
A black-box attack model for visually-aware recommender systems

类型 会议论文

作者 Rami Cohen

作者 Oren Sar Shalom

作者 Dietmar Jannach

作者 Amihood Amir

日期 2021

馆藏目录 Google Scholar

页码 94–102

会议论文集标题 Proceedings of the 14th ACM International Conference on Web Search and Data Mining

添加日期 2022/11/16 下午7:26:43

修改日期 2022/11/17 下午6:42:28

标签:

Black-box, Visual

附件

- Full Text
- Snapshot

A CNN-based Hybrid Model and Architecture for Shilling Attack Detection

类型 会议论文

作者 Mahsa Ebrahimian

作者 Rasha Kashef

日期 2021

馆藏目录 Google Scholar

出版社 IEEE

页码 1–7

会议论文集标题 2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)

添加日期 2022/11/16 下午8:16:59

修改日期 2022/11/16 下午8:32:40

标签:

Computational modeling, Conferences, Data models, Benchmark testing, Compatibility, Computer architecture, Deep learning, Divergence Criterion, Neural Networks, Predictive models, Recommendation systems, Shilling attacks

附件

- IEEE Xplore Abstract Record
- IEEE Xplore Full Text PDF
- Snapshot

A genre trust model for defending shilling attacks in recommender systems

类型 期刊文章

作者 Li Yang

作者 Xinxin Niu

日期 2021

馆藏目录 Google Scholar

其它 Publisher: Springer

页码 1–14

期刊 Complex & Intelligent Systems

添加日期 2022/11/16 下午7:56:31

修改日期 2022/11/16 下午8:32:41

附件

- Full Text
- Yang 和 Niu - 2021 - A genre trust model for defending shilling attacks.pdf

A Regression Framework to Interpret the Robustness of Recommender Systems Against Shilling Attacks.

类型 会议论文

作者 Yashar Deldjoo

作者 Tommaso Di Noia

作者 Eugenio Di Sciascio

作者 Felice Antonio Merra

日期 2021

馆藏目录 Google Scholar

会议论文集标题 IIR

添加日期 2022/11/16 下午7:56:31

修改日期 2022/11/16 下午7:56:31

附件

- Full Text

A survey of attack detection approaches in collaborative filtering recommender systems

类型 期刊文章

作者 Fatemeh Rezaimehr

作者 Chitra Dadkhah

日期 2021

馆藏目录 Google Scholar

其它 Publisher: Springer

卷次 54

页码 2011–2066

期刊 Artificial Intelligence Review

期号 3

添加日期 2022/11/16 下午7:28:34

修改日期 2022/11/16 下午7:28:34

附件

- Snapshot

A survey on adversarial attack in the age of artificial intelligence

类型 期刊文章

作者 Zixiao Kong

作者 Jingfeng Xue

作者 Yong Wang

作者 Lu Huang

作者 Zequn Niu

作者 Feng Li

日期 2021

馆藏目录 Google Scholar

其它 Publisher: Hindawi

卷次 2021

期刊 Wireless Communications and Mobile Computing

添加日期 2022/11/16 下午7:23:52

修改日期 2022/11/16 下午7:23:52

附件

○ Full Text

A Survey on Adversarial Recommender Systems: From Attack/Defense Strategies to Generative Adversarial Networks

类型 期刊文章

作者 Yashar Deldjoo

作者 Tommaso Di Noia

作者 Felice Antonio Merra

摘要 Latent-factor models (LFM) based on collaborative filtering (CF), such as matrix factorization (MF) and deep CF methods, are widely used in modern recommender systems (RS) due to their excellent performance and recommendation accuracy. However, success has been accompanied with a major new arising challenge: Many applications of machine learning (ML) are adversarial in nature [146]. In recent years, it has been shown that these methods are vulnerable to adversarial examples, i.e., subtle but non-random perturbations designed to force recommendation models to produce erroneous outputs. The goal of this survey is two-fold: (i) to present recent advances on adversarial machine learning (AML) for the security of RS (i.e., attacking and defense recommendation models) and (ii) to show another successful application of AML in generative adversarial networks (GANs) for generative applications, thanks to their ability for learning (high-dimensional) data distributions. In this survey, we provide an exhaustive literature review of 76 articles published in major RS and ML journals and conferences. This review serves as a reference for the RS community working on the security of RS or on generative models using GANs to improve their quality.

日期 2022-03-31

语言 en

短标题 A Survey on Adversarial Recommender Systems

馆藏目录 DOI.org (Crossref)

URL <https://dl.acm.org/doi/10.1145/3439729>

访问时间 2022/11/15 上午2:35:53

卷次 54

页码 1-38

期刊 ACM Computing Surveys

DOI 10.1145/3439729

期号 2

刊名缩写 ACM Comput. Surv.

ISSN 0360-0300, 1557-7341

添加日期 2022/11/15 上午2:35:53

修改日期 2022/11/16 下午8:32:45

附件

○ Deldjoo 等 - 2022 - A Survey on Adversarial Recommender Systems From .pdf

A Survey on Adversarial Recommender Systems: From Attack/Defense Strategies to Generative Adversarial Networks: ACM Computing Surveys: Vol 54, No 2

类型 网页

URL <https://dl.acm.org/doi/abs/10.1145/3439729>

访问时间 2022/11/16 下午6:44:08

添加日期 2022/11/16 下午6:44:08

修改日期 2022/11/16 下午6:44:08

附件

- 全文

Adversarial Attacks and Defenses in Deep Learning: from a Perspective of Cybersecurity

类型 期刊文章

作者 Shuai Zhou

作者 Chi Liu

作者 Dayong Ye

作者 Tianqing Zhu

作者 Wanlei Zhou

作者 Philip S. Yu

日期 2021

短标题 Adversarial Attacks and Defenses in Deep Learning

馆藏目录 Google Scholar

其它 Publisher: ACM New York, NY

期刊 ACM Computing Surveys (CSUR)

添加日期 2022/11/16 下午6:49:21

修改日期 2022/11/16 下午8:32:48

附件

- Snapshot
- Zhou 等 - 2022 - Adversarial Attacks and Defenses in Deep Learning.pdf

Adversarial Attacks to API Recommender Systems: Time to Wake Up and Smell the Coffee?

类型 会议论文

作者 Phuong T. Nguyen
作者 Claudio Di Sipio
作者 Juri Di Rocco
作者 Massimiliano Di Penta
作者 Davide Di Ruscio

摘要 Recommender systems in software engineering provide developers with a wide range of valuable items to help them complete their tasks. Among others, API recommender systems have gained momentum in recent years as they became more successful at suggesting API calls or code snippets. While these systems have proven to be effective in terms of prediction accuracy, there has been less attention for what concerns such recommenders' resilience against adversarial attempts. In fact, by crafting the recommenders' learning material, e.g., data from large open-source software (OSS) repositories, hostile users may succeed in injecting malicious data, putting at risk the software clients adopting API recommender systems. In this paper, we present an empirical investigation of adversarial machine learning techniques and their possible influence on recommender systems. The evaluation performed on three state-of-the-art API recommender systems reveals a worrying outcome: all of them are not immune to malicious data. The obtained result triggers the need for effective countermeasures to protect recommender systems against hostile attacks disguised in training data.

日期 2021-11

短标题 Adversarial Attacks to API Recommender Systems

馆藏目录 IEEE Xplore

其它 ISSN: 2643-1572

页码 253-265

会议论文集标题 2021 36th IEEE/ACM International Conference on Automated Software Engineering (ASE)

会议名称 2021 36th IEEE/ACM International Conference on Automated Software Engineering (ASE)

DOI 10.1109/ASE51524.2021.9678946

添加日期 2022/11/16 下午8:30:15

修改日期 2022/11/16 下午8:30:15

标签:

Recommender systems, Task analysis, Adversarial Attacks, Adversarial machine learning, Adversarial Machine Learning, API Mining, Codes, Open source software, Software engineering, Training data

附件

- IEEE Xplore Full Text PDF

Adversarial Attacks to API Recommender Systems: Time to Wake Up and Smell the Coffee

类型 会议论文

作者 Phuong T. Nguyen

作者 Claudio Di Sipio

作者 Juri Di Rocco

作者 Massimiliano Di Penta

作者 Davide Di Ruscio

日期 2021

短标题 Adversarial Attacks to API Recommender Systems

馆藏目录 Google Scholar

出版社 IEEE

页码 253–265

会议论文集标题 2021 36th IEEE/ACM International Conference on Automated Software Engineering (ASE)

添加日期 2022/11/16 下午6:48:03

修改日期 2022/11/16 下午6:48:03

附件

- Snapshot

Adversarial Recommendation: Attack of the Learned Fake Users

类型 预印本

作者 Konstantina Christakopoulou

作者 Arindam Banerjee

摘要 Can machine learning models for recommendation be easily fooled? While the question has been answered for hand-engineered fake user profiles, it has not been explored for machine learned adversarial attacks. This paper attempts to close this gap. We propose a framework for generating fake user profiles which, when incorporated in the training of a recommendation system, can achieve an adversarial intent, while remaining indistinguishable from real user profiles. We formulate this procedure as a repeated general-sum game between two players: an oblivious recommendation system \mathcal{R} and an adversarial fake user generator \mathcal{A} with two goals: (G1) the rating distribution of the fake users needs to be close to the real users, and (G2) some objective $f_{\mathcal{A}}$ encoding the attack intent, such as targeting the top-K recommendation quality of \mathcal{R} for a subset of users, needs to be optimized. We propose a learning framework to achieve both goals, and offer extensive experiments considering multiple types of attacks highlighting the vulnerability of recommendation systems.

