基于特征选择的推荐系统托攻击检测算法

伍之昂[1],庄 毅[2],王有权[3],曹 杰[1,3]
(1.南京财经大学江苏省电子商务重点实验室,江苏南京 210003;
2.浙江工商大学计算机与信息工程学院,浙江杭州 310018;3.南京理工大学计算机科学与技术学院,江苏南京 210094)

摘 要: 基于协同过滤的电子商务推荐系统极易受到托攻击,托攻击者注入伪造的用户模型增加或减少目标对象的推荐频率,如何检测托攻击是目前推荐系统领域的热点研究课题.分析五种类型托攻击对不同协同过滤算法产生的危害性,提出一种特征选择算法,为不同类型托攻击选取有效的检测指标.基于选择出的指标,提出两种基于监督学习的托攻击检测算法,第一种算法基于朴素贝叶斯分类;第二种算法基于 k 近邻分类.最后,通过实验验证了特征选择算法的有效性,及两种算法的灵敏性和特效性.

关键词: 推荐系统; 托攻击检测; 特征选择; 朴素贝叶斯分类; k 近邻分类
中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112 (2012)08-1687-07
电子学报 URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2012.08.031

Shilling Attack Detection Based on Feature Selection for Recommendation Systems

WU Zhi-ang[1], ZHUANG Yi[2], WANG You-quan[3], CAO Jie[1,3]
(1. *Jiangsu Provincial Key Laboratory of E-Business*, *Nanjing University of Finance and Economics*, *Nanjing*, *Jiangsu* 210003, *China*;
2. *College of Computer and Information Engineering*, *Zhejiang Gongshang University*, *Hangzhou*, *Zhejiang* 310018, *China*;
3. *College of Computer Science and Technology*, *Nanjing University of Science and Technology*, *Nanjing*, *Jiangsu* 210094, *China*)

Abstract: Most of the e-business recommender systems are based upon collaborative filtering (CF) algorithms. Since such systems have been shown to be vulnerable to shilling attacks in which malicious user profiles are inserted into the system in order to push or nuke the predictions of some targeted items, shilling attack detection has recently become a hot research topic in recommender systems. Firstly, the effectiveness of five types of attacks against different CF algorithms is analyzed. Secondly, a feature selection algorithm is presented. Two kinds of shilling attack detection algorithms based on supervised learning are then proposed: the first one is based on naïve Bayesian classifier, and the second one is based on k nearest neighbor (kNN) classifier. At last, experimental results show the effectiveness of the feature selection algorithm and the sensitivity and specificity of these two kinds of detection algorithms.

Key words: recommender system; shilling attack detection; feature selection; naïve Bayesian classifier; kNN classifier

1 引言

电子商务的迅速发展使得服务信息呈现"超载"趋势,用户面对海量的商品服务信息束手无策,难以顺利找到自己需要的商品服务信息.推荐系统(recommender system)是信息过滤的重要手段,是解决信息超载问题非常有潜力的方法.协同过滤(collaborative filtering, CF)是应用最广泛的推荐算法[1],很多著名的电子商务推荐系统都是基于协同过滤的,如亚马逊网络书店、GroupLens、TiVo、Netflix、YouTube 和 Facebook 等.然而,基于协同过滤的电子商务推荐系统极易受到托攻击(shilling attack),托攻击者通过伪造用户模型(user profile)干预系统的推荐结果,增加或减少目标对象的推荐频率.比如,某些恶意生产商或店主为了使自己的产品更加畅销,利用托攻击使得推荐系统频繁推荐自己的商品,而减少或不推荐竞争对手的商品.

如何防范和检测托攻击,保证电子商务推荐系统的安全性成为近年来推荐系统领域的一个研究热点.托攻

收稿日期:2011-09-25;修回日期:2012-02-16
基金项目:国家自然科学基金(No.61103229, No.71072172, No.61003074);浙江省自然科学基金(No.Z110822, No.Y1110644, No.Y1110969, No.Y1090165);江苏省科技支撑计划工业部分(No.BE2011198);江苏省高等学校优秀科技创新团队(No.2011013);东南大学江苏省网络与信息安全重点实验室开放课题(No.BM2003201);江苏省高校科研成果产业化推进项目(No.JHB2011-21)

击检测是二元分类问题,现有检测算法大多利用一系列指标来区分托攻击者和正常用户.事实上,一个指标不可能对所有类型的托攻击都有效,而仅仅对某些类型的托攻击有效,托攻击检测的效果很大程度上依赖于所选取的特征指标.并且对于不同类型的托攻击,往往需要选择不同的特征指标.本文提出基于特征选择的推荐系统托攻击检测算法,首先提出一种特征选择算法,基于训练集自动选择有价值的指标,然后分别利用朴素贝叶斯分类和 k 近邻两种分类技术进行托攻击检测.我们在 MovieLens 数据集上验证了特征选择对检测性能的促进作用,并验证了检测算法对付各种类型托攻击的有效性.

