INFERENCES BASED ON PROBABILITIES AND ASSESSMENTS OF THE RELIABILITY OF LINKS ARE USED TO IDENTIFY FRAUDULENT ACTIVITY IN ONLINE RECOMMENDATIONS

KISHORE BEZAWADA, Assistant Professor, Department of CSE**,** SREE DATTHA INSTITUTE OF ENGINEERING AND SCIENCE
G.SIRISHA, Assistant Professor, Department of CSE, SRIINDU COLLEGE OF ENGINEERING AND TECHNOLOGY(AUTONOMOUS)
DR.VENKATESH.E Assistant Professor, Department of EEE, SRIINDU COLLEGE OF ENGINEERING AND TECHNOLOGY(AUTONOMOUS)

Abstract—Because of the inherent weaknesses and openness of recommender systems, their growing usage as customised suggestion services like Amazon, TripAdvisor, and Yelp has increased the need for safe and practical anomaly detection methods. Defense against the ever-growing variety of harmful attacks targeting internet recommendations is a formidable problem. Moreover, there has been a recent uptick in both theoretical and practical investigation into the age-old challenge of describing and assessing sparse rating behaviours. This research takes an innovative approach to threat detection by studying probabilistic inference and evaluating the trustworthiness of behavioural linkages using coupled association networks derived from rating behaviours. Initially, using the users' innate rating motivation and the atomic propagation principles of coupled networks, an association graph is built from the original rating matrix. Then, using a factor graph model of a coupled network, we re-determine linkages of interest in the targeted network and assess the reliability of their behaviours. Lastly, by thoroughly examining the trustworthiness of both linkages and nodes in the targeted network, suspicious persons and things may be deduced experimentally. Comprehensive studies on synthetic data for profile injection assaults and co-visitation injection attacks, as well as real-world data including Amazon and TripAdvisor, illustrate the efficacy of the suggested detection technique compared with competing benchmarks.

Index-Terms— Recommender system, malicious attack, trustworthiness assessment, rating behaviour, and detection of attacks.

1 INTRODUCTION

Several popular websites, like Amazon, TripAdvisor, and eBay, rely on PERSONALIZATION recommender systems to aid users in finding relevant content [1, 2]. In particular, during the last two decades, CRTs (such as UBCF, IBCF, co-visitation based [3], [4]) have been created in various e-commerce systems. CRTs operate on the premise that if two users have exhibited similar interests in the past, they will continue to share those interests in the future. Nevertheless, collaborative recommender systems are especially susceptible to profile injection attacks (also known as shilling attacks) [7], [8], [9], [10], false co-visitation injection attacks [11], etc. due to the openness and basic weaknesses of CRTs [5], [6]. An adversary may degrade the quality of a CRT's suggestions or influence it to generate recommendations favourable to the adversary by injecting false data into the CRT. In practise, an adversary may trick a recommender system into suggesting a targeted item to as many people as possible by injecting fake users with properly prepared phoney ratings. To manipulate recommendations, an attacker may use empirical methods such as giving a higher rating to the targeted item (called a push or promotion attack [12]) or a lower rating to the item (called a nuke or demotion attack [12]), or they may inject fake co-visits into the system to spoof CRTs [11]; this has a significant impact on the public and ultimately shakes consumer and merchant trust in the online marketplace. Hence, it is becoming more important to study countermeasures against CRTs. 1.1 Prospects and Difficulties During the last decade, researchers have focused a lot of attention on both harmful attacks on and countermeasures for recommender systems [11], [12], [13], [14], [15]. Though malicious attacks in online recommender systems are always a risk, they also present opportunities, such as: (1) maintaining the security and fairness of virtual market has become an urgent problem to be solved as personalised recommendation technologies are increasingly adopted in online service platforms (e.g., Amazon, eBay); (2) new attack models, such as co-visitation injection attacks, pollution attacks, etc. [11], are gradually emerging, which presents a window of opportunity for researchers to improve their defences against these types of attacks. Simply said, there is still much utility in many traditional approaches of doing research. [12], [13], [14]. These new approaches are more feasible than their predecessors, but there are still certain problems that need fixing. As there is nothing that can be used to define common behavioural traits in raw data, this makes it difficult to characterise common behavioural features. In addition, current feature representations are often only applicable to some or a limited set of assaults. Second, it's hard to improve the detection model's flexibility in the face of various injection threats. A standardised input structure is necessary for the wide variety of injection behaviours (such as rating behaviours for profile injection attacks and co-visitation behaviours for co-visitation injection attacks). 3) The high demands for detection performance posed by the imbalanced distribution and the sparsity of the original data is another long-standing but unsolved problem. Dimensionality reduction in sparse rating matrices is challenging since it must be done while still preserving relevant information (e.g., well-designed fake profiles). In addition, there are practical constraints on relying on past information to mitigate sparsity. We are energised by the possibilities and challenges presented by this situation: Further improvements to the depiction of rating behaviours (1) are infeasible. It is worthwhile to investigate: (1) how to discover anomalous behaviours from a new angle and how to avoid extracting features directly from the original rating data; and (2) how to mitigate sparsity and unbalanced distribution via prior knowledge in order to reduce the dimensionality of data and the time complexity of algorithm. With this in mind, the value of a methodical approach becomes clear.