日期 2018-09-21

短标题 Adversarial Recommendation

馆藏目录 arXiv.org

URL <http://arxiv.org/abs/1809.08336>

访问时间 2022/11/17 上午12:25:26

其它 arXiv:1809.08336 [cs, stat]
DOI 10.48550/arXiv.1809.08336
仓库 arXiv
存档ID arXiv:1809.08336
添加日期 2022/11/17 上午12:25:26
修改日期 2022/11/17 上午12:25:26

标签:

Computer Science - Information Retrieval, Computer Science - Machine Learning, Statistics - Machine Learning

附件

- arXiv Fulltext PDF
- arXiv.org Snapshot

Adversarial recommender systems: Attack, defense, and advances

类型 图书章节
作者 Vito Walter Anelli
作者 Yashar Deldjoo
作者 Tommaso DiNoia
作者 Felice Antonio Merra
日期 2022
短标题 Adversarial recommender systems
馆藏目录 Google Scholar
出版社 Springer
页码 335–379
书名 Recommender systems handbook
添加日期 2022/11/16 下午6:47:11
修改日期 2022/11/16 下午6:47:11

附件

- Snapshot

An adaptive RNN algorithm to detect shilling attacks for online products in hybrid recommender system

类型 期刊文章
作者 Akanksha Bansal Chopra

作者 Veer Sain Dixit

摘要 Recommender system (RS) depends on the thoughts of numerous users to predict the favourites of potential consumers. RS is vulnerable to malicious information. Unsuitable products can be offered to the user by injecting a few unscrupulous “shilling” profiles like push and nuke attacks into the RS. Injection of these attacks results in the wrong recommendation for a product. The aim of this research is to develop a framework that can be widely utilized to make excellent recommendations for sales growth. This study uses the methodology that presents an enhanced clustering algorithm named as modified density peak clustering algorithm on the consumer review dataset to ensure a well-formed cluster. An improved recurrent neural network algorithm is proposed to detect these attacks in hybrid RS, which uses the content-based RS and collaborative filtering RS. The results are compared with other state of the art algorithms. The proposed method is more suitable for E-commerce applications where the number of customers and products grows rapidly.

日期 2022-10-31

语言 en

馆藏目录 DOI.org (Crossref)

URL <https://www.degruyter.com/document/doi/10.1515/jisys-2022-1023/html>

访问时间 2022/11/16 下午8:16:34

卷次 31

页码 1133-1149

期刊 Journal of Intelligent Systems

DOI 10.1515/jisys-2022-1023

期号 1

ISSN 2191-026X

添加日期 2022/11/16 下午8:16:35

修改日期 2022/11/16 下午8:16:35

附件

- Chopra 和 Dixit - 2022 - An adaptive RNN algorithm to detect shilling attac.pdf

An adaptive RNN algorithm to detect shilling attacks for online products in hybrid recommender system

类型 网页

URL <https://www.degruyter.com/document/doi/10.1515/jisys-2022-1023/html>

访问时间 2022/11/16 下午8:15:24

添加日期 2022/11/16 下午8:15:24

修改日期 2022/11/16 下午8:15:24

附件

- 全文
- An adaptive RNN algorithm to detect shilling attacks for online products in hybrid recommender system

An Empirical Analysis of Collaborative Recommender Systems Robustness to Shilling Attacks

类型 期刊文章

作者 Anu Shrestha

作者 Francesca Spezzano

作者 Maria Soledad Pera

日期 2021

馆藏目录 Google Scholar

添加日期 2022/11/16 下午6:49:21

修改日期 2022/11/16 下午8:32:52

附件

- Full Text
- Snapshot

An Unsupervised Approach for Detecting Group Shilling Attacks in Recommender Systems Based on Topological Potential and Group Behaviour Features

类型 期刊文章

作者 Hongyun Cai

作者 Fuzhi Zhang

编辑 Rutvij Jhaveri

摘要 To protect recommender systems against shilling attacks, a variety of detection methods have been proposed over the past decade. However, these methods focus mainly on individual features and rarely consider the lockstep behaviours among attack users, which suffer from low precision in detecting group shilling attacks. In this work, we propose a three-stage detection method based on strong lockstep behaviours among group members and group behaviour features for detecting group shilling attacks. First, we construct a weighted user relationship graph by combining direct and indirect collusive degrees between users. Second, we find all dense subgraphs in the user relationship graph to generate a set of suspicious groups by introducing a topological potential method. Finally, we use a clustering method to detect shilling groups by extracting group behaviour features. Extensive experiments on the Netflix and sampled Amazon review datasets show that the proposed approach is effective for detecting group shilling attacks in recommender systems, and the F1-measure on two datasets can reach over 99 percent and 76 percent, respectively.

日期 2021-9-27
语言 en
馆藏目录 DOI.org (Crossref)
URL <https://www.hindawi.com/journals/scn/2021/2907691/>
访问时间 2022/11/16 下午8:02:52
卷次 2021
页码 1-18
期刊 Security and Communication Networks
DOI 10.1155/2021/2907691
刊名缩写 Security and Communication Networks
ISSN 1939-0122, 1939-0114
添加日期 2022/11/16 下午8:02:52
修改日期 2022/11/16 下午8:02:52

附件

- Cai 和 Zhang - 2021 - An Unsupervised Approach for Detecting Group Shill.pdf

Attacking black-box recommendations via copying cross-domain user profiles

类型 会议论文
作者 Wenqi Fan
作者 Tyler Derr
作者 Xiangyu Zhao
作者 Yao Ma
作者 Hui Liu
作者 Jianping Wang
作者 Jiliang Tang
作者 Qing Li
日期 2021
馆藏目录 Google Scholar
出版社 IEEE
页码 1583–1594
会议论文集标题 2021 IEEE 37th International Conference on Data Engineering (ICDE)
添加日期 2022/11/16 下午6:45:50
修改日期 2022/11/16 下午6:45:50

附件

- Full Text
- Snapshot

Attacking recommender systems with plausible profile

类型 期刊文章

作者 Xuxin Zhang

作者 Jian Chen

作者 Rui Zhang

作者 Chen Wang

作者 Ling Liu

日期 2021

馆藏目录 Google Scholar

其它 Publisher: IEEE

卷次 16

页码 4788–4800

期刊 IEEE Transactions on Information Forensics and Security

添加日期 2022/11/16 下午6:48:03

修改日期 2022/11/16 下午8:32:53

附件

- Full Text
- Snapshot

Black-Box Attacks on Sequential Recommenders via Data-Free Model Extraction

类型 会议论文

作者 Zhenrui Yue

作者 Zhankui He

作者 Huimin Zeng

作者 Julian McAuley

摘要 We investigate whether model extraction can be used to ‘steal’ the weights of sequential recommender systems, and the potential threats posed to victims of such attacks. This type of risk has attracted attention in image and text classification, but to our knowledge not in recommender systems. We argue that sequential recommender systems are subject to unique vulnerabilities due to the specific autoregressive regimes used to train them. Unlike many existing recommender attackers, which assume the dataset used to train the victim model is exposed to attackers, we consider a data-free setting, where training data are not accessible. Under this setting, we propose an API-based model extraction method via limited-budget synthetic data generation and knowledge distillation. We investigate state-of-the-art models for sequential recommendation and show their vulnerability under model extraction and downstream attacks. We perform attacks in two stages. (1) Model extraction: given different types of synthetic data and their labels retrieved from a black-box recommender, we extract the black-box model to a

white-box model via distillation. (2) Downstream attacks: we attack the black-box model with adversarial samples generated by the white-box recommender. Experiments show the effectiveness of our data-free model extraction and downstream attacks on sequential recommenders in both profile pollution and data poisoning settings.