2 相关工作

Zhang 等人将推荐系统的托攻击分为五类[2],攻击者通过对目标项目(target item)和装填项目(filler item)的评分伪造用户模型,目标项目是攻击者试图提高或降低推荐频率的项目,装填项目是攻击者为了达到目的选择出一部分项目进行打分.本文以攻击者试图提高目标项目的推荐频率,即推攻击(push attack)为例,表 1列出了推攻击中五类托攻击名称及其对目标项目和装填项目的评分方法.五类托攻击都对目标项目评最高分,区别在于攻击者对装填项目的评分方法不同.随机攻击和平均攻击易于理解,分段攻击将与目标项目相似的项目评最高分;流行攻击的基本思想是齐普夫定律(Zip′s law),即少数项目可以吸引大多数人的注意,攻击者将目标项目和流行项目评最高分,以便跟大多数用户相似;抽样攻击也称为拷贝模型攻击.

表 1 推攻击中五类托攻击及其评分方法

攻击模型	评分方法
随机攻击	对目标项目评最高分,对装填项目随机评分
平均攻击	对目标项目评最高分,对装填项目取其平均分
分段攻击	对目标项目及其同类项目评最高分,对装填项目评最低分
流行攻击	对目标项目和流行项目最高评分,对装填项目随机评分
抽样攻击	拷贝已有的用户模型,对目标项目评最高分

Lam 等人首次分析了托攻击并说明随机攻击和平均攻击对不同协同过滤算法的影响[3].Chirita 等人提出了基于多个经验指标的随机攻击检测算法和分段攻击检测算法[4].尽管除随机攻击之外的四种攻击都需获取推荐系统的知识,但是,这些知识是不难获取的,比如,很多推荐系统(如 MovieLens)公开了项目的均分,这为攻击者实施各类更复杂的托攻击提供了条件,因此,需全面研究各种类型托攻击的检测方法.

目前,托攻击检测算法的设计有两种思路:基于监督学习(supervised learning)和基于无监督学习(unsupervised learning).Chirita[4]和 Burke[5]提出的算法是基于决策树的,属于监督学习;Mehta 提出了基于无监督学习的检测算法[6].Hurley 等人利用奈曼-皮尔逊准则来检测托攻击,并分别提出监督学习算法和无监督学习算法[7].本文提出的检测算法是基于监督学习中的惰性学习法,且不像已有算法一样采用预定义的特征指标,而是动态根据训练集选择特征指标,试图使检测算法能对付上述五类托攻击,从而提高算法的普适性.特征选择是在指标集合中选择子集,使得目标函数最优,目标函数由用户根据具体问题定义,特征选择是 NP-hard 问题[8,9].本文提出一种特征选择启发式算法,它能够在训练数据集上快速获得最有价值的几个检测指标,能大幅度提高检测算法的有效性.

3 托攻击对协同过滤算法的危害性分析

协同过滤算法可以分为两个阶段[10]:(1)计算相似度,利用 kNN 算法找出用户或项目的 k-近邻;(2)基于用户或项目的 k-近邻产生预测值.推荐系统的协同过滤算法可分为三类:基于用户的协同过滤(User-based CF,UCF)、基于项目的协同过滤(Item-based CF,ICF)、混合式协同过滤(Hybrid CF,HCF),三种协同过滤算法区别在于第二阶段.UCF 根据用户 u 的 k-近邻对项目 i 的评分进行预测,ICF 根据项目 i 的 k-近邻对用户 u 对其的评分进行预测,HCF 结合了上述两种协同过滤算法.

推荐系统面对托攻击的脆弱性因托攻击类型以及协同过滤算法而异,本节通过实验分析第 2 节所提出的五种托攻击对上述三种协同过滤算法造成的危害性.实验数据集采用 GroupLens 研究小组的 MovieLens*,它包含 943 个用户对 1682 部电影的 100000 条评分信息,分值处于[0,5]区间内,评分越高表示用户对该电影越满意.使用平均预测偏移来衡量推荐系统的脆弱性:

$$\bar{\Delta} = \sum_{(u,i)\in T} \frac{p'_{u,i} - p_{u,i}}{|T|} \tag{1}$$

其中,$|T|$ 为预测的记录数量,p' 是存在托攻击时的预测值,p 是正常情况下的预测值.平均预测偏移越大,说明推荐系统面对托攻击时越脆弱,即托攻击对该推荐系统越有效.本节的实验按照攻击比例(分别取 943 个用户的 10% 和 15%)向 MovieLens 数据集中分别注入五种攻击,随机选择平均分小于 3 分的 30 部电影作为目标项目集合.随机评分遵从均值 3.6、标准差 1.1 的高斯分布.