2 BACKGROUND AND RELATED WORK

Recommender systems play an increasing role in ecommerce systems such as Amazon, eBay, etc. In reality, they are susceptible to malicious attacks including profile injection attacks and co-visitation injection attacks. Defending these threats has attracted much attention from both academic and industry in the past two decades. Promising results can be obtained from previous efforts, which can be briefly summarized in the following aspects: (1) characterizing rating behaviors between users; (2) exploring the elimination of disturbed data and divide-and-conquer detection strategies; and (3) developing detection approaches for real-world application. In what follows, we introduce related researches focused on the above aspects. 2.1 Evaluation of Rating and Link Behaviors With the development of machine learning and online social network, researchers have begun to investigate detection models from the perspectives of supervised, unsupervised and semi-supervised learning. Naturally, designing representative rating features extracted from the original rating data is a primary task. To this end, on the one hand, Chirita et al. [16] developed statistical metrics to reveal rating patterns of shilling attackers. Then, Burke et al. [7] analyzed diverse features such as generic features and modelspecific features to detect shilling attackers. Shortly after, Williams et al. [17] proposed 5 generic features and 9 typespecific features to characterize the underlying rating behaviors of users. Moreover, Zhang et al. [18] further expanded the feature space from the perspective of Hilbert-Huang Transform according to the novelty- and popularity-based rating series for each user profile. On the other hand, graph-based detection methods for spotting fake accounts or Sybil nodes have been well developed. Investigating attributes of link behaviors in association graphs has also attracted much attention. Cao et al. [19] provided a new technique termed SybilRank based on social graph properties to rank users according to their perceived likelihood of being fake (Sybils). Then, Gong et al. [20] proposed a semi-supervised learning framework to detect Sybil nodes. They also investigated algorithms for both link prediction and attribute inference [21]. Additionally, Dong et al. [22] presented a unified link prediction framework (termed CLP) to deal with the problem of heterogeneous interactions in coupled networks. In particular, they mathematically analyzed the feasibility for the inference of associative links in coupled networks. Link prediction in the target network, however, depends heavily on the structure of both the source and cross networks [22]. In this sense, improving the performance of link prediction may be easily constrained by disturbed information (e.g., an inactive user rated an unpopular item in the user-item cross network). After that, Wang et al. investigated a guilt-by-association method to detect fraudulent users in online social networks [23]. They designed a novel pairwise Markov Random Field to model the joint probability distribution of all random variables based on a directed social graph, which provides unique characteristics for detecting fraudulent users. In addition, Wang et al. [24] developed a novel structure-based method to detect sybils in social networks. They proposed a framework to unify the state of the art RW-based and LBPbased methods and designed a new local rule for sybil detection. Recently, Wang et al. [25] provided a novel collective classification framework. Weight learning and propagation of weighted graph have been well developed. These efforts reveal that rating behaviour attributes extracted from the original rating data and linkage attributes based on behavioural association graphs have great potential to characterize internal relevance of nodes (users). Nevertheless, exploring general and representative characteristics facing with diverse attacks while preserving traditional rating attributes and linkage attributes in order to improve the generalization ability of detection model, is still an open issue.

