基于人工免疫危险理论的微博水军用户检测研究

杨 超[1] 秦廷栋[1] 范 波[2] 李 涛[3]

(湖北大学计算机与信息工程学院 武汉 430062)[1] (武汉大学科学技术发展研究院 武汉 430072)[2]

(智能信息处理与实时工业系统湖北省重点实验室(武汉科技大学) 武汉 430065)[3]

摘 要 将人工免疫危险理论引入到用户行为特征的分析中,以有效地识别微博水军用户。以新浪微博为例,分析了新浪微博水军的行为特征,选取微博总数、微博等级、是否认证、阳光信用、粉丝数等特征属性,将属性分析结果作为区别水军与正常用户的特征信号,并基于树突状细胞算法(Dendritic Cells Algorithm,DCA)实现新浪微博水军的识别。使用新浪微博用户的真实数据对算法的有效性进行了验证和对比实验,结果表明该方法能够有效检测出新浪微博中的水军用户,具有较高的检测准确率。

关键词 微博水军,行为特征,人工免疫,危险理论,树突状细胞算法

中图法分类号 TP393 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2018.11.020

Study on Detection of Weibo Spammers Based on Danger Theory in Artificial Immunity System

YANG Chao[1] QIN Ting-dong[1] FAN Bo[2] LI Tao[3]

(School of Computer Science and Information Engineering, Hubei University, Wuhan 430062, China)[1]

(Office of Scientific Research and Development, Wuhan University, Wuhan 430072, China)[2]

(Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan University of Science and Technology, Wuhan 430065, China)[3]

Abstract This paper introduced the danger theory in artificial immunity system into the analysis of user behavior characteristics to identify the spammers in Weibo effectively. Taking Sina Weibo as an example, this paper analyzed the behavior characteristics of Weibo spammers, selected the total number of Weibo, Weibo level, user authentication, sunshine credit and the number of fans as attribute characteristics and used the analysis results of attribute characteristics as the characteristic signals of distinguishing the spammers and the normal users. After that, the recognition of Sina Weibo spammers can be achieved based on Dendritic Cells Algorithm. The real data of Sina Weibo users was used to verify the effectiveness of the proposed algorithm and conducted comparison experiments. The experimental results suggest that this algorithm can effectively detect the spammers in Sina Weibo and has high detection accuracy.

Keywords Weibo spammers, Behavioral characteristics, Artificial immunity, Danger theory, Dendritic cells algorithm

1 引言

近年来,微博作为一种新的社交网络平台迅速崛起并飞速发展,已成为民众获取最新资讯、参与舆论话题、发表言论的重要途径[1]。截至 2017 年 6 月,微博用户规模达到 29017 万,手机微博用户规模达到 24086 万,用户使用率为 35.7%,人均单日访问次数持续快速增长,其中新浪微博人均单日访问次数达到 2.8 次[2]。随着微博类社交平台的日益普及,近年来有组织的微博网络水军开始出现并快速泛滥,其肆意发表的倾向性言论,导致社交网络中的谣言信息盛行,混淆信息的真实性,造成错误的舆论导向,损害公民利益,影响正常网络秩序,危及社会稳定及国家政治,因此进行微博网络水军识别迫在眉睫。

本文总结了各类网络水军识别的相关研究。水军的识别方法主要分为以下 3 类。

(1)基于内容特征的方法。其包括文本分类[3]、文本情感分析[4]以及文本倾向性分析[5]等方法,如金礼仁提出的基于结构与内容的识别方法[6],利用网络拓扑结构构建网络水军的重叠社区结构,计算网络节点中的内容与垃圾信息相似度,综合用户结构与发布内容的特征识别网络水军。由于网络环境的复杂化和各类网络平台实名制的约束,水军由以往的系统批量操作生成,逐渐转变为一种被真实用户操作的新型水军,后者制造的垃圾信息趋向于正常用户,不再具有显著的可识别特征,因此该方法不能有效发现新型网络水军,较难实现对社交网络水军的实时检测[7]。

(2)基于环境特征的方法。其依据水军造成危害时产生

到稿日期:2017-12-04 返修日期:2018-03-16 本文受武汉科技大学智能信息处理与实时工业系统湖北省重点实验室基金(znxx2018MS05)资助。

杨 超(1982-),男,博士,副教授,主要研究方向为信息安全;秦廷栋(1992-),男,硕士生,主要研究方向为信息安全;范 波(1985-),男,硕士,工程师,主要研究方向为计算机网络,E-mail:fanbo@whu.edu.cn(通信作者);李 涛(1979-),男,博士,副教授,主要研究方向为信息安全。