* http://movielens.umn.edu/

表 2 列出 UCF、ICF 和 HCF 三种算法对五类托攻击的平均预测偏移.除分段攻击之外,以 UCF 为算法的推荐系统受到的危害最大,原因在于,UCF 纯粹根据近邻来进行预测,托攻击者容易成为正常用户的近邻.同时,攻击比例越高,对推荐系统的危害也越大,当然,攻击者注入高比例用户模型的代价也越高.

表 2　三种协同过滤算法对五类托攻击的脆弱性

攻击类型	攻击比例(10%)			攻击比例(15%)		
推荐算法	UCF	ICF	HCF	UCF	ICF	HCF
随机攻击	**0.675**	0.483	0.54	0.723	0.733	**0.737**
平均攻击	**0.925**	0.565	0.716	**1.28**	0.762	0.962
分段攻击	0.772	0.779	**0.886**	1.124	**1.361**	1.261
流行攻击	**0.83**	0.485	0.584	**1.139**	0.789	0.862
抽样攻击	**0.65**	0.545	0.551	**0.815**	0.732	0.786

4　托攻击的检测指标选择

托攻击者一般利用正态分布 $N(\mu,\sigma^2)$生成随机评分数据[3,11],托攻击者的随机评分变化幅度较小,而正常用户会根据自己的兴趣偏好评分,评分变化幅度较大.为此,本文定义了熵来描述用户模型评分的变化程度,熵的定义如下:用户模型 Pu 可表示为统计集合 $Xu=\{n_i, i=1,2,\cdots,r_{\max}\}$,其中 i 是评分值,n_i 是评分值 i 在 Pu 中出现的次数.熵 $Entropy(Xu)$的计算公式如式(2)所示:

$$Entropy(X_u)=-\sum_{i=1}^{r_{\max}}\frac{n_i}{S}\log_2\frac{n_i}{S},\text{其中 } S=\sum_{i=1}^{r_{\max}}n_i \tag{2}$$

熵的范围是$[0,\log_2 r_{\max}]$,熵越小,表示评分值变化幅度越小,当所有评分值都相等时,熵为 0;当 n_i 相等时,熵取到最大值 $\log_2 r_{\max}$.

除此之外,本文引用 Williams 和 Burke 等人在文献[5,11]中定义的 9 个指标:RDMA(Rating Deviation from Mean Agreement)、WDA (Weighted Degree of Agreement)、WDMA (Weighted Deviation from Mean Agreement)、ADegSim (Average Degree of Similarity with Top Neighbors)、LengthVar (Length Variance)、FMTD (Filler Mean Target Di?erence)、FMV (Filler Mean Variance)、MeanVar (Mean Variance)和 TMF (Target Model Focus).因此,本文提出的托攻击检测算法一共考虑了 10 个指标.

设有 $n(n=10)$个检测指标,我们需要选择一个子集包含 m 个最有价值的指标($m<n$),算法 1 给出了特征选择算法的伪代码.第 1 行定义了每个检测指标的权重 $W[M_j]$ $(j=1,2,\cdots,n)$,初始值赋为 0.第 2 – 9 行从训练数据集中随机抽出 1 个用户,重复 S_n 次,对于每个被抽出的用户 u,根据 PCC 相似度找出与 u 属于同一类的最近邻 u_s,及与 u 属于同一类的最近邻 u_d,根据两个用户在指标 M_j 上的差值来调整 $W[M_j]$,$diff(M_j,u,v)$表示 u 用户 M_j 值减去 v 用户的 M_j 值.算法 1 所示特征选择算法的思想是某指标在两个不同类用户间的差值越大,其权重就越高(式(3)更新 $W[M_j]$时加上两个不同类用户间的差值 $diff(M_j,u,u_d)$);而某指标在两个同类用户间的差值越大,其权重就越低(式(3)更新 $W[M_j]$时减去两个同类用户间的差值 $diff(M_j,u,u_s)$).本文实验使用 PCC 来衡量用户之间的相似度.

抽样次数 S_n 越大,选择出的特征子集越可信,但是,耗费时间也越长.S_n 不超过训练数据集的总数,且应在两个类中抽样相同次数,避免造成权重更新的不平衡.m 值过小,如 $m=1$,无法有效区分托攻击者和正常用户;m 值过大,如 $m=n$,又包含了很多冗余指标,也将影响检测算法性能.本文通过实验给出了各类托攻击检测指标的排序,以及 m 取值对两种检测算法的影响.本文实验详细分析了特征选择算法对检测算法的影响.