2.2 Determination of Disturbed Information

Despite promising results relying on the aforementioned rating attributes and linkage attributes, a dense distribution between attackers and some authentic users whose rating details are mimicked by the attackers results in high false alarm rates, which makes a challenging issue for the improvement of detection performance. To address this issue, step-by-step detection frameworks are favorable by researchers. One group of active studies focused on target item analysis [28], [29]. They investigated the distribution of items in order to capture suspicious target items. Specially, disturbed rating profiles can be partly removed in advance. In addition, Xia et al. [30] proposed a detection approach using a dynamic time interval segmentation technique. Recently, Yang et al. [31], [32] investigated the distribution patterns between items and users in order to filter out disturbed data. The above efforts confirm that disturbed data, such as inactive users, unpopular items, etc., can be empirically eliminated before detection. Nevertheless, purely analyzing the local distribution of items and users is insufficient to distinguish target items from popular or unpopular items, especially facing with small attack sizes [32], [33]. How to incorporate global properties such as topological properties of behavior association graph, etc., to enhance the recognition of rating intention is worth further investigating.

2.3 Anomaly Detection for Real Application

The ultimate goal of abnormality detection is to serve real application. Discovering anomalous rating behaviors in online recommender systems has attracted much attention. Initially, Cao et al. [34] developed a semi-supervised detection approach for spotting profile injection attacks. Shortly after, Wu et al. [8], [35] further investigated hybrid detection models to separate shilling attackers (a group of fake users elaborately mimic rating details of few authentic users together in order to mix in their neighbors) from authentic users. Additionally, Gunnemann ¨ et al. [10] casted the basic behavior of users on an item as a latent multivariate autoregressive process and proposed a detection framework to discover interesting findings on real-world data. Gunnemann ¨ et al. [9] also provided a Bayesian model to detect the general rating behavior of users. After that, Zhang et al. [36] investigated a unified framework (termed CBS) for detecting shilling attacks in an eye for an eye manner without being bothered by the details of attacks. They developed a probability propagation mechanism of link behaviors based on an original user-item bipartite graph. The lack of eliminating disturbed information and analyzing target items, however, may lead to higher false alarm rates especially facing with small attack sizes. Moreover, detecting covisitation injection attacks can not be directly implemented using CBS. A careful reading of the literature suggests that rating behavior analysis has considerable potential for discovering abnormalities in reality. But, the generalization and adaptability of detection models are insufficient especially for diverse attacks due to the limited representation of existing behavior features. This work, differing from existing studies: (1) aims to convert abnormality detection into probabilistic inference of link behavior, bypassing the limitation of feature characterization; (2) comprehensively estimates the trustworthiness of both link behaviors and node inherent behaviors of probabilistic graph; (3) remains the advantages of removing disturbed information, probability propagation in CBS, and heterogeneous interaction in CLP in order to reduce the scope of detection; and (4) explores a unified detection framework that is applicable to different attacks.

