.

In this study, 10 attack set is added to the dataset separately for each item, attack size and filler size. These different attack sets are predicted with fitted model. Then performance metrics are calculated.

Algorithm 1: Shilling Attack Detection wit OCSVM

Inputs: User-Item Matrix (X)

for $k \leftarrow 1$ to 50 **do**
  Find most rated 50 items
  Construct item's matrix for each item
  Construct train, test, and attack sets
  Obtain optimal parameter for each train set for OCSVM
  for $j \leftarrow 1$ to 5 **do**
    fit OCSVM model for item i train set j
    Predict train set j
    Calculate train accuracy rate
    for $k \leftarrow 1$ to 10 do
      Predict attack set k, injected to item i
    end for
    Calculate precision and recall values
  end for
end for

EXPERIMENTAL EVALUATION

In this study, MovieLens100K dataset is used as an experimental example. This dataset includes 100 000 ratings form 943 users to 1652 movies. Ratings take values between 0 and 5.

The dataset used in this study consists of four columns of data. The first column is the user id, the second column is the item number, the third column is the score given and the last column is the time stamp at which the score was given. In the data pre-processing step, these data were transformed into matrices for each item. The column values of the resulting matrices are the ids of the movies and the rows are all the scores of the users who rated that product. In this way, the data was made ready for OCSVM.

In this study, k-fold cross validation is applied for model validation. In the training dataset only genuine profiles are used. 80% of genuine profiles are used for training dataset and remain 20% percent of genuine profiles are used for test dataset. Additionally, attack profiles are tested with constructed model. Accuracy rates of training dataset is given in Table 2. All training datasets are used for all types attacks for different filler and attack sizes. However, small changes are observed for accuracy rates. The reason of mentioned small changes are k-fold cross validation. Different parts are selected randomly as train and test samples for each replication. The flowchart of proposed algorithm is given in Figure 2.

Table 2. *MovieLens100K dataset training dataset accuracy rates*

	Attack Size	1			3			5		
	Filler Size	1	3	5	1	3	5	1	3	5
Accuracy	Random	0,9795	0,9806	0,9795	0,9812	0,9805	0,9792	0,9821	0,9821	0,9804
	Average	0,9800	0,9798	0,9795	0,9813	0,9806	0,9800	0,9827	0,9810	0,9807
	Mixed-Attack	0,9793	0,9802	0,9795	0,9821	0,9801	0,9795	0,9825	0,9798	0,9800

Three types of attacks are injected to data, random, average and mixed attacks. Mixed attack consists of random and average attacks. These attacks are injected to most rated 50 items in the dataset. 10 replications are provided for each item. Finally, Table 3 is obtained for all test results. In Table 3, recall, precision, and F1-measure scores are given for all attack models for different attack and filler size. In general, above 96 percent

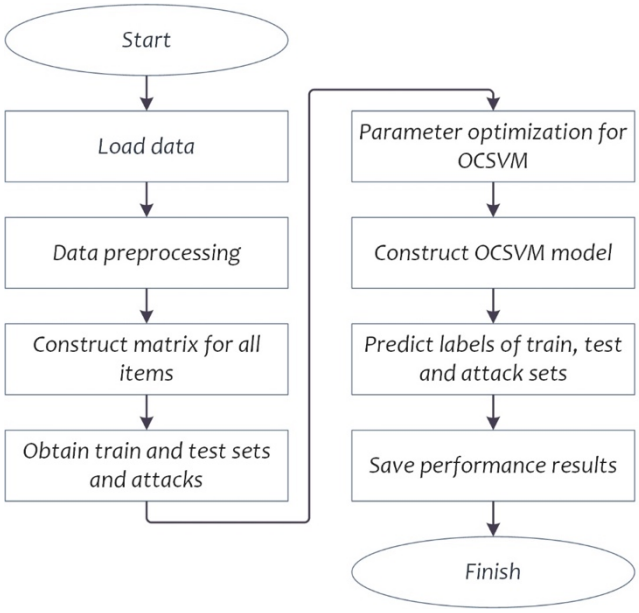

Figure 2. *Flowchart of proposed algorithm*

results are obtained for F1-Measure. Especially, lower filler sizes give better results than high filler sizes. In this study, quick model selection [46] method is used to predict hyperparameters of OCSVM. Different hyperparameter estimation methods can achieve higher scores for their performance metrics.