算法 1　特征选择算法

输入:训练数据集,n 个检测指标,抽样次数 S_n

输出:m 个检测指标

1.将 n 个检测指标的初始权重 $W[M_j]$设为 0 $(j=1,2,\cdots,n)$

2.for $s\leftarrow 1$ to S_n　　/* 从训练数据集中抽样 S_n 次 */

3.　随机选择训练数据集中的 1 个用户,记为 u

4.　找出与用户 u 属于同一类的最近邻居,记为 u_s

5.　找出与用户 u 不属于同一类的最近邻居,记为 u_d

6.　for $j\leftarrow 1$ to n　　/* 更新 n 个指标的权重 */

7.　　$W[M_j]\leftarrow W[M_j]-diff(M_j,u,u_s)+diff(M_j,u,u_d)$　　(3)

8.　end for

9.end for

10.选择权重最大的 m 个检测指标作为输出结果

5　基于监督学习的托攻击检测算法

检测算法用于发现推荐系统中的异常用户模型,理想状况是托攻击者模型集合和异常用户模型集合等同,这样,若能检测所有的异常用户模型,就可检测出所有的托攻击.但是,托攻击者模型集合并不经常与异常用户模型集合等同,合法用户模型也可能属于异常用户模型集合.本文假设托攻击者模型是异常用户模型集合的子集,检测算法就是构造异常活动集合,并从中发现托攻击者模型子集.本节提出基于朴素贝叶斯分类和基于 k 近邻分类的两种托攻击检测算法.

5.1　基于朴素贝叶斯分类的托攻击检测算法

贝叶斯分类法来自统计法,可以预测特定用户模

型属于一个特定类的概率,由于本文第4节所选出的一系列指标间是相互独立的,因此,我们可以使用朴素贝叶斯分类(Naïve Bayesian Classifier)来检测托攻击,具体过程如算法2所示:

算法2 基于朴素贝叶斯分类的托攻击检测算法

输入:训练数据集,待检测的用户模型 u

输出:用户模型 u 的类别

1.对训练数据集预处理

2.计算 u 的 n 个指标值,记为 $\{x_1, x_2, \cdots, x_n\}$

3.利用算法1所示的特征选择算法选择 m 个指标

4.根据式(4)计算 $P(x_k \mid i)$,μ_{ik} 和 σ_{ik} 分别是 i 类训练数据第 k 个指标的均值和标准差

$$P(x_k \mid i) = g(x_k, \mu_{ik}, \sigma_{ik}),$$

$$\text{其中 } g(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4)$$

5.根据式(5)计算判断托攻击者的分值

$$\Omega = \ln \frac{P(S \mid u)}{P(N \mid u)} = \ln \frac{P(S)}{P(N)} + \sum_{k=1}^{m} \ln \frac{P(x_k \mid S)}{P(x_k \mid N)} \quad (5)$$

6.根据 Ω 的值判断 u 是否为托攻击者.

对训练数据集预处理(第1行)首先由用户模型以及式10个指标的值,得到图1所示格式的数据,然后调用算法1所示特征选择算法获得最有价值的 m 个指标.

UID/Profile	Class S/N	Metric 1 *H(X)*	Metric 2 *ADegSim*	.....	Metric *n* *TMF*

图1 预处理后的数据格式

由于定义的10个指标都是连续值,本文首先计算第 i 类训练数据上第 k 个指标的均值和标准差 μ_{ik} 和 σ_{ik},代入高斯分布 $g(x, \mu, \sigma)$ 函数从而得到 $P(x_k \mid i)$.

5.2 基于 k 近邻分类的托攻击检测算法

本节提出基于 k 近邻分类的托攻击检测算法,是一种基于实例的学习法(instance-based learner),该算法只是对训练数据集存储并稍加处理,当有了待分类用户模型时,才进行泛化,根据存储的训练数据集的相似性对该用户模型进行分类.

算法3 基于 k 近邻分类的托攻击检测算法

输入:训练数据集,待检测的用户模型 u

输出:用户模型 u 的类别

1.对训练数据集预处理

2.计算 u 的 n 个指标值,记为 $\{x_1, x_2, \cdots, x_n\}$

3.利用算法1所示的特征选择算法选择 m 个指标

4.计算用户模型 u 与训练数据集中用户模型的欧氏距离

5.找出用户模型 u 在训练数据集中最近的 k 个数据点

6.根据多数决定原则(majority rule)预测用户模型 u 的类别.