日期 2021-09-13

语言 en

馆藏目录 DOI.org (Crossref)

URL <https://dl.acm.org/doi/10.1145/3460231.3474275>

访问时间 2022/11/17 上午12:33:38

地点 Amsterdam Netherlands

出版社 ACM

ISBN 978-1-4503-8458-2

页码 44-54

会议论文集标题 Fifteenth ACM Conference on Recommender Systems

会议名称 RecSys '21: Fifteenth ACM Conference on Recommender Systems

DOI 10.1145/3460231.3474275

添加日期 2022/11/17 上午12:33:38

修改日期 2022/11/17 上午12:33:38

附件

- Yue 等 - 2021 - Black-Box Attacks on Sequential Recommenders via D.pdf

Black-Box Attacks on Sequential Recommenders via Data-Free Model Extraction | Proceedings of the 15th ACM Conference on Recommender Systems

类型 网页

URL <https://dl.acm.org/doi/abs/10.1145/3460231.3474275>

访问时间 2022/11/17 上午12:30:37

添加日期 2022/11/17 上午12:30:37

修改日期 2022/11/17 上午12:30:37

附件

- 全文
- Black-Box Attacks on Sequential Recommenders via Data-Free Model Extraction | Proceedings of the 15th ACM Conference on Recommender Systems

Comprehensive Privacy Analysis on Federated Recommender System against Attribute Inference Attacks

类型 期刊文章

作者 Shijie Zhang

作者 Hongzhi Yin

日期 2022

馆藏目录 Google Scholar

期刊 arXiv preprint arXiv:2205.11857

添加日期 2022/11/16 下午7:58:52

修改日期 2022/11/16 下午7:58:52

附件

- Full Text
- Snapshot

Data Poisoning Attack against Recommender System Using Incomplete and Perturbed Data

类型 会议论文

作者 Hengtong Zhang

作者 Changxin Tian

作者 Yaliang Li

作者 Lu Su

作者 Nan Yang

作者 Wayne Xin Zhao

作者 Jing Gao

日期 2021

馆藏目录 Google Scholar

页码 2154–2164

会议论文集标题 Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining

添加日期 2022/11/16 下午7:28:34

修改日期 2022/11/16 下午7:28:34

附件

- Full Text
- Snapshot

Data poisoning attacks on neighborhood-based recommender systems

类型 期刊文章

作者 Liang Chen

作者 Yangjun Xu

作者 Fenfang Xie

作者 Min Huang

作者 Zibin Zheng

日期 2021

馆藏目录 Google Scholar

其它 Publisher: Wiley Online Library

卷次 32

页码 e3872

期刊 Transactions on Emerging Telecommunications Technologies

期号 6

添加日期 2022/11/16 下午6:45:50

修改日期 2022/11/16 下午6:45:50

附件

○ Full Text

○ Snapshot

Data Poisoning Attacks on Stochastic Bandits

类型 会议论文

作者 Fang Liu

作者 Ness Shroff

摘要 Stochastic multi-armed bandits form a class of online learning problems that have important applications in online recommendation systems, adaptive medical treatment, and many others. Even though potential attacks against these learning algorithms may hijack their behavior, causing catastrophic loss in real-world applications, little is known about adversarial attacks on bandit algorithms. In this paper, we propose a framework of offline attacks on bandit algorithms and study convex optimization based attacks on several popular bandit algorithms. We show that the attacker can force the bandit algorithm to pull a target arm with high probability by a slight manipulation of the rewards in the data. Then we study a form of online attacks on bandit algorithms and propose an adaptive attack strategy against any bandit algorithm without the knowledge of the bandit algorithm. Our adaptive attack strategy can hijack the behavior of the bandit algorithm to suffer a linear regret with only a logarithmic cost to the attacker. Our results demonstrate a significant security threat to stochastic bandits.

日期 2019-05-24

语言 en

馆藏目录 proceedings.mlr.press

URL <https://proceedings.mlr.press/v97/liu19e.html>

访问时间 2022/11/17 上午12:37:04

其它 ISSN: 2640-3498

出版社 PMLR

页码 4042-4050

会议论文集标题 Proceedings of the 36th International Conference on Machine Learning

会议名称 International Conference on Machine Learning

添加日期 2022/11/17 上午12:37:04

修改日期 2022/11/17 上午12:37:04

附件

- Full Text PDF
- Supplementary PDF

Data Poisoning Attacks on Stochastic Bandits

类型 网页

URL <http://proceedings.mlr.press/v97/liu19e.html>

访问时间 2022/11/17 上午12:36:52

添加日期 2022/11/17 上午12:36:52

修改日期 2022/11/17 上午12:36:52

附件

- 全文
- Data Poisoning Attacks on Stochastic Bandits

Data poisoning attacks to deep learning based recommender systems

类型 期刊文章

作者 Hai Huang

作者 Jiaming Mu

作者 Neil Zhenqiang Gong

作者 Qi Li

作者 Bin Liu

作者 Mingwei Xu

日期 2021

短标题 数据中毒攻击数据深度学习为基础的推荐系统

馆藏目录 Google Scholar

期刊 arXiv preprint arXiv:2101.02644

添加日期 2022/11/16 下午6:45:50

修改日期 2022/11/17 上午12:37:33

附件

- Full Text
- Snapshot

Debiasing Learning for Membership Inference Attacks Against Recommender Systems

类型 会议论文
作者 Zihan Wang
作者 Na Huang
作者 Fei Sun
作者 Pengjie Ren
作者 Zhumin Chen
作者 Hengliang Luo
作者 Maarten de Rijke
作者 Zhaochun Ren
日期 2022

馆藏目录 Google Scholar

页码 1959–1968

会议论文集标题 Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining

添加日期 2022/11/16 下午7:34:14

修改日期 2022/11/16 下午7:34:14

附件

- Full Text
- Snapshot

Deep Model Poisoning Attack on Federated Learning

类型 期刊文章
作者 Xingchen Zhou
作者 Ming Xu
作者 Yiming Wu
作者 Ning Zheng

摘要 Federated learning is a novel distributed learning framework, which enables thousands of participants to collaboratively construct a deep learning model. In order to protect confidentiality of the training data, the shared information between server and participants are only limited to model parameters. However, this setting is vulnerable to model poisoning attack, since the participants have permission to modify the model parameters. In this paper, we perform systematic investigation for such threats in federated learning and propose a novel optimization-based model

poisoning attack. Different from existing methods, we primarily focus on the effectiveness, persistence and stealth of attacks. Numerical experiments demonstrate that the proposed method can not only achieve high attack success rate, but it is also stealthy enough to bypass two existing defense methods.

日期 2021-03-14
语言 en
馆藏目录 DOI.org (Crossref)
URL <https://www.mdpi.com/1999-5903/13/3/73>
访问时间 2022/11/17 下午6:00:53
卷次 13
页码 73
期刊 Future Internet
DOI 10.3390/fi13030073
期号 3
刊名缩写 Future Internet
ISSN 1999-5903
添加日期 2022/11/17 下午6:00:53
修改日期 2022/11/17 下午6:00:53

附件

- Zhou 等 - 2021 - Deep Model Poisoning Attack on Federated Learning.pdf

Defending a Music Recommender Against Hubness-Based Adversarial Attacks

类型 期刊文章
作者 Katharina Hoedt
作者 Arthur Flexer
作者 Gerhard Widmer
日期 2022
馆藏目录 Google Scholar
期刊 arXiv preprint arXiv:2205.12032
添加日期 2022/11/16 下午7:58:52
修改日期 2022/11/16 下午7:58:52

附件

- Full Text
- Snapshot

Defending Substitution-Based Profile Pollution Attacks on Sequential Recommenders

类型 会议论文

作者 Zhenrui Yue

作者 Huimin Zeng

作者 Ziyi Kou

作者 Lanyu Shang

作者 Dong Wang

日期 2022

馆藏目录 Google Scholar

页码 59–70

会议论文集标题 Proceedings of the 16th ACM Conference on Recommender Systems

添加日期 2022/11/16 下午8:18:26

修改日期 2022/11/16 下午8:18:26

附件

- Full Text
- Snapshot

Detecting Group Shilling Attacks In Online Recommender Systems

类型 期刊文章

作者 B. Sharmila

作者 D. Narmada

作者 Krishna Keerthana

作者 K. L. Mounika

日期 2021

馆藏目录 Google Scholar

卷次 12

期刊 journal of engineering science

期号 05

添加日期 2022/11/16 下午7:54:32

修改日期 2022/11/16 下午7:54:32

附件

- Full Text

Detecting group shilling attacks in recommender systems based on maximum dense subtensor mining

类型 会议论文

作者 Hongtao Yu

作者 Haihong Zheng

作者 Yishu Xu

作者 Ru Ma

作者 Dingli Gao

作者 Fuzhi Zhang

日期 2021

馆藏目录 Google Scholar

出版社 IEEE

页码 644–648

会议论文集标题 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)

添加日期 2022/11/16 下午7:56:31

修改日期 2022/11/17 下午5:33:25

标签:

Conferences, Data models, dual-input convolutional neural network, Feature extraction, Fuses, group shilling attacks, Knowledge engineering, maximum dense subtensor mining, recommender systems, Tensors, Time series analysis

附件

- IEEE Xplore Full Text PDF
- Snapshot

Detecting shilling groups in online recommender systems based on graph convolutional network

类型 期刊文章

作者 Shilei Wang

作者 Peng Zhang

作者 Hui Wang

作者 Hongtao Yu

作者 Fuzhi Zhang

日期 2022

短标题 在线检测先令团体推荐系统基于图像卷积网络

馆藏目录 Google Scholar

其它 Publisher: Elsevier

卷次 59

页码 103031

期刊 Information Processing & Management

期号 5

添加日期 2022/11/16 下午8:11:12

修改日期 2022/11/17 下午5:40:12

附件

- Snapshot
- Wang 等 - 2022 - Detecting shilling groups in online recommender sy.pdf

Detecting shilling groups in recommender systems based on hierarchical topic model

类型 会议论文

作者 Shilei Wang

作者 Hui Wang

作者 Hongtao Yu

作者 Fuzhi Zhang

日期 2021

馆藏目录 Google Scholar

出版社 IEEE

页码 832–837

会议论文集标题 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)

添加日期 2022/11/16 下午7:58:52

修改日期 2022/11/17 下午5:40:13

标签:

Artificial intelligence, Clustering, Clustering algorithms, Computational modeling, Computer applications, Conferences, Group shilling attacks, Hierarchical topic model, Natural language processing, Optics, Recommender systems, Shilling group detection

附件

- IEEE Xplore Abstract Record
- IEEE Xplore Full Text PDF
- Snapshot

Detection of shilling attack in recommender system for YouTube video statistics using machine learning techniques

类型 期刊文章

作者 Shalli Rani

作者 Manpreet Kaur
作者 Munish Kumar
作者 Vinayakumar Ravi
作者 Uttam Ghosh
作者 Jnyana Ranjan Mohanty
日期 2021

馆藏目录 Google Scholar

其它 Publisher: Springer

页码 1–13

期刊 Soft Computing

添加日期 2022/11/16 下午6:48:03

修改日期 2022/11/16 下午6:48:03

附件

- Snapshot

Detection of Trust Shilling Attacks in Recommender Systems

类型 期刊文章

作者 Xian CHEN

作者 Xi DENG

作者 Chensen HUANG

作者 Hyoseop SHIN

日期 2022

馆藏目录 Google Scholar

其它 Publisher: The Institute of Electronics, Information and Communication Engineers

卷次 105

页码 1239–1242

期刊 IEICE TRANSACTIONS on Information and Systems

期号 6

添加日期 2022/11/16 下午7:54:32

修改日期 2022/11/16 下午7:54:32

附件

- Full Text
- Snapshot

FedAttack: Effective and Covert Poisoning Attack on Federated Recommendation via Hard Sampling

类型 期刊文章

作者 Chuhan Wu

作者 Fangzhao Wu

作者 Tao Qi

作者 Yongfeng Huang

作者 Xing Xie

日期 2022

短标题 FedAttack

馆藏目录 Google Scholar

期刊 arXiv preprint arXiv:2202.04975

添加日期 2022/11/16 下午7:54:32

修改日期 2022/11/16 下午7:54:32

附件

- Full Text
- Snapshot

FedRecAttack: Model Poisoning Attack to Federated Recommendation

类型 预印本

作者 Dazhong Rong

作者 Shuai Ye

作者 Ruoyan Zhao

作者 Hon Ning Yuen

作者 Jianhai Chen

作者 Qinming He

摘要 Federated Recommendation (FR) has received considerable popularity and attention in the past few years. In FR, for each user, its feature vector and interaction data are kept locally on its own client thus are private to others. Without the access to above information, most existing poisoning attacks against recommender systems or federated learning lose validity. Benefiting from this characteristic, FR is commonly considered fairly secured. However, we argue that there is still possible and necessary security improvement could be made in FR. To prove our opinion, in this paper we present FedRecAttack, a model poisoning attack to FR aiming to raise the exposure ratio of target items. In most recommendation scenarios, apart from private user-item interactions (e.g., clicks, watches and purchases), some interactions are public (e.g., likes, follows and comments). Motivated by this point, in FedRecAttack we make use of the public interactions to approximate users' feature vectors, thereby attacker can generate poisoned gradients accordingly and control malicious users to upload the poisoned gradients in a well-designed way. To evaluate the effectiveness and side effects of FedRecAttack, we conduct extensive experiments on three real-world datasets of different sizes from two completely different scenarios. Experimental results demonstrate that our proposed FedRecAttack achieves the state-of-the-art effectiveness while its side effects are negligible. Moreover, even with small

proportion (3%) of malicious users and small proportion (1%) of public interactions, FedRecAttack remains highly effective, which reveals that FR is more vulnerable to attack than people commonly considered.

日期 2022-10-13

语言 en

短标题 FedRecAttack

馆藏目录 arXiv.org

URL <http://arxiv.org/abs/2204.01499>

访问时间 2022/11/15 上午2:35:49

其它 arXiv:2204.01499 [cs]

仓库 arXiv

存档ID arXiv:2204.01499

添加日期 2022/11/15 上午2:35:49

修改日期 2022/11/16 下午8:33:17

标签:

Computer Science - Cryptography and Security, Computer Science - Machine Learning

笔记:

Comment: This paper has been accepted by IEEE International Conference on Data Engineering 2022 (Second Research Round)

Comment: This paper has been accepted by IEEE International Conference on Data Engineering 2022 (Second Research Round)

附件

- arXiv.org Snapshot
- Rong 等 - 2022 - FedRecAttack Model Poisoning Attack to Federated .pdf

Fight Fire with Fire: Towards Robust Recommender Systems via Adversarial Poisoning Training

类型 会议论文

作者 Chenwang Wu

作者 Defu Lian

作者 Yong Ge

作者 Zhihao Zhu

作者 Enhong Chen

作者 Senchao Yuan

日期 2021

短标题 Fight Fire with Fire

馆藏目录 Google Scholar

页码 1074–1083

会议论文集标题 Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval

添加日期 2022/11/16 下午7:54:32

修改日期 2022/11/16 下午7:54:32

附件

- Full Text
- Snapshot

Fusing hypergraph spectral features for shilling attack detection

类型 期刊文章

作者 Hao Li

作者 Min Gao

作者 Fengtao Zhou

作者 Yueyang Wang

作者 Qilin Fan

作者 Linda Yang

日期 2021

馆藏目录 Google Scholar

其它 Publisher: Elsevier

卷次 63

页码 103051

期刊 Journal of Information Security and Applications

添加日期 2022/11/16 下午6:48:35

修改日期 2022/11/17 下午6:31:25

附件

- Li 等 - 2021 - Fusing hypergraph spectral features for shilling a.pdf
- Li 等 - 2021 - Fusing hypergraph spectral features for shilling a.pdf
- Snapshot

Gray-Box Shilling Attack: An Adversarial Learning Approach

类型 期刊文章

作者 Zongwei Wang

作者 Min Gao

作者 Jundong Li

作者 Junwei Zhang

作者 Jiang Zhong

日期 2022

短标题 Gray-Box Shilling Attack

馆藏目录 Google Scholar

其它 Publisher: ACM New York, NY

期刊 ACM Transactions on Intelligent Systems and Technology (TIST)

添加日期 2022/11/16 下午7:34:14

修改日期 2022/11/16 下午7:34:14

附件

◦ Full Text

◦ Snapshot

Hybrid convolutional neural network (CNN) and long-short term memory (LSTM) based deep learning model for detecting shilling attack in the social-aware network

类型 期刊文章

作者 Kalimuthu Vivekanandan

作者 Narayanan Praveena

日期 2021

馆藏目录 Google Scholar

其它 Publisher: Springer

卷次 12

页码 1197–1210

期刊 Journal of Ambient Intelligence and Humanized Computing

期号 1

添加日期 2022/11/16 下午8:11:12

修改日期 2022/11/16 下午8:11:12

附件

◦ Snapshot


IEEE13-AdvAttack A Novel Dataset for Benchmarking the Power of Adversarial Attacks against Fault Prediction Systems in Smart Electrical Grid

类型 会议论文

作者 Carmelo Ardito

作者 Yashar Deldjoo

作者 Tommaso Di Noia
作者 Eugenio Di Sciascio
作者 Fatemeh Nazary
日期 2022

短标题  IEEE13-AdvAttack小说为基准数据集的力量对抗攻击智能电网故障预测系统

馆藏目录 Google Scholar

页码 3817–3821

会议论文集标题 Proceedings of the 31st ACM International Conference on Information & Knowledge Management