Table 3. *MovieLens100K Dataset Test Results*

	Attack Size	1			3			5		
	Filler Size	1	3	5	1	3	5	1	3	5
Recall	Random	1,0000	0,9994	0,9946	1,0000	0,9991	0,9840	1,0000	0,9982	0,9741
	Average	1,0000	0,9997	0,9972	1,0000	0,9996	0,9904	1,0000	0,9990	0,9843
	Mixed-Attack	1,0000	0,9999	0,9974	1,0000	0,9995	0,9896	1,0000	0,9989	0,9849
Precision	Random	0,9431	0,9424	0,9423	0,9479	0,9445	0,9486	0,9486	0,9475	0,9430
	Average	0,9465	0,9431	0,9424	0,9493	0,9458	0,9510	0,9510	0,9479	0,9457
	Mixed-Attack	0,9465	0,9431	0,9424	0,9493	0,9458	0,9510	0,9510	0,9479	0,9457
F1-Measure	Random	0,9707	0,9701	0,9678	0,9733	0,9710	0,9625	0,9736	0,9722	0,9583
	Average	0,9725	0,9706	0,9690	0,9740	0,9720	0,9664	0,9749	0,9728	0,9646
	Mixed-Attack	0,9720	0,9705	0,9691	0,9734	0,9714	0,9658	0,9749	0,9725	0,9641

The results show that the accuracy rates, which are limited to 98% in the learning set, increase up to 100% in the test set. There are relative decreases in performance metrics as the attack size and fill rate increase. In this study, attack users are considered positive observations. However, handled data is a sparse data. The filler rates of the data are about 6%. If an appropriate hyperparameter selection method is choose for this data, results will be better.

CONCLUSIONS

Recommender systems are a powerful technique to cope with information overload problems. However, shilling attacks pose a significant challenge to the trustworthiness and integrity of recommender systems. It is crucial to emphasize the importance of researching, understanding, and mitigating these attacks. We can foster more reliable, fair, and user-centric recommender systems by developing robust defenses and advancing the field's knowledge. Ultimately, this will lead to enhanced user satisfaction, increased trust, and more meaningful recommendations that align with individual preferences and interests.

In the literature, supervised, semi-supervised, and unsupervised methods are proposed to detect shilling attacks. In the study, a mathematical detection algorithm is adapted to recommender system to shilling attack detection. Incorporating One-Class Support Vector Machines into research and analysis can greatly enhance the understanding and detection of anomalies in various domains. Their ability to identify and characterize rare events, outliers, and unusual patterns provides valuable insights and can contribute to advancements in fields ranging from computer science and engineering to social sciences and natural sciences. By leveraging the power of One-Class SVMs, academics can effectively address the challenges posed by anomaly detection. Different hyperparameter selection methods are investigated and an unsupervised method is selected for OCSVM hyperparameters. In this study, OCSVM method is applied for different attack and filler sizes for MovieLens100K dataset. Accuracy rates are given for all scenarios.

Generally, dataset used for recommender systems are sparse datasets. For this reason, new hyperparameter selection algorithms for OCSVM can propose in future studies. This study uses movie database as experimental evaluation, different datasets can be used in further studies. Sigmoid kernel function is used in this study. In the future studies, different kernel functions can be presented for shilling attacks. Additionally, boundary-based one class classification algorithm is used in this study. For future studies, different classifiers can be used for classification method.

Acknowledgements

This study is supported by Eskisehir Technical University, Scientific Research Projects Committee (ESTUBAP- 20DRP076).

REFERENCES

[1] Y. Ge, S. Zhao, H. Zhou, C. Pei, F. Sun, W. Ou, Y. Zhang, Understanding Echo Chambers in E-commerce Recommender Systems, *SIGIR 2020 - Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, *Association for Computing Machinery, Inc*, Rutgers University, Piscataway, United States, 2020: ss. 2261-2270. doi:10.1145/3397271.3401431.

[2] R.M. Frey, D. Vučkovac, A. Ilic, A secure shopping experience based on blockchain and beacon technology, içinde: G. I., S. A. (Ed.), *CEUR Workshop Proceedings*, *CEUR-WS*, ETH Zurich, Weinbergstrasse 56/58, Zurich, 8092, Switzerland, 2016.