算法3对训练数据集预处理(第1行)首先仍然是得到图1所示格式的数据,并需要对 n 个指标值规范化,同样调用算法1所示特征选择算法获得最有价值的 m 个指标;然后使用交叉确认方法(cross-validation)确定 k 值,一般地,k 值随着训练数据集的增加而增大.本文简单选择欧式距离作为用户模型的距离.多数决定原则指待分类用户模型指派到它的 k 个最近邻中的多数类.

6 仿真实验

为了评估本文所提出算法的有效性,我们沿用评估分类器的两个常用度量:灵敏性(sensitivity)和特效性(specificity).对于托攻击检测问题,我们定义灵敏性和特效性两个评价指标:

$$sensitivity = \frac{\#\ true\ positives}{\#\ true\ positives + \#\ false\ negatives} \quad (6)$$

$$specificity = \frac{\#\ true\ negatives}{\#\ true\ negatives + \#\ false\ positives} \quad (7)$$

本文采用 Chirita 提出检测随机攻击和平均攻击的算法(称为 Chirita-SDA)[4]作为比较对象,本节通过四组实验说明特征选择算法对托攻击检测的有效性,以及本文所提出的两种算法检测各类托攻击的效果.

表3 各托攻击类型检测指标的 Top-5 排序结果

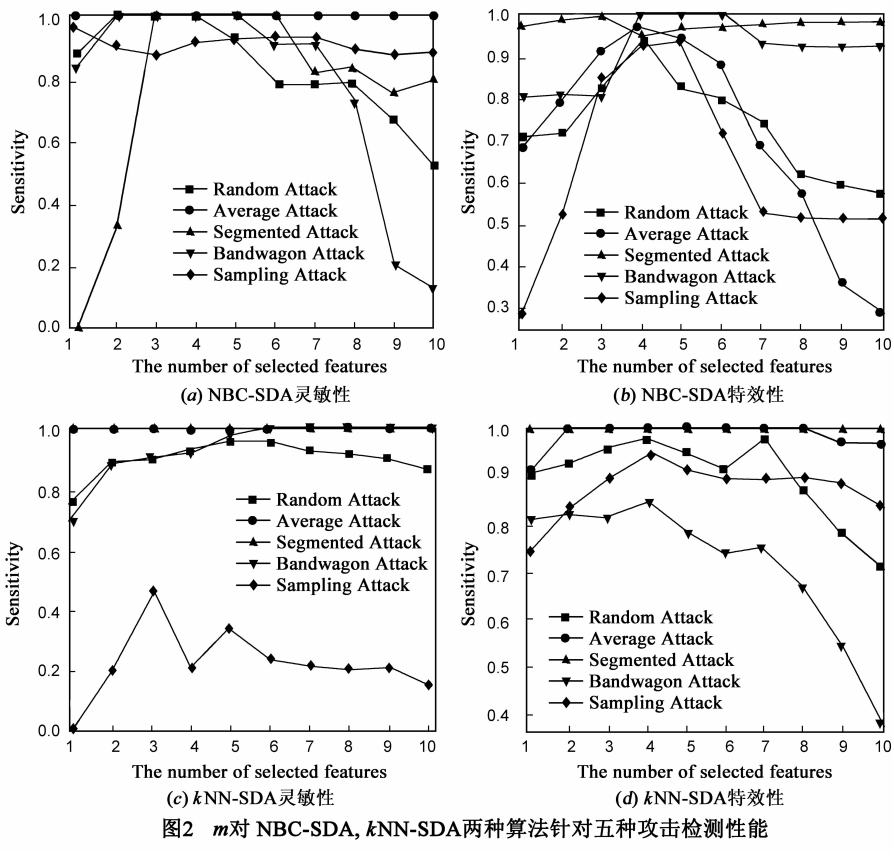
类型	检测指标排序				
	No.1	No.2	No.3	No.4	No.5
随机攻击	ADegsim	FMTD	MeanVar	LengthVar	Entropy
平均攻击	WDA	Entropy	ADegsim	MeanVar	FMV
分段攻击	FMTD	Entropy	FMV	MeanVar	RDMA
流行攻击	ADegsim	MeanVar	WDA	TMF	WDMA
抽样攻击	TMF	ADegsim	LengthVar	FMV	Entropy

第一组实验给出了检测每种攻击类型的各托攻击检测指标的排序结果.实验将 MovieLens 训练集中的用户标记为正常用户,根据表1定义生成五类攻击类型的用户模型,将攻击者用户模型注入 MovieLens 的训练集,这样就得到了实验的训练数据集,它包含100个正常用户和50个攻击者,装填比例为10%(即168部电影).表3是调用4.2节描述的特征选择算法得出的各攻击类型的检测指标 Top-5 排序结果,可以看出每类攻击的排序结果差异性很大,如 TMF 指标对随机攻击、平均攻击和分段攻击类型的检测作用很小,但对抽样攻击和流

行攻击类型的检测结果却很好,而ADegSim检测指标对分段攻击类型的检测效果很差,对其他托攻击类型效果却很好.因此,如果我们盲目地选取检测指标,很难应对所有种类的托攻击.