3 THREAT MODELS

Due to the fundamental vulnerabilities and openness of recommender systems, the rating intention of users and co-visitations between items have been concerned by malicious attackers. In this paper, two types of representative threat models including co-visitation injection attacks focused on co-visitation recommender systems [11] and profile injection attacks (a.k.a., shilling attacks) [7], [12] are implemented in diverse cases. All presented attacks are briefly summarized in Table 1. For co-visitation injection attacks, due to the intention of attackers (i.e., promotion or demotion) and the background knowledge of attackers (or termed bounded resources), we only exploited 4 different attack scenarios in the experiments including promotion attack with high knowledge (PH), promotion attack with medium knowledge (PM), demotion attack with high knowledge (DH), and demotion attack with medium knowledge (DM) [11]. Taking the PH attack for example, attackers select a set of anchor items that can be successfully attacked using bounded resources to maximize the threat. The threat of promotion attack (termed increased probability of top-k user impression, IUI) can be formally defined as, IUI = Pi∈Jit pi , which means that, suppose one target item it is originally among the top-k recommendation list in a set of items which we denote as Iit , this set of items can be enlarged to be Jit after the promotion attack. There into, pi = w_i w_1+w_2+··· ,w_n , where n and wi denote the total number of items and the popularity of i in the past, respectively. Regarding the high knowledge of attack, an attacker can access a co-visitation graph G and a popularity threshold τ used to filter out unpopular items when producing the recommendation lists. For the medium knowledge of attack, comparatively, the attacker can access the popularity of each item and see its recommendation list, but the attacker cannot obtain the number of co-visitations between items nor the popularity threshold [11]. In order to make it appear in the top-k recommendation list of an anchor item j, attackers need to inject mjk fake co-visitations between items it and j. Two conditions need to be satisfied, namely s 0 jit > s0 jkj , wit + mjk ≥ τ, ∀j ∈ Vk, where Vk denotes the set of items whose top-k recommendation lists do not include it. s 0 jit denotes the similarity between items j and it. s 0 jkj represents the similarity between j and the k-th ranked item kj after the attack [11]. The goal of PH attack is to maximize IUI based on bounded resources, which can be formally defined as, max IUI = P j∈Vk aj · pj , s.t. P j∈Vk aj · mjk ≤ m, where m denotes the resource constraint that m items whose recommendation lists were collected by the attackers. aj ∈ {0, 1}, aj = 1 means j is an anchor item. Note that, demotion attacks can be converted to promotion attacks. Due to space limit, we provide the other attack scenarios and briefly describe each of them as shown in Table 1. With regard to profile injection attacks (or termed shilling attacks), shilling attackers empirically insert welldesigned rating profiles into the system in order to manipulate the recommendation or diminish the performance of recommendation on behalf of their goals [7]. The rationale behind shilling attacks is that a user may like an item that his neighbors who have similar preferences. In this paper, we only exploited 4 representative shilling attack models in the experiments, including Love/Hate attack, PUA-NR attack, etc. [26], [27] as described in Table 1. Generally, shilling attackers focus on a targeted item it with the highest score Rmax (called push attacks) or the lowest score Rmin (called nuke attacks). Formally, the attack profiles consist of a d-dimensional vector of ratings, where d is the total number of items in the system. The rating profile is partitioned in three parts from the perspective of item distribution, including selected items IS, filler items IF and target items IT . Concretely, TS receives ratings as specified by the function δ, due to the fact that some attacks require identifying a group of items for special treatment during the attacks. IF is used to control the length of rating vector for mimicking similar rating vectors of neighbors of users who have been concerned by the attackers. The corresponding ratings are added as specified by the function σ. Finally, IT contains one item (termed single-target attacks) or multiple items (termed multi-target attacks). A rating assigned to a target item it is generally set as R_{max} or R_{min}. It is worth emphasizing that ratings assigned to IS or IF can be obtained using the normal distribution of item i or the global normal distribution, i.e., N(R¯ i , σ¯ 2 i) and N(R¯, σ¯ 2) [26]. Specially, power user attacks (PUA) [37], [38], [39] and

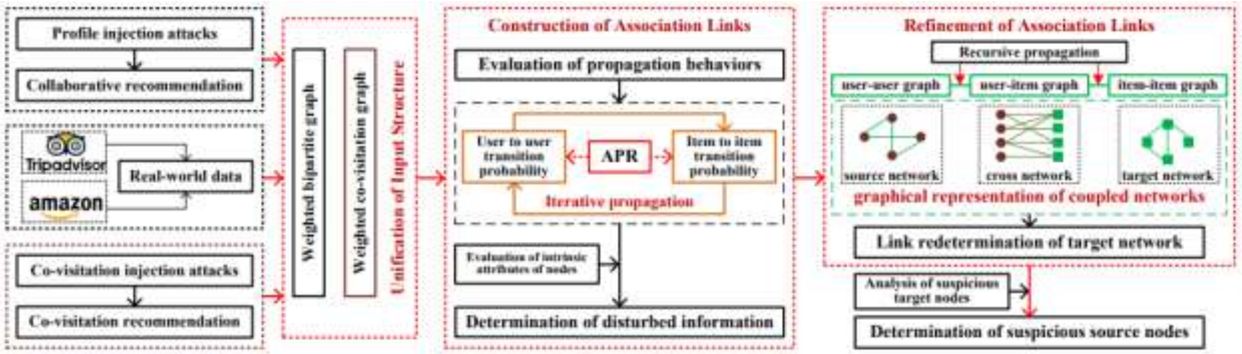
Profile injection attacks
Collaborative recommendation
Tripadvisor
amazon
Real-world data
Co-visitation injection attacks
Co-visitation recommendation
Weighted bipartite graph
Weighted co-visitation graph
Unification of Input Structure
Construction of Association Links
Evaluation of propagation behaviors
User to user transition probability
APR
Item to item transition probability
Iterative propagation
Evaluation of intrinsic attributes of nodes
Determination of disturbed information
Refinement of Association Links
Recursive propagation
user-user graph
user-item graph
item-item graph
source network
cross network
target network
graphical representation of coupled networks
Link redetermination of target network
Analysis of suspicious target nodes
Determination of suspicious source nodes

Fig. 1: The basic framework of the proposed detection approach, which consists of four stages including the unified representation of injection behaviors, the construction of association links, the refinement of association graphs, and the determination of suspicious nodes.