[3] Y. Han, Z. Han, J. Wu, Y. Yu, S. Gao, D. Hua, A. Yang, Artificial Intelligence Recommendation System of Cancer Rehabilitation Scheme Based on IoT Technology, *IEEE Access*. 8 (2020), 44924-44935. doi:10.1109/ACCESS.2020.2978078.

[4] C.C. Yang, L. Jiang, Enriching User Experience in Online Health Communities Through Thread Recommendations and Heterogeneous Information Network Mining, *IEEE Transactions on Computational Social Systems*. 5 (2018), 1049-1060. doi:10.1109/TCSS.2018.2879044.

[5] A. Alsalemi, Y. Himeur, F. Bensaali, abbes amira, C. Sardianos, I. Varlamis, G. Dimitrakopoulos, achieving Domestic Energy Efficiency Using Micro-Moments and Intelligent Recommendations, *IEEE Access*. 8 (2020), 15047-15055. doi:10.1109/aCCESS.2020.2966640.

[6] C. Sardianos, I. Varlamis, G. Dimitrakopoulos, D. Anagnostopoulos, A. Alsalemi, F. Bensaali, Y. Himeur, A. Amira, REHAB-C: Recommendations for Energy HABits Change, *Future Generation Computer Systems*. 112 (2020), 394-407. doi:10.1016/j.future.2020.05.041.

[7] Q. Li, J. Kim, A deep learning-based course recommender system for sustainable development in education, *Applied Sciences (Switzerland)*. 11 (2021). doi:10.3390/app11198993.

[8] P. V. Kulkarni, S. Rai, R. Kale, Recommender System in eLearning: A Survey, *Proceeding of International Conference on Computational Science and Applications*. (2020), 119-126. doi:10.1007/978-981-15-0790-8_13.

[9] S. Puglisi, J. Parra-Arnau, J. Forné, D. Rebollo-Monedero, On content-based recommendation and user privacy in social-tagging systems, *Computer Standards and Interfaces*. 41 (2015), 17-27. doi:10.1016/j.csi.2015.01.004.

[10] A. Bilge, Z. Ozdemir, H. Polat, A novel shilling attack detection method, *Procedia Computer Science*, 31 (2014), 165-174. doi:10.1016/j.procs.2014.05.257.

[11] M. Si, Q. Li, Shilling attacks against collaborative recommender systems: a review, *Artificial Intelligence Review*. 53 (2020), 291-319. doi:10.1007/s10462-018-9655-x.

[12] F. Rezaimehr, C. Dadkhah, A survey of attack detection approaches in collaborative filtering recommender systems, *Artificial Intelligence Review*. 54 (2021), 2011-2066. doi:10.1007/s10462-020-09898-3.

[13] Z. Batmaz, B. Yilmazel, C. Kaleli, Shilling attack detection in binary data: a classification approach, *Journal of Ambient Intelligence and Humanized Computing*. 11 (2020), 2601-2611. doi:10.1007/s12652-019-01321-2.

[14] F. Zhang, Q. Zhou, HHT-SVM: An online method for detecting profile injection attacks in collaborative recommender systems, *Knowledge-Based Systems*. 65 (2014), 96-105. doi:10.1016/j.knosys.2014.04.020.

[15] S. Zahra, M.A. Ghazanfar, A. Khalid, M.A. Azam, U. Naeem, A. Prugel-Bennett, Novel centroid selection approaches for KMeans-clustering based recommender systems, *Information Sciences*. 320 (2015), 156-189. doi:10.1016/j.ins.2015.03.062.

[16] F. Zhang, S. Wang, Detecting Group Shilling Attacks in Online Recommender Systems Based on Bisecting K-Means Clustering, *IEEE Transactions on Computational Social Systems*. 7 (2020), 1189-1199. doi:10.1109/TCSS.2020.3013878.

[17] R. Bhaumik, B. Mobasher, R. Burke, A Clustering Approach to Unsupervised Attack Detection in Collaborative Recommender Systems, *Proceedings of the 7th IEEE international conference on data mining*. (2011), 181-187.