的环境特征进行水军识别。由于环境特征无法被修改、掩饰,因此该方法识别准确率较高。如 Ramachandran 等提出的水军网络级特征分析方法[8],从被水军污染的领域中追踪收集 1000 万封垃圾邮件信息,将该数据与 TCP 脚印信息、IP 黑名单信息、机器人网站命令追踪以及路由信息等联系起来对水军的网络级特征进行分析,实现水军追踪。但该方法需要相应的大量实验数据集,可推广性较低,不适用于微博等社交网络的水军检测。

(3)基于用户特征的方法。其包括用户行为特征分析[9]和用户关系特征分析[10]。如张艳梅等[11]提出的基于贝叶斯模型的识别方法,利用相关特征属性设计水军识别分类器,运用遗传算法得到判别水军的阈值矩阵,并利用贝叶斯模型建立概率矩阵来表示水军特征在不同阈值内的情况,完成对水军的检测识别。该方法能够很好地发现潜藏的网络水军,且较为适用于社交网络平台环境下的水军检测。但其存在特征描述不全面、对于多指标的海量数据处理效率较低等问题。

社交网络环境中,各类安全问题的出现往往是由用户的异常行为所导致的,用户异常行为的分析和检测与入侵检测具有较高的相似性。人工免疫系统在入侵检测领域具有较为广泛的应用,如利用人工免疫危险理论中的树突状细胞算法(Dendritic Cell Algorithm,DCA)构建集成入侵检测(RSAI-IID)模型[12],或进行垃圾邮件群发检测[13]与 Web 服务器异常检测[14]等。

本文拟将人工免疫的思想应用于微博用户行为特征的检测中,使用基于用户行为特征的分析方法刻画定义网络水军行为,将人工免疫危险理论的信号处理机制应用于网络水军检测,采用危险理论的核心算法——树突状细胞算法(DCA),来检测微博中的水军用户。实验结果表明,采用合理的水军用户特征指标,可以有效检测出微博平台中的水军用户。

2　树突状细胞算法概述

Matzinger 于 1994 年提出了全新的"危险理论"生物免疫应答模式[15],Aickelin 等将其引入并成功应用于人工免疫系统中[16],随后该小组的 Aickelin 等于 2005 年提出了基于"危险理论"的树突状细胞算法(Dendritic Cell Algorithm,DCA)[17]。

树突状细胞(Dendritic Cell,DC)是来自于生物免疫系统中骨髓内髓样前体细胞的专职抗原提呈细胞,其功能强大,是生物机体抵抗体外有害抗原入侵的一道坚固壁垒。DCA 是基于 DC 群体的算法[18],群体中的 DC 参与收集组织中的抗原时,摄取病原体相关分子模式信号 PAMP、危险信号 DS、安全信号 SS 和致炎信号 IS 4 种环境信号,并将摄取到的 4 种环境信号作为输入值,使用相关函数及权值矩阵进行融合处理,在输出协同刺激信号 CSM、半成熟信号 SEMI 和成熟信号 MAT 等 3 种信号后,对各类信号分别进行累加,DC 在 CSM 达到迁移阈值时发生迁移,此时细胞环境状态取决于 SEMI 和 MAT 的浓度较大的一方,其中 SEMI 和 MAT 分别表示当前抗原环境的安全程度和危险程度。当所有抗原达到判别次数后,使用 SEMI 和 MAT 值来计算环境中的抗原异常程度 MCAV,并与已设定的异常阈值进行比较,从而判断该抗原是否异常[19]。

生物免疫学中 7 种信号在算法应用中的含义[20]如表 1 所列。

表 1　7 种信号及其抽象信号与含义

Table 1　Abstract signals and meanings of seven signals

生物学信号	功能	抽象信号	含义
PAMP	表明存在病原体	PAMP	表明特征存在异常
坏死信号	表明组织坏死	DS	表明异常的可能性高
凋亡信号	表明组织健康	SS	表明正常的可能性高
促炎细胞因子	表明组织总体上存在损伤	IS	与所有其他信号相乘的因子
CD80/86	协同刺激分子	CSM	协同刺激信号
IL-10	smDC 分泌的细胞因子	SEMI	正常信号
IL-12	mDC 分泌的细胞因子	MAT	异常信号

DCA 算法具有不依赖自体库、计算效率高、可减少误报率和漏报率等特点[20]。基于此,本文将微博网络环境抽象为生物机体,分析微博水军用户的行为特征,将各类特征映射并定义为 DCA 算法模型中的各种信号,设计出一种新的微博网络水军用户检测方法。

3　基于 DCA 算法的微博水军检测

3.1　微博水军特征定义

微博用户的行为虽然有一定的随机性,但蕴含着很多特定的行为模式,如用户注册、发布微博、转发、评论、点赞等。用户在使用过程中会产生相应的注册时间、终端、个人信息、信用度、认证等信息。本文以目前国内用户规模最大的新浪微博平台为研究对象,通过对近几年微博水军特征研究的分析,结合相关文献的观点,得出如下分析结论。