第二组实验研究选择的检测指标数 m 对 NBC-SDA 和 kNN-SDA 性能的影响.实验根据表 1 生成各攻击类型的用户模型,装填比例为10%,注入 MovieLens 数据集,形成训练集和测试集,NBC-SDA 和 kNN-SDA 依据训练集检测测试集中的托攻击用户.训练集中正常用户数量为 100,托攻击者数量为 50.而测试集中正常用户数量为 943,攻击者数量为 94.实验改变 m 值,得到 NBC-SDA 和 kNN-SDA 对五种攻击类型的灵敏性和特效性.MovieLens 共给出 5 对训练集和测试集,本实验最终取 5 对数据集上灵敏性和特效性的平均值.图 2 给出了实验结果,图 2(a)和(b)分别展示了 NBC-SDA 的灵敏性和特效性,当 m 取 3~5 时,NBC-SDA 灵敏性和特效性达到最大.从图 2(c)看出,除了抽样攻击之外,kNN-SDA 灵敏性接近于 1,因此,m 值变化未引起 kNN-SDA 灵敏性的波动.但是,m 值变化对 kNN-SDA 特效性影响颇大(见图 2(d)),同样 m 取 3~5 时,kNN-SDA 特效性达到最大.从图 2 也可看出,选取所有检测指标时(m = 10),检测性能不佳.这说明利用本文 4.2 节所述的特征选择算法选取适中数量的指标能大幅度提高托攻击检测效果.

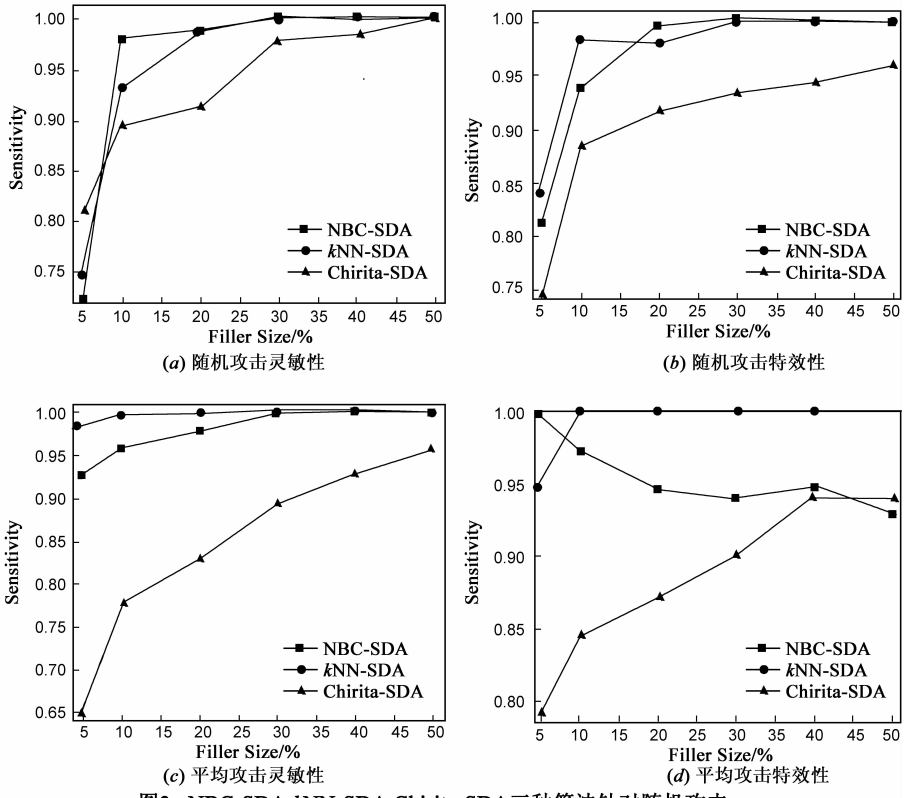
由于 Chirita-SDA 是针对随机攻击和平均攻击的,因此第三组实验比较了 NBC-SDA、kNN-SDA 和 Chirita-SDA 针对随机攻击和平均攻击的性能.用与第二组实验中所述的同样方法生成训练集和测试集,设置 m = 4,并统计 5 对数据集上的平均值作为最终结果.图 3 给出了实验结果,NBC-SDA 和 kNN-SDA