4 THE PROPOSED APPROACH

In this section, we first provide an overview of the proposed approach. Then, we introduce several related definitions. Based on the definitions, we investigate the construction and evaluation of association links to determine disturbed information and prune the concerned networks. Finally, we polish the refined networks to further reduce the scope of detection and infer suspicious nodes. In this paper, we use bold uppercase characters for matrices (e.g., A), bold lowercase characters for vectors (e.g., x), calligraphic fonts for sets (e.g., U), and normal lowercase characters for scalars (e.g., d). Note that, Table 2 describes the notations used throughout the paper.

5 CONCLUSIONS AND FUTURE WORK

Dimensionality reduction and the association representation of sparse rating data has been a long-standing but unsolved challenge when it comes to defending recommender systems against malicious injection assaults. It is easy to slip into the trap of misleading correlation, which might lead to erroneous discovery and insight, if one does not have a firm grasp on the distinctions between real and phoney injection profiles. In this research, we provide a unified detection system for detecting malicious assaults such as co-visitation injection attacks, profile injection attacks, etc. based on a divide-and-conquer approach. Based on empirical evidence, (1) it is possible to reduce the scope of detection and the computational cost by eliminating disturbed data in advance (as determined by user activity, item popularity, and transition probability between users and items); and (2) the detection accuracy can be significantly improved by evaluating the trustworthiness of associative links in coupled networks after the disturbance information has been removed. Future study will investigate parameter sensitivity in a variety of scale-free, small-world, and other synthetic networks. It is also unclear how to create a uniform detection framework to cope with novel threats specific to recommender systems, such as data poisoning attacks against factorization-based collaborative filtering [55], poisoning attacks on graph-based recommender systems [15], etc. In addition, it's important to zero in on the right problems by conducting thorough forensic investigations of anomalous data.

REFERENCES

[1] X. Luo, M. Zhou, Y. Xia, Q. Zhu, A. Ammari, and A. Alabdulwahab, “Generating highly accurate predictions for missing qosdata via aggregating non-negative latent factor models,” IEEE Transactions on Neural Networks and Learning Systems, vol. 27, no. 3, pp. 524–537, 2016.

[2] X. Luo, M. Zhou, S. Li, and M. Shang, “An inherently non-negative latent factor model for high-dimensional and sparse matrices from industrial applications,” IEEE Transactions on Industrial Informatics, 2017.

[3] G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions,” IEEE Transactions on Knowledge and Data Engineering, vol. 17, no. 6, pp. 734–749, 2005.

[4] X. Su and T. Khoshgoftaar, “A survey of collaborative filtering techniques,” Advance in Artificial Intelligence, pp. 1–19, 2009.

[5] J. Calandrino, A. Kilzer, A. Narayanan, E. Felten, and V. Shmatikov, “You might also like: Privacy risks of collaborative filtering,” IEEE Symposium on Security and Privacy (SP), pp. 231–246, 2011.

[6] X. Xing, W. Meng, D. Doozan, A. Snoeren, N. Feamster, and W. Lee, “Take this personally: pollution attacks on personalized services,” USENIX Security, pp. 671–686, 2013.

[7] R. Burke, B. Mobasher, and C. Williams, “Classification features for attack detection in collaborative recommender systems,” International Conference on Knowledge Discovery and Data Mining, pp. 17–20, 2006.

[8] Z. Wu, J. Wu, J. Cao, and D. Tao, “HySAD: A semi-supervised hybrid shilling attack detector for trustworthy product recommendation,” ACM conference on KDD, pp. 985–993, 2012.

[9] S. Gunnemann, N. Gu¨nnemann, and C. Faloutsos, “Detecting anomalies in dynamic rating data: A robust probabilistic model for rating evolution,” In KDD’2014, pp. 841–850, 2014.

[10] N. Gunnemann, S. Gu¨nnemann, and C. Faloutsos, “Robust multivariate autoregression for anomaly detection in dynamic product ratings,” Proceedings of the conference on World Wide Web, pp. 361– 372, 2014.

[11] 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.