(1)由于水军用户关注和评论其他用户通常都带有一定的目的性,而其有价值的原创微博数少且影响力弱,造成微博水军用户的关注数远高于粉丝数,同时水军用户与其粉丝之间的相互关注数极少[21]。

(2)水军用户本身的影响力极小,通常是由群体行为造成较大的舆论、社交关系等影响。

(3)普通用户在建立微博后,一定周期内平均发布微博的频率和时间段符合一定的幂率规律,而水军用户则与正常用户存在较大差异[22]。

(4)水军用户一般为了达到某种宣传的目的,在短期内申请一定数量的账号,并在某时间段内发布大量的微博,或聚集在某条微博下进行大量的评论[23]。

(5)水军发布的微博通常没有用户留言评论,且缺乏完善的简介、认证等,同时等级数、阳光信用值偏低。"阳光信用"属性被新浪微博官方定义为:对用户的信用历史、社交关系、消费偏好等进行综合分析评价,其结果作为衡量微博用户在网络上积极表达、阳光讨论、理性交流的标尺[24]。该属性是新浪微博为了评价用户信用度新增的一个特征值,可以增强水军识别能力,对于区分水军具有关键性作用。

综上所述,本文选取粉丝数、关注数、微博总数、原创微博数、是否认证、微博等级、有无简介、注册时间、阳光信用、互相关注数、参与话题数、评论数、转发数和点赞数共 14 种用户行为特征,并将其融合成 6 种能够区分水军与非水军的重要属性,具体的融合方法请见定义 1—定义 6。

基于上述微博用户的特点，分析一定时间段内的微博用户情况，并以用户近期发布的前 20 条微博数据作为样本，采取以下指标作为抗原信号，通过多次对比实验和经验总结，确定定义公式中各项指标的权重系数。

定义 1(阳光信用(*SC*)) 其分为极低(300～419)、较低(420～450)、一般(451～570)、较好(571～690)、极好(691～900) 5 个等级，在实验中分别使用数值 1－5 来表示。

定义 2(活跃度(*AT*)) 其涉及多个属性变量，包括微博总数(*M*)、参与话题数(*T*)、注册时间(*Z*)、当前时间(*N*)，其中“*N*－*Z*”表示用户已经使用微博时间，以“天”为单位，计算方式如式(1)所示：

$$AT=(0.7M+0.3T)/(N-Z) \tag{1}$$

定义 3(身份评价(*IE*)) 涉及多个属性变量，分别为有无简介(*I*)、是否认证(*C*)和等级数(*G*)，各个属性的权重分别为 0.2，0.4，0.4，计算方式如式(2)所示：

$$IE=0.2I+0.4C+0.4G \tag{2}$$

定义 4(影响力(CI)) 涉及多个属性变量，分别为评论数(*J*)、转发数(*R*)、点赞数(*F*)，及用户所发微博的被评论数、被转发数和被点赞数，各个属性的权重分别为 0.3，0.5，0.2，计算方式如式(3)所示：

$$CI=0.3J+0.5R+0.2F \tag{3}$$

定义 5(粉丝关注比(*FF*)) 每个用户的粉丝数(*Fans*)与关注数(*Followers*)的比值，计算方法如式(4)所示：

$$FF=Fans/Followers \tag{4}$$

定义 6(原创微博比(*OM*)) 用户所发送微博中原创微博(*Weibo_Original*)所占微博总数(*M*)的比例，计算方式如式(5)所示：

$$OM=(Weibo\_Original)/M \tag{5}$$

3.2 数据预处理及算法实现

3.2.1 信号映射

对前文提到的微博用户行为的各项特征指标数据进行预处理，统一将其规格化到 1 到 10 之间。定义规格化函数 $f(x)$如式(6)所示：

$$f(x)=\begin{cases}\dfrac{x-m}{n-m}*10, & x\in[m,n]\\ 10, & x>n\end{cases} \tag{6}$$

其中，x 是原始信号值，当 $x\in[m,n]$时，进行线性映射，当 $x\in[n,\infty]$时，信号取最大值 10。

现将检测新浪微博水军用户的相关指标进行如下映射。

(1)病原体相关分子模式 *PAMP*：表明用户行为非健康，存在水军行为的特征，定义 $PAMP=\{\langle SC,IE,FF\rangle\}$；

(2)危险信号 *DS*：表明用户亚健康，异常的可能性较高，虽然可能只是正常的行为改变，但存在水军行为的可能，定义 $DS=\{\langle AT,CI,OM\rangle\}$；

