|| Volume 5 || Issue 12 || April 2021 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH & ENGINEERING TRENDS
Multidisciplinary Journal
Double-Blind Peer Reviewed Refereed Open Access International Journal

IDENTIFICATION OF MALICIOUS INJECTION ATTACKS IN PROFILE INJECTION AND CO-VISITATION BEHAVIORS ON PRODUCT REVIEWS

Sayali Walunjkar[1], Nikita Deshmane[2], Rutuja Supekar[3], Vaishnavi Dalvi[4], Prof. Pravin Avhad[5]

[1]Savitribai Phule Pune University,
Dept. of computer engineering, S.C.S.M.C.O.E, Nepti, Maharashtra, India
sayaliwalunjkar123@gmail.com

[2]Savitribai Phule Pune University,
Dept. of computer engineering, S.C.S.M.C.O.E, Nepti, Maharashtra, India
nikitadeshmane11@gmail.com

[3]Savitribai Phule Pune University,
Dept. of computer engineering, S.C.S.M.C.O.E, Nepti, Maharashtra, India
rutujasupekar1232@gmail.com

[4]Savitribai Phule Pune University,
Dept. of computer engineering, S.C.S.M.C.O.E, Nepti, Maharashtra, India
vaishanvidalvi111@gmail.com

[5]Savitribai Phule Pune University,
Professor, Dept. of computer engineering, S.C.S.M.C.O.E, Nepti, Maharashtra, India
pravin123.avhad@gmail.com

----------***----------

Abstract: Recommender systems have become an essential component in a wide range of web services. It is believed that recommender systems recommend a user items (e.g., products on Amazon) that match the user's preference. Now a days, a serious a part of everyone trusts on content in social media like opinions and feedbacks of a subject or a product. The liability that anyone can begin a survey provides a brilliant chance to fake co-visitations to compose spam surveys about products and services for various interests. Recognizing these fake co-visitations and therefore the fake content may be a wildly debated issue of research and in spite of the very fact that a powerful number of studies are done as lately toward this end, yet thus far the procedures set forth still scarcely distinguish fake reviews, fake co-visitations, and none of them demonstrate the importance of every extracted feature type. During this investigation, we propose a completely unique structure, named Fake review detection system, which uses spam highlights for demonstrating review datasets to style fake co-visitations detection method into a classification issue in such networks. Utilizing the importance of fake features help we to accumulate better outcomes regarding different metrics on review datasets. We also discuss strategies to mitigate our attacks.

Keywords: *Amazon, Machine learning, Product.*

----------***----------

I INTRODUCTION

In a recommender system, we have a set of users (e.g., registered users, unregistered visitors) and items. Two widely used recommendation tasks are user-to-item recommendation and item to- item recommendation. In a user-to-item recommendation, the system recommends items to a user based on the user's profile (e.g., the browsing history, the items the user liked or disliked). In an item-to-item recommendation, a list of items are recommended to a user when the user is visiting an item. This recommendation is commonly known as features like "People who viewed this also viewed". One particular category of recommender system to implement the two recommendation tasks, which we call co-visitation recommender system, is likely being widely used by web service providers (e.g. Amazon) due to its effectiveness and simplicity. Co-visitation recommender systems leverage covisitation information between items, and the key idea is that two items that were frequently co-visited in the past are likely to be co-visited in the future.

It was widely believed that recommender systems should recommend a user items that match the user's preference. However, recently proposed pollution attacks to user-to-item

recommendation, in which the recommender system is spoofed to recommend any target item (e.g., a video advertisement on YouTube) to a victim user. Their key idea is to inject fake information (e.g. fake reviews), which is related to the target item, into the victim user's profile via cross-site request forgery (CSRF) attacks.

In this work, we propose new attacks to spoof recommender systems to make recommendations as an attacker desires. Our attacks do not rely on CSRF, can be performed at a large scale, and are applicable to both user-to-item and item-to-item recommendations. In particular, we focus on covisitation recommender systems. Our key idea is to inject fake co-visitations to the system, and we call our attacks fake covisitation injection attacks. We note that attacking co-visitation recommender systems via injecting fake co-visitations is a natural idea. Our key contribution is to perform the first formal and systematic study on fake co-visitation injection attacks.

Organization of paper

Section 1: In this chapter, introduction of the project i.e. what is need, relevance and what is actual project idea.
Section 2: In this chapter, we briefly review the related work on mental disorder detection and their different techniques.
Chapter 3: Describe open issues.
Chapter 4: in this chapter, conclusion of the project, Future Scope of the project.