(a) NBC-SDA灵敏性　(b) NBC-SDA特效性

(c) kNN-SDA灵敏性　(d) kNN-SDA特效性

图2　m对 NBC-SDA, kNN-SDA两种算法针对五种攻击检测性能



(a) 随机攻击灵敏性　(b) 随机攻击特效性

(c) 平均攻击灵敏性　(d) 平均攻击特效性

图3　NBC-SDA,kNN-SDA,Chirita-SDA三种算法针对随机攻击

表现出了很好的性能,尤其是当装填项目比例较大时,灵敏性和特效性趋向于 100%.而 Chirita-SDA 检测随机攻击的灵敏性较好,而特效性较差,因为,存在近一半正常用户的 RDMA 高于平均值.Chirita-SDA 检测平均攻击的性能更差,因为平均攻击者模型的 RDMA 极小,容易被误认为是正常用户.

第四组实验比较了 NBC-SDA 和 kNN-SDA 针对分段攻击、流行攻击和抽样攻击的性能,实验结果如图 4 所示.抽样攻击没有装填项目,因此,实验分别考察了不同长度用户模型的情形,同样设置 $m=4$.两种算法各有千秋,kNN-SDA 对分段攻击的检测优于 NBC-SDA,而对流行攻击却不如 NBC-SDA.而且 kNN-SDA 不适用于检测抽样攻击,原因在于抽样攻击拷贝了正常用户模型,其近邻大部分仍然是正常用户,kNN-SDA 极易将攻击者判定为正常用户,而正常用户也很容易认为是正常用户,导致 kNN-SDA 对抽样攻击的灵敏性很低,特效性却较高(如图 4(e)和(f)所示).