[12] I. Gunes, C. Kaleli, A. Bilge, and H. Polat, “Shilling attacks against recommender systems: A comprehensive survey,” Artificial Intelligence Review, vol. 42, no. 4, pp. 1–33, 2012.

[13] Z. Wu, Y. Wang, and J. Cao, “A survey on shilling attack models and detection techniques for recommender systems,” Science China, vol. 59, no. 7, pp. 551–560, 2014.

[14] M. Si and Q. Li, "Shilling attacks against collaborative recommender systems: a review," Artificial Intelligence Review, vol. 9, pp. 1–29, 2018.

[15] M. Fang, G. Yang, N. Gong, and J. Liu, "Poisoning attacks to graph-based recommender systems," ACSAC, arXiv preprint arXiv:1809.04127, 2018.

[16] P. Chirita, W. Nejdl, and C. Zamfir, "Preventing shilling attacks in online recommender systems," In Proceedings of the 7th annual ACM WIDM, pp. 67–74, 2005.

[17] C. A. Williams, B. Mobasher, R. Burke, and R. Bhaumik, "Detecting profile injection attacks in collaborative filtering: a classificationbased approach," Advances in Web Mining and Web Usage Analysis, pp. 167–186, 2007.

[18] F. Zhang and Q. Zhou, "HHT-SVM: An online method for detecting profile injection attacks in collaborative recommender systems," Knowledge-Based Systems, vol. 65, pp. 96–105, 2014.

[19] Q. Cao, M. Sirivianos, X. Yang, and T. Pregueiro, "Aiding the detection of fake accounts in large scale social online services," Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation (NSDI'12), pp. 15–15, 2012.

[20] N. Gong, M. Frank, and P. Mittal, "SybilBelief: A semi-supervised learning approach for structure-based sybil detection," IEEE Transactions on Information Forensics and Security, vol. 9, no. 6, pp. 976–987, 2014.

[21] N. Gong, A. Talwalkar, L. Mackey, L. Huang, E. Shin, E. Stefanov, E. Shi, and D. Song, "Joint link prediction and attribute inference using a social-attribute network," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 5, no. 2, pp. 1–20, 2014.

[22] Y. Dong, J. Zhang, J. Tang, N. V. Chawla, and B. Wang, "Coupledlp: Link prediction in coupled networks," In Proceedings of the TwentyFirst ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'15), pp. 199–208, 2015.

[23] 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 (ICDM), pp.465–474, 2017.

[24] B. Wang, J. Jia, L. Zhang, and N. Gong, "Structure-based Sybil detection in social networks via local rule-based propagation," IEEE Transactions on Network Science and Engineering, vol. 6, no. 3, pp. 523–537, 2019.

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

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

[27] C. E. Seminario and D. C. Wilson, “Attacking item-based recommender systems with power items,” ACM Conference on Recommender Systems, pp. 57–64, 2014.

[28] W. Zhou, Y. S. Koh, J. H. Wen, S. Burki, and G. Dobbie, “Detection of abnormal profiles on group attacks in recommender systems,” Proceedings of the 37th international ACM SIGIR conference on Research on development in information retrieval, vol. 1, pp. 955–958,2014.

[29] W. Zhou, J. Wen, Q. Xiong, M. Gao, and J. Zeng, “SVM-TIA: A shilling attack detection method based on SVM and target item analysis in recommender systems,” Neurocomputing, vol. 210, pp.197–205, 2016.

[30] H. Xia, B. Fang, M. Gao, H. Ma, Y. Tang, and J.Wen, “A novel item anomaly detection approach against shilling attacks in collaborative recommendation systems using the dynamic time interval segmentation technique,” Information Sciences, vol. 306, no. 10, pp.150–165, 2015.

[31] Z. Yang, Q. Sun, Y. Zhang, and B. Zhang, “Uncovering anomalous rating behaviors for rating systems,” Neurocomputing, vol. 308, pp. 205–226, 2018.

[32] Z. Yang, Z. Cai, and X. Guan, “Estimating user behavior toward detecting anomalous ratings in rating systems,” Knowledge-Based Systems, vol. 111, pp. 144–158, 2016.

[33] Z. Yang, Z. Cai, and Y. Yang, “Spotting anomalous ratings for rating systems by analyzing target users and items,” Neurocomputing, vol. 240, pp. 25–46, 2017.

[34] J. Cao, Z. Wu, B. Mao, and Y. Zhang, “Shilling attack detection utilizing semi-supervised learning method for collaborative recommender system,” World Wide Web, vol. 16, pp. 729–748, 2012.