We are going to use following UML Diagrams:-

1. Use Case Diagram
2. Activity Diagram
3. Sequence Diagram
4. Class Diagram

II RELATED WORK.

In this paper Author develop a unified detection approach named IMIA-HCRF, to progressively discriminate malicious injection behaviors for recommender systems. First, disturbed data are empirically eliminated by implementing both the construction of association graph and enhancement of dense behaviors, which can be adapted to different attacks. Then, the smooth boundary of co-visitation behaviors is further segmented using higher order potentials, which is Finally leveraged to determine the concerned injection behaviors[1].

In this work, author propose new attacks to recommender systems. Attacks exploit fundamental vulnerabilities of recommender systems and can spoof a recommender system to make recommendations as an attacker desires. Key idea is to inject fake co-visitations to the system[2].

Author improve detection performance from following two aspects. Firstly, extract well-designed features from user profiles based on the statistical properties of the diverse attack models, making hard detection scenarios become easier to perform. Then, refer to the general idea of re-scale Boosting (RBoosting) and AdaBoost, then apply a variant of AdaBoost, called the re-scale AdaBoost (RAdaBoost) as detection method based on the extracted features [3].

In this work, author propose GANG, a guilt-by-association method on directed graphs, to detect fraudulent users in OSNs. GANG is based on a novel pairwise Markov Random Field that we design to capture the unique characteristics of the fraudulent-user-detection problem in directed OSNs. In the basic version of GANG, given a training dataset, we leverage Loopy Belief Propagation (LBP) to estimate the posterior probability istribution for each user and uses it to predict a user's label[4].

In this paper, author examine the detection of shilling attacks in privacy-preserving collaborative filtering systems. Authors utilize four attack-detection methods to filter out fake profiles produced by six well-known shilling attacks on perturbed data. They evaluate these detection methods with respect to their ability to identify bogus profiles. Real data-based experiments are performed. Empirical outcomes demonstrate that some of the detection methods are very successful at filtering out fake profiles in privacy-preserving collaborating filtering schemes[5].

Author provide a formulation for learning to attack a recommender as a repeated general-sum game between two players, i.e., an adversary and a recommender oblivious to the adversary's existence. We consider the challenging case of poison¬ing attacks, which focus on the training phase of the recommender model. Author generate adversarial user profles targeting subsets of users or items, or generally the top-K recommendation quality. Moreover, author ensure that the adversarial user profles remain unno¬ticeable by preserving proximity of the real user rating/interaction distribution to the adversarial fake user distribution. Author offer a wide range of experiments, in¬stantiating the proposed method for the case of the classic popular approach of a low-rank recommender, and illustrating the extent of the recommender's vulnerability to a variety of adversarial intents[6].

In this work, author present ProNE - a fast, scalable, and effective model, whose single-thread version is 10–400_ faster than efficient network embedding benchmarks with 20 threads, including LINE, DeepWalk, node2vec, GraRep, and HOPE. As a concrete example, the single-thread ProNE requires only 29 hours to embed a network of hundreds of millions of nodes while it takes LINE weeks and Deep- Walk months by using 20

threads. To achieve this, ProNE first initializes network embeddings efficiently by formulating the task as sparse matrix factorization. The second step of ProNE is to enhance the embeddings by propagating them in the spectrally modulated space[7].

In this work, author propose Ianus, a Sybil detection method that leverages account registration information. Ianus aims to catch Sybils immediately after they are registered. First, using a realworld registration dataset with labeled Sybils from WeChat (the largest online social network in China), author perform a measurement study to characterize the registration patterns of Sybils and benign users. Author find that Sybils tend to have synchronized and abnormal registration patterns. Second, based on our measurement results, model Sybil detection as a graph inference problem, which allows us to integrate heterogeneous features[8].

In this work, Author propose SybilSCAR, a new structure based method to perform Sybil detection in OSNs. SybilSCAR maintains the advantages of existing methods while overcoming their limitations. Specifically, SybilSCAR is Scalable, Convergent, Accurate, and Robust to label noises. author first propose a framework to unify RW-based and LBP-based methods. Under our framework, these methods can be viewed as iteratively applying a (different) local rule to every user, which propagates label information among a social graph. Second, author design a new local rule, which SybilSCAR iteratively applies to every user to detect Sybils. We compare SybilSCAR with a state-of-the art RW-based method and a state-of-the-art LBP-based method, using both synthetic Sybils and large-scale social network datasets with real Sybils[9].