添加日期 2022/11/16 下午6:49:21

修改日期 2022/11/17 上午12:23:36

附件

- Full Text
- Snapshot

IMPACT ANALYSIS OF PROFILE INJECTION ATTACKS IN RECOMMENDER SYSTEM

类型 期刊文章

作者 M. Ashish Kumar

作者 Yudhvir Singh

作者 Vikas Siwach

作者 Harkesh Sehrawat

日期 2021

馆藏目录 Google Scholar

卷次 9

页码 472–478

期刊 INFORMATION TECHNOLOGY IN INDUSTRY

期号 1

添加日期 2022/11/16 下午7:58:52

修改日期 2022/11/16 下午7:58:52

附件

- Full Text
 - Snapshot
-

Improving Deep Learning-Based Recommendation Attack Detection Using Harris Hawks Optimization


类型 期刊文章

作者 Quanqiang Zhou

作者 Cheng Huang

作者 Liangliang Duan

日期 2022

短标题  改善深上优于推荐攻击检测使用哈里斯鹰优化

馆藏目录 Google Scholar

其它 Publisher: MDPI

卷次 12

页码 10135

期刊 Applied Sciences

期号 19

添加日期 2022/11/16 下午8:11:12

修改日期 2022/11/17 下午6:31:27

附件

- Snapshot
- Zhou 等 - 2022 - Improving Deep Learning-Based Recommendation Attac.pdf

Influence Function based Data Poisoning Attacks to Top-N Recommender Systems

类型 会议论文

作者 Minghong Fang

作者 Neil Zhenqiang Gong

作者 Jia Liu

摘要 Recommender system is an essential component of web services to engage users. Popular recommender systems model user preferences and item properties using a large amount of crowdsourced user-item interaction data, e.g., rating scores; then top-N items that match the best with a user's preference are recommended to the user. In this work, we show that an attacker can launch a data poisoning attack to a recommender system to make recommendations as the attacker desires via injecting fake users with carefully crafted user-item interaction data. Specifically, an attacker can trick a recommender system to recommend a target item to as many normal users as possible. We focus on matrix factorization based recommender systems because they have been widely deployed in industry. Given the number of fake users the attacker can inject, we formulate the crafting of rating scores for the fake users as an optimization problem. However, this optimization problem is challenging to solve as it is a non-convex integer programming problem. To address the challenge, we develop several techniques to approximately solve the

optimization problem. For instance, we leverage influence function to select a subset of normal users who are influential to the recommendations and solve our formulated optimization problem based on these influential users. Our results show that our attacks are effective and outperform existing methods.

日期 四月 20, 2020

馆藏目录 ACM Digital Library

URL <https://doi.org/10.1145/3366423.3380072>

访问时间 2022/11/16 上午8:00:00

地点 New York, NY, USA

出版社 Association for Computing Machinery

ISBN 978-1-4503-7023-3

页码 3019–3025

系列 WWW '20

会议论文集标题 Proceedings of The Web Conference 2020

DOI 10.1145/3366423.3380072

添加日期 2022/11/17 上午12:34:56

修改日期 2022/11/17 上午12:35:00

标签:

adversarial machine learning., Adversarial recommender systems, data poisoning attacks

附件

- 全文

Injection Shilling Attack Tool for Recommender Systems

类型 会议论文

作者 Fatemeh Rezaimehr

作者 Chitra Dadkhah

日期 2021

馆藏目录 Google Scholar

出版社 IEEE

页码 1–4

会议论文集标题 2021 26th International Computer Conference, Computer Society of Iran (CSICC)

添加日期 2022/11/16 下午7:34:14

修改日期 2022/11/16 下午7:34:14

Item-triggered recommendation for identifying potential customers of cold sellers in supermarkets

类型 期刊文章

作者 San Diego

语言 en

馆藏目录 Zotero

页码 105

添加日期 2022/11/17 上午12:27:59

修改日期 2022/11/17 上午12:28:34

附件

- Diego - Beyond Personalization 2005.pdf

Link farm, an effective attack to Page Rank in algorithm in graph-based recommender systems

类型 期刊文章

作者 Sima Iranmanesh

作者 Mohammad-Reza Pajoohan

日期 2021

馆藏目录 Google Scholar

其它 Publisher: Babol Noshirvani University of Technology

卷次 10

页码 53–67

期刊 Journal of Soft Computing and Information Technology

期号 2

添加日期 2022/11/16 下午7:34:14

修改日期 2022/11/16 下午7:34:14

附件

- Snapshot

LOKI: A Practical Data Poisoning Attack Framework against Next Item Recommendations

类型 期刊文章

作者 Hengtong Zhang

作者 Yaliang Li

作者 Bolin Ding

作者 Jing Gao

摘要 Due to the openness of the online platform, recommendation systems are vulnerable to data poisoning attacks, where malicious samples are injected into the training set of the recommendation system to manipulate its recommendation results. Existing attack approaches are either based on heuristic rules or designed against specific recommendation approaches. The former suffers unsatisfactory performance, while the latter requires strong knowledge of the target system. In this paper, we propose a practical poisoning attack approach named LOKI against blackbox recommendation systems. The proposed LOKI utilizes the reinforcement learning algorithm to train the attack agent, which can be used to generate user behavior samples for data poisoning. In real-world recommendation systems, the cost of retraining recommendation models is high, and the interaction frequency between users and a recommendation system is restricted. Thus, we propose to let the agent interact with a recommender simulator instead of the target recommendation system and leverage the transferability of the generated adversarial samples to poison the target system. We also use the influence function to efficiently estimate the influence of injected samples on recommendation results, without re-training the models. Extensive experiments on multiple datasets against four representative recommendation models show that the proposed LOKI outperforms existing method. We also discuss the characteristics of vulnerable users/items, and evaluate whether anomaly detection methods can be used to mitigate the impact of data poisoning attacks.

日期 2022

短标题 LOKI

馆藏目录 IEEE Xplore

其它 Conference Name: IEEE Transactions on Knowledge and Data Engineering

页码 1-13

期刊 IEEE Transactions on Knowledge and Data Engineering

DOI 10.1109/TKDE.2022.3181270

ISSN 1558-2191

添加日期 2022/11/16 下午8:05:50

修改日期 2022/11/16 下午8:33:14

标签:

Adversarial Learning, Behavioral sciences, Collaborative filtering, Data Poisoning, Detectors, Recommendation System, Recurrent neural networks, Reinforcement learning, Task analysis, Training

附件

- IEEE Xplore Abstract Record
- IEEE Xplore Full Text PDF

LOKI: A Practical Data Poisoning Attack Framework against Next Item Recommendations | IEEE Journals & Magazine | IEEE Xplore

类型 网页

URL <https://ieeexplore.ieee.org/abstract/document/9806383>

访问时间 2022/11/16 下午8:06:46

添加日期 2022/11/16 下午8:06:46

修改日期 2022/11/16 下午8:06:46

附件

- LOKI: A Practical Data Poisoning Attack Framework against Next Item Recommendations | IEEE Journals & Magazine | IEEE Xplore

Msap: Multi-step adversarial perturbations on recommender systems embeddings

类型 会议论文

作者 Vito Walter Anelli

作者 Alejandro Bellogin

作者 Yashar Deldjoo

作者 Tommaso Di Noia

作者 Felice Antonio Merra

日期 2021

短标题 Msap

馆藏目录 Google Scholar

页码 1–6

会议论文集标题 The 34th International FLAIRS Conference. The Florida AI Research Society (FLAIRS), AAAI Press

添加日期 2022/11/16 下午7:56:31

修改日期 2022/11/16 下午7:56:31

MSPLD: Shilling Attack Detection Model Based on Meta Self-Paced Learning

类型 会议论文

作者 Yanjing Yang

作者 Min Gao

作者 Yuerang Li

作者 Fan Wu

作者 Jia Wang

作者 Quanwu Zhao

日期 2021

短标题 MSPLD

馆藏目录 Google Scholar

出版社 IEEE

页码 1–8

会议论文集标题 2021 International Joint Conference on Neural Networks (IJCNN)

添加日期 2022/11/16 下午8:19:46

修改日期 2022/11/17 下午6:31:29

标签:

Feature extraction, Training, Adaptation models, Analytical models, meta self-paced learning, Metadata, Neural networks, Reliability theory, shilling attack detection, visual theoretical reliability analysis

附件

- IEEE Xplore Abstract Record
- IEEE Xplore Full Text PDF
- Snapshot

On the feasibility of crawling-based attacks against recommender systems 1

类型 期刊文章

作者 Fabio Aiolli

作者 Mauro Conti

作者 Stjepan Picek

作者 Mirko Polato

日期 2022

馆藏目录 Google Scholar

其它 Publisher: IOS Press

卷次 30

页码 599–621

期刊 Journal of Computer Security

期号 4

添加日期 2022/11/16 下午8:09:45

修改日期 2022/11/16 下午8:09:45

附件

- Snapshot

PipAttack: Poisoning Federated Recommender Systems for Manipulating Item Promotion