[18] R. Bhaumik, C. Williams, B. Mobasher, R. Burke, Securing collaborative filtering against malicious attacks

through anomaly detection, *AAAI Workshop - Technical Report*. WS-06-10 (2006), 50-59.

[19] Z. Wu, J. Wu, J. Cao, D. Tao, HySAD: A semi-supervised hybrid shilling attack detector for trustworthy product recommendation, içinde: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2012: ss. 985-993. doi:10.1145/2339530.2339684.

[20] Z. Yang, Z. Cai, X. Guan, Estimating user behavior toward detecting anomalous ratings in rating systems, *Knowledge-Based Systems*. 111 (2016), 144-158. doi:10.1016/j.knosys.2016.08.011.

[21] B. Mobasher, R. Burke, R. Bhaumik, J.J. Sandvig, Attacks and remedies in collaborative recommendation, *IEEE Intelligent Systems*. 22(3) (2007), 56-63. doi:10.1109/MIS.2007.45.

[22] S.K. Lam, J. Riedl, Shilling recommender systems for fun and profit, içinde: *Thirteenth International World Wide Web Conference Proceedings, WWW2004*, *Association for Computing Machinery*, GroupLens Research, Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455, United States, 2004: ss. 393-402. doi:10.1145/988672.988726.

[23] R. Burke, B. Mobasher, R. Zabicki, Runa. Bhaumik, Identifying Attack Models for Secure Recommendation, *Beyond Personalization*. (2005), 19-25.

[24] B. Mobasher, R. Burke, R. Bhaumik, C. Williams, Effective Attack Models for Shilling Item-Based Collaborative Filtering System, *WEBKDD*. 2005, (2005).

[25] M.P. O'Mahony, N.J. Hurley, G.C.M. Silvestre, Recommender systems: Attack types and strategies, içinde: *Proceedings of the National Conference on Artificial Intelligence*, 2005: ss. 334-339.

[26] C. Williams, B. Mobasher, Thesis: Profile Injection Attack Detection for Securing Collaborative Recommender Systems, (2006), 1-47.

[27] A. Pektaş, O. İnan, Ağaç Tohum Algoritmasının Kümeleme Problemlerine Uygulanması, *Necmettin Erbakan Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*. 4 (2022), 1-10. doi:10.47112/neufmbd.2022.8.

[28] F. Zhang, Z. Zhang, P. Zhang, S. Wang, UD-HMM: An unsupervised method for shilling attack detection based on hidden Markov model and hierarchical clustering, *Knowledge-Based Systems*. 148 (2018), 146-166. doi:10.1016/j.knosys.2018.02.032.

[29] H. Cai, F. Zhang, An unsupervised method for detecting shilling attacks in recommender systems by mining item relationship and identifying target items, *Computer Journal*. 62 (2019), 579-597. doi:10.1093/comjnl/bxy124.

[30] L. Yang, W. Huang, X. Niu, Defending shilling attacks in recommender systems using soft co-clustering, *IET Information Security*. 11 (2017), 319-325. doi:10.1049/iet-ifs.2016.0345.

[31] A. Davoudi, M. Chatterjee, Detection of profile injection attacks in social recommender systems using outlier analysis, *Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017*. 2018-Janua (2017), 2714-2719. doi:10.1109/BigData.2017.8258235.

[32] C. Panagiotakis, H. Papadakis, P. Fragopoulou, Unsupervised and supervised methods for the detection of hurriedly created profiles in recommender systems, *International Journal of Machine Learning and Cybernetics*. 11 (2020), 2165-2179. doi:10.1007/s13042-020-01108-4.

[33] J.M. Alostad, Improving the Shilling Attack Detection in Recommender Systems Using an SVM Gaussian Mixture Model, *Journal of Information and Knowledge Management*. 18(01) (2019). doi:10.1142/S0219649219500114.

[34] L. Zhang, Y. Yuan, Z. Wu, J. Cao, Semi-SGD: Semi-Supervised Learning Based Spammer Group Detection in Product Reviews, *Proceedings - 5th International Conference on Advanced Cloud and Big Data, CBD 2017*. 2017, 368-373. doi:10.1109/CBD.2017.70.

[35] Q. Zhou, L. Duan, Semi-supervised recommendation attack detection based on Co-Forest, *Computers and Security*. 109 (2021), 102390. doi:10.1016/j.cose.2021.102390.