In this work, author propose a novel collective classification framework to address this long-standing challenge. author first formulate learning edge weights as an optimization problem, which quantifies the goals about the final reputation scores that aim to achieve. However, it is computationally hard to solve the optimization problem because the final reputation scores depend on the edge weights in a very complex way. To address the computational challenge, author propose to jointly learn the edge weights and propagate the reputation scores, which is essentially an approximate solution to the optimization problem[10].

III PROPOSED WORK

First, we propose to enhance profile injection behaviors and co-visitation injection behaviors via the elimination of disturbed data and representation of sparse behaviors, which also provides a possibility for the integrated detection of different injection attack behaviors. Second, we explore attributes of both nodes and edges of behavior association graph, and propose to incorporate unary potential and pairwise potential of higher order conditional random fields for informative representations of rating and co-visitation behaviors. Third, we develop a unified detection approach to identify both profile injection attacks and co-visitation injection attacks. Additionally, mixed profile injection attacks and mixed co-visitation injection attacks with different cases are implemented.

IV SYSTEM ARCHITECTURE

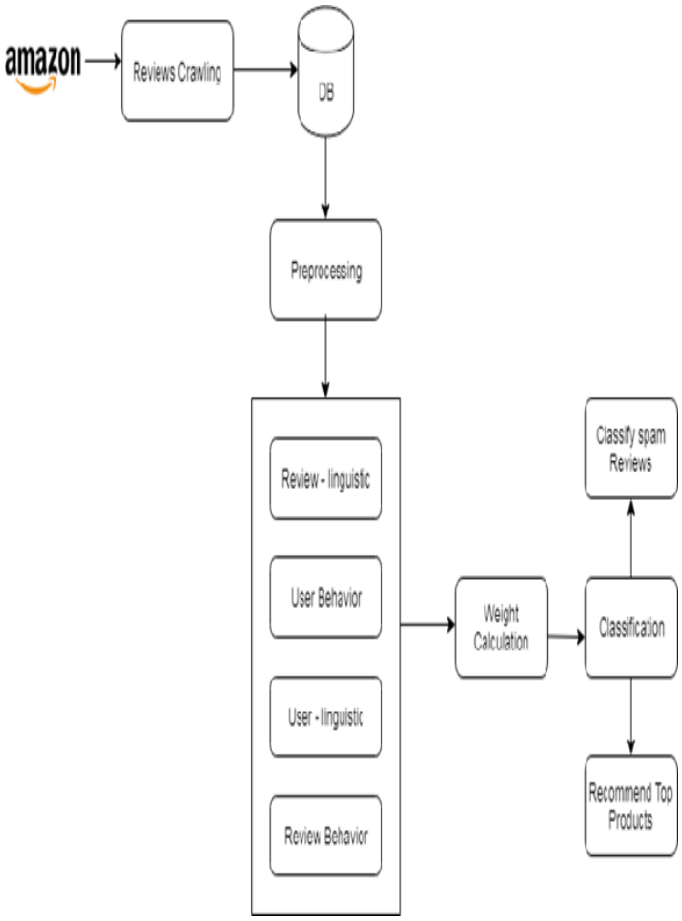

Fig 1: System Architecture

1.The attacker generates fake ratings and reviews by users toward selected products.

2.If an unpopular item suddenly has many co-visitations with some items, then it is possible that an attacker is trying to promote this item via our fake co-visitation injection attacks.

3.via analyzing temporal dynamic of visits and co-visits, the service provider could detect certain fake co-visitation and mitigate our attacks.

1. Use Case Diagram with necessary information

Use case diagram is used to show which operations are performed by the user and which operation are performed by the system.

|| Volume 5 || Issue 12 || April 2021 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH & ENGINEERING TRENDS
Multidisciplinary Journal
Double-Blind Peer Reviewed Refereed Open Access International Journal