类型 期刊文章

作者 Shijie Zhang

作者 Hongzhi Yin
作者 Tong Chen
作者 Zi Huang
作者 Quoc Viet Hung Nguyen
作者 Lizhen Cui
日期 2021

短标题 PipAttack

馆藏目录 Google Scholar

期刊 arXiv preprint arXiv:2110.10926

添加日期 2022/11/16 下午7:34:14

修改日期 2022/11/16 下午7:34:14

附件

- Full Text
- Snapshot

Poison Attacks against Graph Representation Learning in Recommender Systems

类型 期刊文章

作者 Thanh Toan Nguyen

作者 Khang Nguyen Duc Quach

作者 Thanh Tam Nguyen

作者 Thanh Trung Huynh

作者 Viet Hung Vu

作者 Phi Le Nguyen

作者 Jun Jo

作者 Quoc Viet Hung Nguyen

日期 2022

馆藏目录 Google Scholar

添加日期 2022/11/16 下午7:56:31

修改日期 2022/11/16 下午7:56:31

附件

- Full Text

Poisoning Attacks to Graph-Based Recommender Systems

类型 会议论文

作者 Minghong Fang

作者 Guolei Yang

作者 Neil Zhenqiang Gong

作者 Jia Liu

摘要 Recommender system is an important component of many web services to help users locate items that match their interests. Several studies showed that recommender systems are vulnerable to poisoning attacks, in which an attacker injects fake data to a recommender system such that the system makes recommendations as the attacker desires. However, these poisoning attacks are either agnostic to recommendation algorithms or optimized to recommender systems (e.g., association-rule-based or matrix-factorization-based recommender systems) that are not graph-based. Like association-rule-based and matrix-factorization-based recommender systems, graph-based recommender system is also deployed in practice, e.g., eBay, Huawei App Store (a big app store in China). However, how to design optimized poisoning attacks for graph-based recommender systems is still an open problem. In this work, we perform a systematic study on poisoning attacks to graph-based recommender systems. We consider an attacker's goal is to promote a target item to be recommended to as many users as possible. To achieve this goal, our attacks inject fake users with carefully crafted rating scores to the recommender system. Due to limited resources and to avoid detection, we assume the number of fake users that can be injected into the system is bounded. The key challenge is how to assign rating scores to the fake users such that the target item is recommended to as many normal users as possible. To address the challenge, we formulate the poisoning attacks as an optimization problem, solving which determines the rating scores for the fake users. We also propose techniques to solve the optimization problem. We evaluate our attacks and compare them with existing attacks under white-box (recommendation algorithm and its parameters are known), gray-box (recommendation algorithm is known but its parameters are unknown), and blackbox (recommendation algorithm is unknown) settings using two real-world datasets. Our results show that our attack is effective and outperforms existing attacks for graph-based recommender systems. For instance, when 1% of users are injected fake users, our attack can make a target item recommended to 580 times more normal users in certain scenarios.

日期 十二月 3, 2018

馆藏目录 ACM Digital Library

URL <https://doi.org/10.1145/3274694.3274706>

访问时间 2022/11/16 上午8:00:00

地点 New York, NY, USA

出版社 Association for Computing Machinery

ISBN 978-1-4503-6569-7

页码 381–392

系列 ACSAC '18

会议论文集标题 Proceedings of the 34th Annual Computer Security Applications Conference

DOI 10.1145/3274694.3274706

添加日期 2022/11/17 上午12:34:19

修改日期 2022/11/17 上午12:34:19

标签:

Adversarial recommender systems, adversarial machine learning, poisoning attacks

附件

- 已提交版本

Poisoning Deep Learning based Recommender Model in Federated Learning Scenarios

类型 期刊文章

作者 Dazhong Rong

作者 Qinming He

作者 Jianhai Chen

日期 2022

馆藏目录 Google Scholar

期刊 arXiv preprint arXiv:2204.13594

添加日期 2022/11/16 下午7:54:32

修改日期 2022/11/16 下午7:54:32

附件

- Full Text
- Snapshot

Poisoning GNN-based Recommender Systems with Generative Surrogate-based Attacks

类型 期刊文章

作者 Thanh Toan Nguyen

作者 Khang Nguyen Duc Quach

作者 Thanh Tam Nguyen

作者 Thanh Trung Huynh

作者 Viet Hung Vu

作者 Phi Le Nguyen

作者 Jun Jo

作者 Quoc Viet Hung Nguyen

日期 2022

馆藏目录 Google Scholar

其它 Publisher: ACM New York, NY

期刊 ACM Transactions on Information Systems

添加日期 2022/11/16 下午7:54:32

修改日期 2022/11/16 下午7:54:32

附件

○ Full Text

○ Snapshot

Practical Data Poisoning Attack against Next-Item Recommendation

类型 会议论文

作者 Hengtong Zhang

作者 Yaliang Li

作者 Bolin Ding

作者 Jing Gao

摘要 Online recommendation systems make use of a variety of information sources to provide users the items that users are potentially interested in. However, due to the openness of the online platform, recommendation systems are vulnerable to data poisoning attacks. Existing attack approaches are either based on simple heuristic rules or designed against specific recommendations approaches. The former often suffers unsatisfactory performance, while the latter requires strong knowledge of the target system. In this paper, we focus on a general next-item recommendation setting and propose a practical poisoning attack approach named LOKI against blackbox recommendation systems. The proposed LOKI utilizes the reinforcement learning algorithm to train the attack agent, which can be used to generate user behavior samples for data poisoning. In real-world recommendation systems, the cost of retraining recommendation models is high, and the interaction frequency between users and a recommendation system is restricted. Given these real-world restrictions, we propose to let the agent interact with a recommender simulator instead of the target recommendation system and leverage the transferability of the generated adversarial samples to poison the target system. We also propose to use the influence function to efficiently estimate the influence of injected samples on the recommendation results, without re-training the models within the simulator. Extensive experiments on two datasets against four representative recommendation models show that the proposed LOKI achieves better attacking performance than existing methods.

日期 2020-04-20

语言 en

馆藏目录 DOI.org (Crossref)

URL <https://dl.acm.org/doi/10.1145/3366423.3379992>

访问时间 2022/11/16 下午6:39:51

地点 Taipei Taiwan

出版社 ACM

ISBN 978-1-4503-7023-3

页码 2458-2464

会议论文集标题 Proceedings of The Web Conference 2020
会议名称 WWW '20: The Web Conference 2020
DOI 10.1145/3366423.3379992
添加日期 2022/11/16 下午6:39:51
修改日期 2022/11/16 下午8:33:08

附件

- Zhang 等 - 2020 - Practical Data Poisoning Attack against Next-Item .pdf

Rank List Sensitivity of Recommender Systems to Interaction Perturbations

类型 会议论文
作者 Sejoon Oh
作者 Berk Ustun
作者 Julian McAuley
作者 Srijan Kumar
摘要 Prediction models can exhibit sensitivity with respect to training data: small changes in the training data can produce models that assign conflicting predictions to individual data points during test time. In this work, we study this sensitivity in recommender systems, where users' recommendations are drastically altered by minor perturbations in other unrelated users' interactions. We introduce a measure of stability for recommender systems, called Rank List Sensitivity (RLS), which measures how rank lists generated by a given recommender system at test time change as a result of a perturbation in the training data. We develop a method, CASPER, which uses cascading effect to identify the minimal and systematical perturbation to induce higher instability in a recommender system. Experiments on four datasets show that recommender models are overly sensitive to minor perturbations introduced randomly or via CASPER - even perturbing one random interaction of one user drastically changes the recommendation lists of all users. Importantly, with CASPER perturbation, the models generate more unstable recommendations for low-accuracy users (i.e., those who receive low-quality recommendations) than high-accuracy ones.