[36] Y. Wang, R. Zhang, N. Masoud, H.X. Liu, Anomaly detection and string stability analysis in connected automated vehicular platoons, *Transportation Research Part C: Emerging Technologies*. 151 (2023), 104114. doi:10.1016/j.trc.2023.104114.

[37] C. Li, L. Mo, H. Tang, R. Yan, Lifelong condition monitoring based on NB-IoT for anomaly detection of

machinery equipment, *Procedia Manufacturing*. 49 (2020), 144-149. doi:10.1016/j.promfg.2020.07.010.

[38] C. Cao, M. Liu, B. Li, Y. Wang, Mechanical fault diagnosis of high voltage circuit breakers utilizing VMD based on improved time segment energy entropy and a new hybrid classifier, *IEEE Access*. 8 (2020), 177767-177781. doi:10.1109/ACCESS.2020.3027478.

[39] A. Karasmanoglou, M. Antonakakis, M. Zervakis, ECG-Based Semi-Supervised Anomaly Detection for Early Detection and Monitoring of Epileptic Seizures, *International Journal of Environmental Research and Public Health*. 20 (2023), 5000. doi:10.3390/ijerph20065000.

[40] A.A. Abdulhussein, M.F. Nasrudin, S.M. Darwish, Z.A.A. Alyasseri, A Genetic Algorithm Based One Class Support Vector Machine Model for Arabic Skilled Forgery Signature Verification, *Journal of Imaging*. 9 (2023). doi:10.2139/ssrn.4303232.

[41] T. Cheng, A. Dairi, F. Harrou, Y. Sun, T. Leiknes, Monitoring influent conditions of wastewater treatment plants by nonlinear data-based techniques, *IEEE Access*. 7 (2019), 108827-108837. doi:10.1109/ACCESS.2019.2933616.

[42] M. Karakoyun, A. Özkış, Transfer Fonksiyonları Kullanarak İkili Güve-Alev Optimizasyonu Algoritmalarının Geliştirilmesi ve Performanslarının Karşılaştırılması, *Necmettin Erbakan Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*. 3 (2021), 1-10.

[43] M. Erişoğlu, N. Yaman, Ridge Tahminine Dayalı Kantil Regresyon Analizinde Yanlılık Parametresi Tahminlerinin Performanslarının Karşılaştırılması, *Necmettin Erbakan University Journal of Science and Engineering*. 1 (2019), 103-111.

[44] Y. Guerbai, Y. Chibani, Y. Meraihi, Techniques for Selecting the Optimal Parameters of One-Class Support Vector Machine Classifier for Reduced Samples, *International Journal of Applied Metaheuristic Computing*. 13 (2021,) 1-15. doi:10.4018/ijamc.290533.

[45] S. Wang, Q. Liu, E. Zhu, F. Porikli, J. Yin, Hyperparameter selection of one-class support vector machine by self-adaptive data shifting, *Pattern Recognition*. 74 (2018) 198-211. doi:10.1016/j.patcog.2017.09.012.

[46] Z. Ghafoori, S.M. Erfani, S. Rajasegarar, J.C. Bezdek, S. Karunasekera, C. Leckie, Efficient Unsupervised Parameter Estimation for One-Class Support Vector Machines, *IEEE Transactions on Neural Networks and Learning Systems*. 29 (2018), 5057-5070. doi:10.1109/TNNLS.2017.2785792.

[47] S. Wang, Q. Liu, E. Zhu, J. Yin, W. Zhao, MST-GEN: An Efficient Parameter Selection Method for One-Class Extreme Learning Machine, *IEEE Transactions on Cybernetics*. 47 (2017), 3266-3279. doi:10.1109/TCYB.2017.2707463.

[48] A. Anaissi, A. Braytee, M. Naji, Gaussian Kernel Parameter Optimization in One-Class Support Vector Machines, *Proceedings of the International Joint Conference on Neural Networks.* 2018, doi:10.1109/IJCNN.2018.8489383.

[49] Y. Xiao, H. Wang, W. Xu, Parameter selection of gaussian kernel for one-class SVM, *IEEE Transactions on Cybernetics*. 45 (2015), 941-953. doi:10.1109/TCYB.2014.2340433.