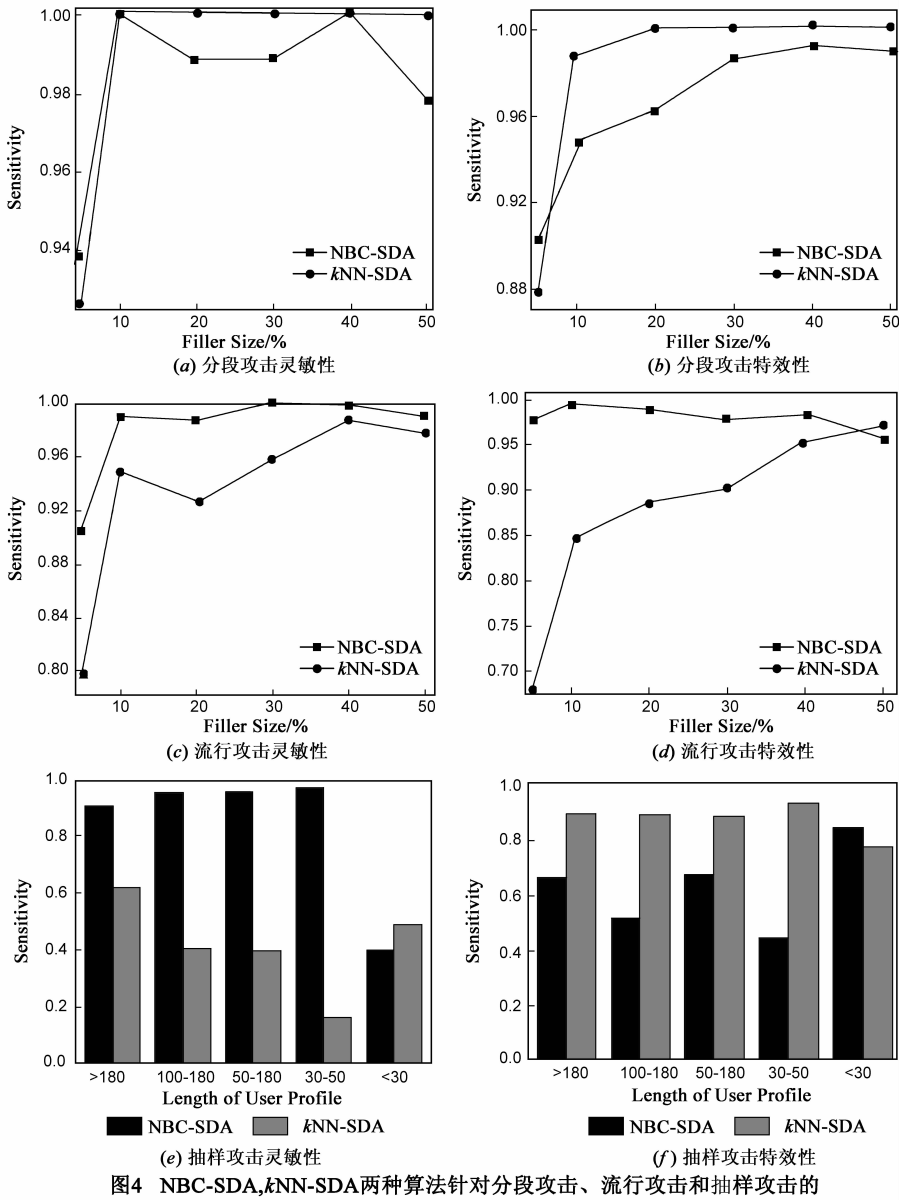


图4 NBC-SDA,kNN-SDA两种算法针对分段攻击、流行攻击和抽样攻击的灵敏性和特效性对比图

7 结束语

本文研究基于协同过滤推荐系统中的托攻击检测问题,基于五种托攻击的分类,首先,分析了五种托攻击对不同协同过滤算法的危害性.其次,提出一种特征选择算法,基于训练集选取有效的检测指标.再次,提出两种基于监督学习的托攻击检测算法:NBC-SDA 和 kNN-SDA,根据选取出的指标检测托攻击.最后,在 MovieLens 数据集上验证了算法的有效性,实验结果表明,不同类型托攻击的检测指标排序大相径庭,利用本文提出的特征选择算法选取适中数量的指标能大幅度提高托攻击检测效果.NBC-SDA 和 kNN-SDA 明显优于以前的算法,两种算法检测不同托攻击时各有优劣.

在下一步的工作中,我们计划将本文成果应用到实际电子商务系统中,检测隐藏在海量用户中的托攻击者.

参考文献

[1] 张锋,孙雪冬,常会友,赵淦森.两方参与的隐私保护协同过滤推荐研究[J].电子学报,2009,37(1):84 - 89.
F Zhang, X Sun, H Chang, G Zhao. Research on privacy-preserving two-party collaborative filtering recommendation[J]. Acta Electronic Sinica, 2009, 37(1): 84 - 89. (in Chinese)

[2] S Zhang, A Chakrabarti, J Ford, F Makedon. Attack detection in time series for recommender systems [A]. Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD′06)[C]. Philadelphia, Pennsylvania, USA, 2006.

[3] SK Lam, J Riedl. Shilling recommender systems for fun and

profit [A]. Proceedings of the 13th World Wide Web (WWW' 04)[C]. New York, USA, 2004, 393 – 402.

[4] P A Chirita, W. Nejdl, C Zamfir. Preventing shilling attacks in online recommender systems [A]. Proceedings of ACM Int. Workshop on Web Information and Data Management[C]. ACM Press, New York, NY, USA, 2005. 67 – 74.

[5] R Burke, B Mobasher, C Williams R Bhaumik. Classification features for attack detection in collaborative recommendation systems [A]. Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'06)[C]. Philadelphia, Pennsylvania, USA, 2006, 542 – 547.

[6] B Mehta, W Nejdl. Unsupervised strategies for shilling detection and robust collaborative filtering [J]. User Modeling and User-Adapted Interaction, 2009, 19(1): 65 – 97.

[7] N Hurley, Z Cheng, M Zhang. Statistical attack detection [A]. Proceedings of ACM Conference on Recommender Systems (RecSys '09)[C] . New York, USA, 2009, 149 – 156.

[8] H Liu, L Yu. Toward integrating feature selection algorithms for classification and clustering [J]. IEEE Transactions on Knowledge and Data Engineering, 2005, 17(4): 491 – 502.

[9] 蒋盛益,郑琪,张倩生. 基于聚类的特征选择方法[J]. 电子学报, 2008, 37(12): 157 – 160.
S Jiang, Q Zheng, Q Zhang. Clustering-based feature selection [J]. Acta Electronica Sinica, 2008, 37(12): 157 – 160. (in Chinese)

[10] J L Herlocker, J A Konstan, LG Terveen, JT Riedl. Evaluating collaborative filtering recommender systems[J]. ACM Transaction on Information Systems, 2004, 22(1): 5 – 53.

[11] C Williams. Profile Injection Attack Detection for Securing Collaborative Recommender Systems[R]. DePaul University CTI Technical Report, 2006.

作者简介

伍之昂　男, 1982 年生于江苏宜兴, 博士, 现为南京财经大学江苏省电子商务重点实验室副教授, 主要研究领域为推荐系统, 云计算和数据挖掘.
E-mail: zawu@seu.edu.cn

庄　毅　男, 1978 年生于浙江杭州, 博士, 浙江工商大学副教授, 获 2008 中国计算机学会优秀博士论文奖, 研究方向为不确定数据管理、多媒体数据库等.
E-mail: zhuang@mail.zjgsu.edu.cn