日期 十月 17, 2022

馆藏目录 ACM Digital Library

URL <https://doi.org/10.1145/3511808.3557425>

访问时间 2022/11/16 上午8:00:00

地点 New York, NY, USA

出版社 Association for Computing Machinery

ISBN 978-1-4503-9236-5

页码 1584–1594

系列 CIKM '22

会议论文集标题 Proceedings of the 31st ACM International Conference on Information & Knowledge Management

DOI 10.1145/3511808.3557425

添加日期 2022/11/17 上午12:31:52

修改日期 2022/11/17 上午12:31:52

标签:

recommender systems, input data perturbation, model stability

附件

- Full Text PDF

Rating behavior evaluation and abnormality forensics analysis for injection attack detection

类型 期刊文章

作者 Zhihai Yang

作者 Qindong Sun

作者 Zhaoli Liu

作者 Jinpei Yan

作者 Yaling Zhang

日期 2022

馆藏目录 Google Scholar

其它 Publisher: Springer

卷次 59

页码 93–119

期刊 Journal of Intelligent Information Systems

期号 1

添加日期 2022/11/16 下午7:58:52

修改日期 2022/11/16 下午7:58:52

附件

- Snapshot

Ready for emerging threats to recommender systems? A graph convolution-based generative shilling attack

类型 期刊文章

作者 Fan Wu

作者 Min Gao

作者 Junliang Yu

作者 Zongwei Wang

作者 Kecheng Liu

作者 Xu Wang
日期 2021
短标题 Ready for emerging threats to recommender systems?
馆藏目录 Google Scholar
其它 Publisher: Elsevier
卷次 578
页码 683–701
期刊 Information Sciences
添加日期 2022/11/16 下午7:28:34
修改日期 2022/11/16 下午7:28:34

附件

- Full Text
- Snapshot

Reverse Attack: Black-box Attacks on Collaborative Recommendation

类型 会议论文

作者 Yihe Zhang

作者 Xu Yuan

作者 Jin Li

作者 Jiadong Lou

作者 Li Chen

作者 Nian-Feng Tzeng

摘要 Collaborative filtering (CF) recommender systems have been extensively developed and widely deployed in various social websites, promoting products or services to the users of interest. Meanwhile, work has been attempted at poisoning attacks to CF recommender systems for distorting the recommend results to reap commercial or personal gains stealthily. While existing poisoning attacks have demonstrated their effectiveness with the offline social datasets, they are impractical when applied to the real setting on online social websites. This paper develops a novel and practical poisoning attack solution toward the CF recommender systems without knowing involved specific algorithms nor historical social data information a priori. Instead of directly attacking the unknown recommender systems, our solution performs certain operations on the social websites to collect a set of sampling data for use in constructing a surrogate model for deeply learning the inherent recommendation patterns. This surrogate model can estimate the item proximities, learned by the recommender systems. By attacking the surrogate model, the corresponding solutions (for availability and target attacks) can be directly migrated to attack the original recommender systems. Extensive experiments validate the generated surrogate model's reproductive capability and demonstrate the effectiveness of our attack upon various CF recommender algorithms.

日期 十一月 13, 2021

短标题 Reverse Attack

馆藏目录 ACM Digital Library

URL <https://doi.org/10.1145/3460120.3484805>

访问时间 2022/11/16 上午8:00:00

地点 New York, NY, USA

出版社 Association for Computing Machinery

ISBN 978-1-4503-8454-4

页码 51–68

系列 CCS '21

会议论文集标题 Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security

DOI 10.1145/3460120.3484805

添加日期 2022/11/17 上午12:29:50

修改日期 2022/11/17 上午12:29:50

标签:

poisoning attack, recommender system

附件

◦ Full Text PDF

Revisiting Injective Attacks on Recommender Systems

类型 期刊文章

作者 Haoyang Li

作者 Shimin Di

作者 Lei Chen

摘要 Recent studies have demonstrated that recommender systems (RecSys) are vulnerable to injective attacks. Given a limited fake user budget, attackers can inject fake users with carefully designed behaviors into the open platforms, making RecSys recommend a target item to more real users for profits. In this paper, we first revisit existing attackers and reveal that they suffer from the difficulty-agnostic and diversity-deficit issues. Existing attackers concentrate their efforts on difficult users who have low tendencies toward the target item, thus reducing their effectiveness. Moreover, they are incapable of affecting the target RecSys to recommend the target item to real users in a diverse manner, because their generated fake user behaviors are dominated by large communities. To alleviate these two issues, we propose a difficulty and diversity aware attacker, namely DADA. We design the difficulty-aware and diversity-aware objectives to enable easy users from various communities to contribute more weights when optimizing attackers. By incorporating these two objectives, the proposed attacker DADA can concentrate on easy users while also affecting a broader range of real users simultaneously, thereby boosting the effectiveness. Extensive experiments on three real-world datasets demonstrate the effectiveness of our proposed attacker.

语言 en
馆藏目录 Zotero
页码 14
添加日期 2022/11/16 下午8:05:32
修改日期 2022/11/16 下午8:05:33

附件

- Li 等 - Revisiting Injective Attacks on Recommender System.pdf

Revisiting Item Promotion in GNN-based Collaborative Filtering: A Masked Targeted Topological Attack Perspective

类型 期刊文章
作者 Yongwei Wang
作者 Yong Liu
作者 Zhiqi Shen
日期 2022
短标题 Revisiting Item Promotion in GNN-based Collaborative Filtering
馆藏目录 Google Scholar
期刊 arXiv preprint arXiv:2208.09979
添加日期 2022/11/16 下午8:19:46
修改日期 2022/11/16 下午8:19:46

附件

- Full Text
- Snapshot

RGRecSys: A Toolkit for Robustness Evaluation of Recommender Systems

类型 会议论文
作者 Zohreh Ovaisi
作者 Shelby Heinecke
作者 Jia Li
作者 Yongfeng Zhang
作者 Elena Zheleva
作者 Caiming Xiong
日期 2022
短标题 RGRecSys
馆藏目录 Google Scholar
页码 1597–1600

会议论文集标题 Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining

添加日期 2022/11/16 下午7:54:32

修改日期 2022/11/16 下午7:54:32

附件

- Full Text
- Snapshot

RMPD: Method for enhancing the robustness of recommendations with attack environments

类型 期刊文章

作者 Qi Ding

作者 Peiyu Liu

作者 Zhenfang Zhu

作者 Huajuan Duan

作者 Fuyong Xu

日期 2021

短标题 RMPD

馆藏目录 Google Scholar

其它 Publisher: IEEE

卷次 9

页码 17843–17853

期刊 IEEE Access

添加日期 2022/11/16 下午7:34:14

修改日期 2022/11/16 下午7:34:14

SARCP: Exploiting Cyber-Attack Prediction Through Socially-Aware Recommendation

类型 期刊文章

作者 Nana Yaw Asabere

作者 Elikem Fiamavle

作者 Joseph Agyiri

作者 Wisdom Kwawu Torgby

作者 Joseph Eyram Dzata

作者 Nina Pearl Doe

日期 2022

短标题 SARCP

馆藏目录 Google Scholar
其它 Publisher: IGI Global
卷次 14
页码 1–21
期刊 International Journal of Decision Support System Technology (IJDSSST)
期号 1
添加日期 2022/11/16 下午7:58:52
修改日期 2022/11/16 下午7:58:52

附件

- Snapshot

Semi-supervised recommendation attack detection based on Co-Forest

类型 期刊文章
作者 Quanqiang Zhou
作者 Liangliang Duan
摘要 In recommendation attack, malicious users attempt to bias the recommendation results by injecting fake profiles into the rating database. To detect such attack, three types of methods, i.e., unsupervised, supervised and semi-supervised, have been proposed. Among these works, the advantage of semi-supervised methods is that they can use the unlabeled user profiles to improve the detection performance. However, the existing semi-supervised methods suffer from low precision. Aiming at this problem, in this paper, we propose a semi-supervised detection approach named SSADR-CoF based on the Co-Forest algorithm. Being different from the existing semi-supervised methods which only use a few of features to train a single classifier for the detection, the proposed approach uses a series of features to train an ensemble of classifiers to detect the recommendation attack. We first use the window dividing and rating behavior statistical methods to extract a series of user rating behavior mode features for training the detection model. Then, we use a small number of labeled user profiles to initialize an ensemble of classifiers, and use the ensemble of classifiers to assign labels to the unlabeled user profiles. Finally, we use the labeled and the newly labeled user profiles to iteratively update the classifiers for the detection. Experiments conducted on three benchmark datasets MovieLens 10M, MovieLens 25M, and Amazon show that the proposed approach can effectively improve the precision of the semi-supervised methods under the condition of maintaining high recall and AUC.