Fig 2: Use Case Diagram

2. Activity Diagram with necessary information

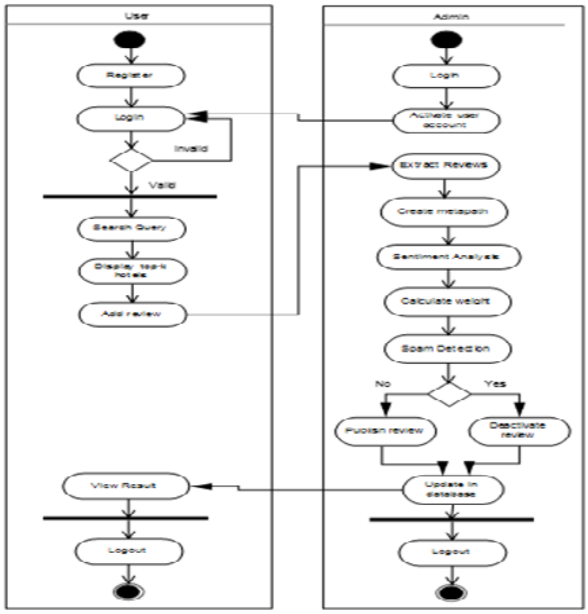

Fig 3: Activity Diagram

Activity Diagram shows the active flow of the system. In above diagram the flow of our project is shown actually how the data flow.

3. Sequence Diagram with necessary information

In sequence diagram step by step sequence of steps is shown. In above diagram first preprocess all train data and test data. Then by applying the train data train the machine and build the module and at the last apply machine learning algorithm on it . For testing purpose apply the test data on module and see the classification either fake or real.

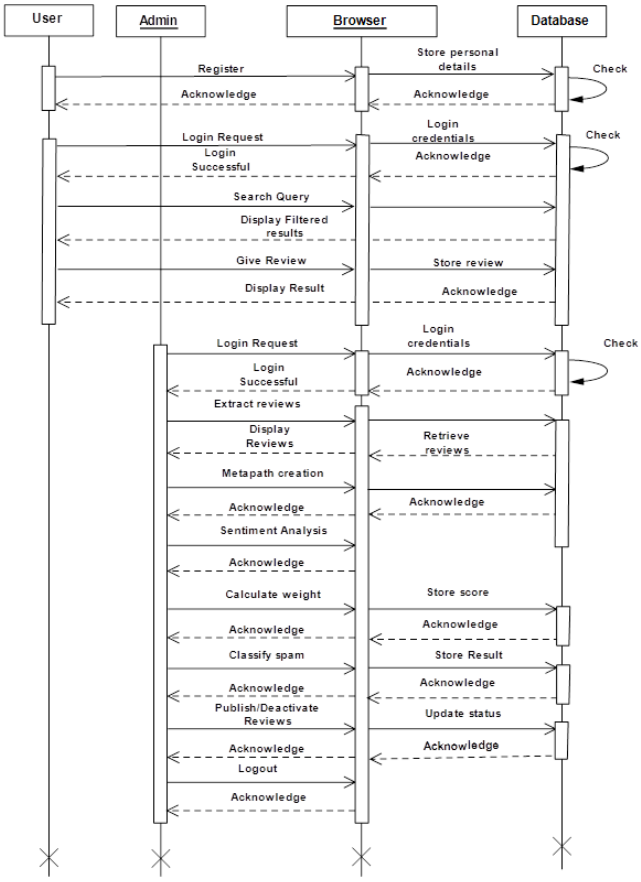

Fig 4: Sequence Diagram

4. Class Diagram with necessary information

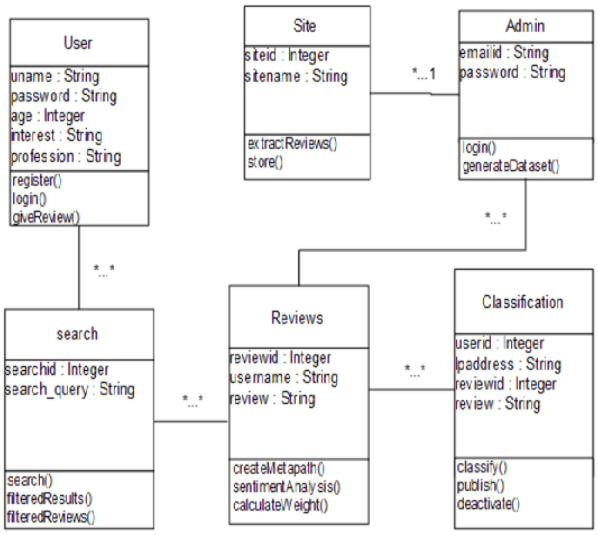

Fig 5: Class Diagram

V ALGORITHM

Input: Text File (comment or review) T , The sentiment lexicon L .