[35] Z. Wu, J. Cao, B. Mao, and Y. Wang, “Semi-SAD: Applying semisupervised learning to shilling attack detection,” ACM Conference on Recommender Systems, vol. 6, no. 6, pp. 289–292, 2011.

[36] Y. Zhang, Y. Tan, M. Zhang, Y. Liu, T. Chua, and S. Ma, “Catch the black sheep: Unified framework for shilling attack detection based on fraudulent action propagation,” Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015), pp. 2408–2414, 2015.

[37] C. E. Seminario, “Accuracy and robustness impacts of power user attacks on collaborative recommender systems,” ACM Conference on Recommender Systems, pp. 447–450, 2013.

[38] D. C. Wilson and C. E. Seminario, “Evil twins: Modeling power users in attacks on recommender systems,” User Modeling, Adaptation, and Personalization, pp. 231–242, 2014. [39] D. Wilson and C. E.

Seminario, "Mitigating power user attacks on a user-based collaborative recommender system," Association for the Advancement of Artificial Intelligence, pp. 513–318, 2015.

[40] L. Adamic and B. Huberman, "Power-law distribution of the world wide web," Science, vol. 287, no. 5461, p. 2115, 2000.

[41] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins, "Propagation of trust and distrust," Proceedings of the 13th international conference on World Wide Web (WWW'04), pp. 403–412, 2004.

[42] L. Adamic and E. Adar, "Friends and neighbors on the web," Social Networks, vol. 25, no. 3, pp. 211–230, 2003.

[43] T. Haveliwala, "Topic-sensitive pagerank: A context-sensitive ranking algorithm for web search," IEEE Transactions on Knowledge and Data Engineering, vol. 15, no. 4, pp. 784–796, 2003.

[44] R. Lichtenwalter, J. Lussier, and N. Chawla, "New perspectives and methods in link prediction," Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD'10), pp. 243–252, 2010.

[45] Y. Sun, J. Han, C. C. Aggarwal, and N. V. Chawla, "When will it happen?: relationship prediction in heterogeneous information networks," Proceedings of the fifth ACM international conference on Web search and data mining (WSDM'12), pp. 663–672, 2012.

[46] J. E. Hopcroft, T. Lou, and J. Tang, "Who will follow you back? reciprocal relationship prediction," In Proceedings of the Twenty Conference on Information and Knowledge Management, pp. 1137– 1146, 2011.

[47] N. Chandgotia, "Generalisation of the hammersley-clifford theorem on bipartite graphs," Transactions of the American Mathematical Society, vol. 369, pp. 7107–7137, 2017.

[48] J. Tang, T. Lou, and J. Kleinberg, "Inferring social ties across heterogenous networks," Proceedings of the fifth ACM international conference on Web search and data mining, pp. 743–752, 2012.

[49] Y. Dong, N. V. Chawla, J. Tang, and Y. Yang, "User modeling on demographic attributes in big mobile social networks," ACM Transactions on Information Systems, vol. 35, no. 4, p. 35, 2017.

[50] D. Nowell and J. Kleinberg, "The link prediction problem for social networks," Journal of the American Society for Information Science and Technology (JASIST), vol. 58, no. 7, pp. 1019–1031, 2007.

[51] J. Benesty, J. Chen, and Y. Huang, "On the importance of the pearson correlation coefficient in noise reduction," IEEE Transactions on Audio Speech and Language Processing, vol. 16, no. 4, pp. 757–765, 2008.

[52] J. McAuley and J. Leskovec, “Hidden factors and hidden topics: understanding rating dimensions with review text,” ACM Conference on Recommender Systems (RecSys), pp. 165–172, 2013.

[53] C. Chung, P. Hsu, and S. Huang, “βP: A novel approach to filter out malicious rating profiles from recommender systems,” Decision Support Systems, vol. 55, no. 1, pp. 314–325, 2013.

[54] N. Hurley, Z. Cheng, and M. Zhang, “Statistical attack detection,” Proceedings of the third ACM conference on Recommender systems, pp. 149–156, 2009.

[55] B. Li, Y. Wang, A. Singh, and Y. Vorobeychik, “Data poisoning attacks on factorization-based collaborative filtering,” 29th Conference on Neural Information Processing Systems (NIPS 2016), pp. 1–13, 2016.