日期 2021-10-01

语言 en

馆藏目录 ScienceDirect

URL <https://www.sciencedirect.com/science/article/pii/S0167404821002145>

访问时间 2022/11/16 下午8:13:01

卷次 109

页码 102390

期刊 Computers & Security

DOI 10.1016/j.cose.2021.102390

刊名缩写 Computers & Security

ISSN 0167-4048

添加日期 2022/11/16 下午8:13:01

修改日期 2022/11/16 下午8:13:01

标签:

Attack detection, Co-Forest algorithm, Collaborative recommender system, Recommendation attack, Semi-supervised learning

附件

- ScienceDirect Full Text PDF

Sequential Attack Detection in Recommender Systems

类型 期刊文章

作者 Mehmet Aktukmak

作者 Yasin Yilmaz

作者 Ismail Uysal

日期 2021

馆藏目录 Google Scholar

其它 Publisher: IEEE

卷次 16

页码 3285–3298

期刊 IEEE Transactions on Information Forensics and Security

添加日期 2022/11/16 下午7:28:34

修改日期 2022/11/17 下午6:31:31

标签:

Computational modeling, Recommender systems, Data models, cyber-attack detection, Data integration, Detection algorithms, Hidden Markov models, History, latent variable model, quickest detection, variational inference

附件

- IEEE Xplore Full Text PDF
- Snapshot

Shilling Attack Detection System for Online Recommenders

类型 会议论文

作者 J. R. V. Jeny

作者 R. Sowmya

作者 G. Sai Kiran

作者 M. Kiran Babu

作者 Ch Arjun

日期 2022

馆藏目录 Google Scholar

出版社 IEEE

页码 988–992

会议论文集标题 2022 International Conference on Inventive Computation Technologies (ICICT)

添加日期 2022/11/16 下午8:19:46

修改日期 2022/11/17 下午6:31:33

标签:

Clustering algorithms, Recommender systems, Computational efficiency, DBSCAN clustering algorithm, Detecting Group Shilling attacks, Group Shilling attacks, Online recommendation system, Recommender System

附件

- IEEE Xplore Full Text PDF
- Snapshot

Targeted Data Poisoning Attack on News Recommendation System by Content Perturbation

类型 预印本

作者 Xudong Zhang

作者 Zan Wang

作者 Jingke Zhao

作者 Lanjun Wang

摘要 News Recommendation System(NRS) has become a fundamental technology to many online news services. Meanwhile, several studies show that recommendation systems(RS) are vulnerable to data poisoning attacks, and the attackers have the ability to mislead the system to perform as their desires. A widely studied attack approach, injecting fake users, can be applied on the NRS when the NRS is treated the same as the other systems whose items are fixed. However, in the NRS, as each item (i.e. news) is more informative, we propose a novel approach to poison the NRS, which is to perturb contents of some browsed news that results in the manipulation of

the rank of the target news. Intuitively, an attack is useless if it is highly likely to be caught, i.e., exposed. To address this, we introduce a notion of the exposure risk and propose a novel problem of attacking a history news dataset by means of perturbations where the goal is to maximize the manipulation of the target news rank while keeping the risk of exposure under a given budget. We design a reinforcement learning framework, called TDP-CP, which contains a two-stage hierarchical model to reduce the searching space. Meanwhile, influence estimation is also applied to save the time on retraining the NRS for rewards. We test the performance of TDP-CP under three NRSs and on different target news. Our experiments show that TDP-CP can increase the rank of the target news successfully with a limited exposure budget.

日期 2022-03-09
馆藏目录 arXiv.org
URL <http://arxiv.org/abs/2203.03560>
访问时间 2022/11/17 上午12:25:55
其它 arXiv:2203.03560 [cs]
DOI 10.48550/arXiv.2203.03560
仓库 arXiv
存档ID arXiv:2203.03560
添加日期 2022/11/17 上午12:25:55
修改日期 2022/11/17 上午12:25:55

标签:

Computer Science - Cryptography and Security, Computer Science - Information Retrieval, Computer Science - Machine Learning, Computer Science - Artificial Intelligence

附件

- arXiv Fulltext PDF
- arXiv.org Snapshot

Three Birds With One Stone: User Intention Understanding and Influential Neighbor Disclosure for Injection Attack Detection

类型 期刊文章

作者 Zhihai Yang

作者 Qindong Sun

作者 Zhaoli Liu

摘要 Recommender system, as a data-driven way to help customers locate products that match their interests, is increasingly critical for providing competitive customer suggestions in many web services. However, recommender systems are highly vulnerable to malicious injection attacks due to their fundamental vulnerabilities and openness. With the endless emergence of new attacks, how to provide a feasible way for defending different malicious threats against online recommendations is still an under-explored issue. In this paper, we explore a new way to defend malicious

injection attacks through user intention understanding and influential neighbour disclosure. Specifically, we propose a detection approach, termed TBOS (Three Birds with One Stone), to deal with different malicious threats. In TBOS, we first develop the discrimination of attack target by combining global influence evaluation and risk attitude estimation of users. In order to make TBOS controllable, second, we propose to incorporate an optimal denoising mechanism to remove disturbed information before detection. To enhance the representativeness and predictability of detection model, finally, we propose to leverage a behavioral label propagation mechanism based on constructed label space for the determination of malicious injection behaviors. Extensive experiments on both synthetic and real data demonstrate that TBOS outperforms all baselines in different cases. Particularly, the detection performance of TBOS can achieve an improvement of 6.08% FAR (false alarm rate) for optimal-injection attacks, an improvement of 3.83% FAR in average for co-visitation injection attacks, as well as an improvement of 2.3% for profile injection attacks over benchmarks in terms of FAR while keeping the highest DR (detection rate). Additional experiments on real-world data show that TBOS brings an improvement with the advantage of 6.5% FAR in average compared with baselines.

日期 2022

短标题 Three Birds With One Stone

馆藏目录 IEEE Xplore

其它 Conference Name: IEEE Transactions on Information Forensics and Security

卷次 17

页码 531-546

期刊 IEEE Transactions on Information Forensics and Security

DOI 10.1109/TIFS.2022.3146769

ISSN 1556-6021

添加日期 2022/11/16 下午8:13:39

修改日期 2022/11/16 下午8:13:39

标签:

Recommender systems, Predictive models, attack detection, behavior representation, Birds, Estimation, Injection attack, performance analysis, Position measurement, Sun, Technological innovation

附件

- IEEE Xplore Abstract Record
- IEEE Xplore Full Text PDF

Towards Adversarially Superior Malware Detection Models: An Adversary Aware Proactive Approach using Adversarial Attacks and Defenses

类型 期刊文章

作者 Hemant Rathore

作者 Adithya Samavedhi

作者 Sanjay K. Sahay

作者 Mohit Sewak

日期 2022

短标题 Towards Adversarially Superior Malware Detection Models

馆藏目录 Google Scholar

其它 Publisher: Springer

页码 1–21

期刊 Information Systems Frontiers

添加日期 2022/11/16 下午6:49:21

修改日期 2022/11/16 下午6:49:21

附件

- Snapshot

Triple Adversarial Learning for Influence based Poisoning Attack in Recommender Systems

类型 会议论文

作者 Chenwang Wu

作者 Defu Lian

作者 Yong Ge

作者 Zhihao Zhu

作者 Enhong Chen

摘要 As an important means to solve information overload, recommender systems have been widely applied in many fields, such as e-commerce and advertising. However, recent studies have shown that recommender systems are vulnerable to poisoning attacks; that is, injecting a group of carefully designed user profiles into the recommender system can severely affect recommendation quality. Despite the development from shilling attacks to optimization-based attacks, the imperceptibility and harmfulness of the generated data in most attacks are arduous to balance. To this end, we propose a triple adversarial learning for influence based poisoning attack (TrialAttack), a flexible end-to-end poisoning framework to generate non-notable and harmful user profiles. Specifically, given the input noise, TrialAttack directly generates malicious users through triple adversarial learning of the generator, discriminator, and influence module. Besides, to provide reliable influence for TrialAttack training, we explore a new approximation approach for estimating each fake user's influence. Through theoretical analysis, we prove that the distribution characterized by TrialAttack approximates to the rating distribution of real users under the premise of performing an efficient attack. This property allows the injected users to attack in an unremarkable way. Experiments on three real-world datasets show that TrialAttack's attack performance outperforms state-of-the-art attacks, and the generated fake profiles are more difficult to detect compared to baselines.

日期 八月 14, 2021

馆藏目录 ACM Digital Library

URL <https://doi.org/10.1145/3447548.3467335>

访问时间 2022/11/16 上午8:00:00

地点 New York, NY, USA

出版社 Association for Computing Machinery

ISBN 978-1-4503-8332-5

页码 1830–1840

系列 KDD '21

会议论文集标题 Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining

DOI 10.1145/3447548.3467335

添加日期 2022/11/17 上午12:31:32

修改日期 2022/11/17 下午6:31:34

标签:

recommender systems, poisoning attacks, adversarial learning

附件

- Wu 等 - 2021 - Triple Adversarial Learning for Influence based Po.pdf

UA-FedRec: Untargeted Attack on Federated News Recommendation

类型 期刊文章

作者 Jingwei Yi

作者 Fangzhao Wu

作者 Bin Zhu

作者 Yang Yu

作者 Chao Zhang

作者 Guangzhong Sun

作者 Xing Xie

日期 2022

短标题 UA-FedRec

馆藏目录 Google Scholar

期刊 arXiv preprint arXiv:2202.06701

添加日期 2022/11/17 上午12:32:11

修改日期 2022/11/17 上午12:32:11

附件

- Full Text
- Snapshot