Output: Smt = {P,Ng and} and strength S where P: Positive, Ng:Negative, N:Neutral

Initialization:SumPos=SumNeg=0,where,SumPos:accumlates the polarity of positive tokens tismt in T, SumNeg: accumulates the polarity of negative tokens ti-smt in T, Begin

1.For each ti ∈ T do

2.Search for ti in L

3. If ti ∈ Pos−list then

4. SumPos ← SumPos + ti−smt

5. Else ifti ∈ Pos−list then

6. SumNeg ← SumNeg + ti−smt

7. End If

8. End For

9. If SumPos > |SumNeg| then

10. Smt = P

11. S=SumPos/(SumPos+SumNeg)

12. Else If SumPos < |SumNeg| then

13. Smt = Ng

14. S=SumNeg/(SumPos+SumNeg)

15. Else

16. Smt = N

17. S=SumPos/(SumPos+SumNeg)

18. End If End

VI ADVANTAGES

1. To enhance dense rating behaviours and co-visitation injection behaviours via the elimination of disturbed data and representation of sparse behaviours, which also provides a possibility for the integrated detection of different injection attack behaviours.

2. To develop a novel detection approach to identify both profile injection attacks and co-visitation injection attacks.

3. To display only trusted reviews to the users.

4. To identify spam and spammers as well as different type of analysis on this topic.

VII CONCLUSION

This work presents a machine learning strategy to detect profile injection attacks and co-visitation injection attacks for online recommender systems. Experimental results on both synthetic data and real-world data show that the elimination of disturbed data, determination of disturbed of dense behaviors, and potential segmentation exhibit considerable stability and discriminability among nodes (users or items) for detecting malicious injection behaviors.

ACKNOWLEDGEMENT

We take this apportunity to express my hearty thanks to all those who helped me in the completion of the project stage-1 on this topic. We whould especially like to ex- press my sincere gratitude to Prof. P.S.Avhad, my Guide and Prof. J.U.Lagad HOD Department of Computer Engineering who extended their moral support, inspiring guidance and encouraging independence throughout this task. We would also thank our Principal Dr. M.P.Nagarkar for his great insight and motivation. Last but not least, we whould like to thank my colleagues for their valuable suggestions.

9.REFERENCES

[1]Zhihai Yang, Qindong Sun, Yaling Zhang, and Wei Wang" Identification of Malicious Injection Attacks in Dense Rating and Co-visitation Behaviors" IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 12, 2020.

[2]G. Yang, N. Gong, and Y. Cai, "Fake co-visitation injection attacks to recommender systems," Network & Distributed System Security Symposium (NDSS), pp. 1–15, 2017.

[3]Z. Yang, L. Xu, Z. Cai, and Z. Xu, "Re-scale AdaBoost for attack detection in collaborative filtering recommender systems," Knowledge- Based Systems, vol. 100, pp. 74–88, 2016.

[4]B. Wang, N. Gong, and H. Fu, "GANG: Detecting fraudulent users in online social networks via guilt-by-association on directed graphs," IEEE International Conference on Data Mining, pp. 465–474, 2017.

[5]Gunes and H. Polat, "Detecting shilling attacks in private environments," Information Retrieval Journal, vol. 19, no. 6, pp. 1–26, 2016.

[6]K. Christakopoulou and A. Banerjee, "Adversarial attacks on an oblivious recommender," Proceedings of the 13th ACM Conference on Recommender Systems, pp. 322–330, 2019.

[7]J. Zhang, Y. Dong, Y. Wang, J. Tang, and M. Ding, "Prone: Fast and scalable network representation learning," In Proceedings of the 28th International Joint Conference on Artificial Intelligence, pp. 1–7, 2019.

[8]D. Yuan, Y. Miao, N. Gong, Z. Yang, Q. Li, D. Song, Q. Wang, and X. Liang, "Detecting fake accounts in online social networks at the time of registrations," In ACM Conference on Computer and Communications Security (CCS), pp. 1423–1438, 2019.

[9]B. Wang, L. Zhang, and N. Gong, "SybilSCAR: Sybil detection in online social networks via local rule based propagation," In IEEE Conference on Computer Communications (INFOCOM), pp. 1–9, 2017.

[10]B. Wang, J. Jia, and N. Gong, "Graph-based security and privacy analytics via collective classification with joint weight learning and propagation," In ISOC Network and Distributed System Security Symposium (NDSS), pp. 1–15, 2019.