(3)安全信号 *SS*：表示用户正常的可能性较高，并处于正常状态，定义 $SS=\{\langle SC,IE\rangle\}$；

(4)致炎因子 *IS*：表明当前用户总体上存在异常，起到放大 *PAMP*，*DS*，*SS* 信号的作用，定义 $IS=\{\langle CI\rangle\}$。

3.2.2 权值矩阵及算法公式

如本文第 2 节所述，DC 收集组织中的抗原，对输入信号 *PAMP*，*DS*，*SS*，*IS* 进行计算处理，输出 *CSM*，*SEMI*，*MAT* 信号。由输入信号值到输出信号值的转换权值矩阵如表 2 所列。

表 2 DCA 信号转换权值矩阵

Table 2 Weight matrix of DCA signals conversion

权值(*W*)	*CSM*	*SEMI*	*MAT*
PAMP	4	0	8
DS	2	0	4
SS	3	1	−6

表 2 中的权值矩阵表明，*PAMP* 和 *DS* 对 *SEMI* 信号的影响值为 0，但对 *CSM* 信号及 *MAT* 信号为正影响，起促进作用。*SS* 对 *MAT* 信号则具有负影响，对 *CSM* 信号和 *SEMI* 信号值具有正影响。

输入信号值转换为输出信号值的计算如式(7)所示：

$$C_{[CSM,SEMI,MAT]}=\frac{(W_P\times C_P)+(W_D\times C_D)+(W_S\times C_S)\times(1+IS)}{|W_P|+|W_D|+|W_S|} \tag{7}$$

其中，$(1+IS)$为放大信号，输入信号 *PAMP*，*DS*，*SS* 对应的值和权值分别为 C_P，C_D，C_S 以及 W_P，W_D，W_S。权值表示某信号在激活树突状细胞功能中的影响力，根据多次实验结果对其进行调整并确定最终的数值，权值越大，影响力就越大。

算法综合评价阶段，采用细胞成熟抗原值(*MCAV*)对抗原进行最终的评价，其计算公式如式(8)所示，*MCAV* 的值越接近 1，该抗原的异常程度就越高。

$$MCAV=MAT/(SEMI+MAT) \tag{8}$$

3.2.3 算法流程

根据文献[20]，本文中微博水军用户检测所使用的 DCA 算法的流程如图 1 所示。

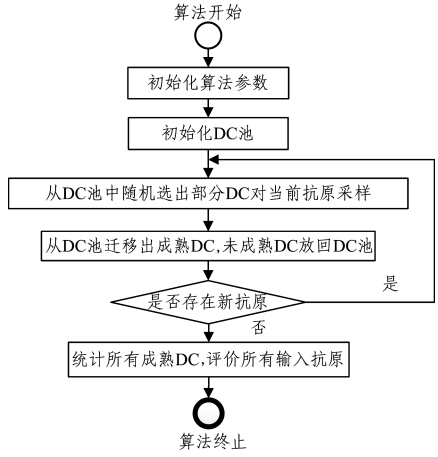

图 1 DCA 算法流程

Fig. 1 Flowchart of DCA algorithm

算法 1 DCA 算法

Begin

初始化抗原池采集数目和 DC 细胞种群；

For 每一个 DC

If CSM＜迁移阈值 Then

Begin

在抗原池中随机采集未被标记的抗原，并计算 PAMP，DS，SS 及 IS 信号值 and 依据权值矩阵计算 CSM，semiMature，Mature 并累加；

End

```
If CSM>=迁移阈值 Then
    Begin
    比较 semiMature，Mature 大小，并标记该 DC 的状态 and 标记该
    DC 采集的抗原状态 and DC 发生迁移；
    If 抗原判定总次数=抗原判别阈值 Then
      Begin
      计算 MCAV=抗原被判为异常次数/抗原判定总次数，比较
      MCAV 与异常阈值的大小；
      If MCAV>=N then 抗原标记为异常；
      Else 抗原标记为正常；
      End
    End
    If DC 采集抗原次数>=DC 生命周期 Then
    抛弃该 DC；
End
```

4　实验

4.1　数据集获取

借鉴文献[25]中的新浪微博数据获取方法，本文的实验数据通过调用新浪微博 API 接口和 Python 语言编写网络爬虫两种方式进行获取。爬虫程序中模拟登录的部分代码的描述如下：

```
defget_info():
url='http://login.sina.com.cn/'
    data=urllib2.urlopen(url).read()
    p=re.compile('\((.*)\)')
    try:
json_data=p.search(data).group(1)
    data=json.loads(json_data)
servertime=str(data['servertime'])
    nonce=data['nonce']
publicKey=data['pubkey']
rsakey=data['rsakv']
    return servertime,nonce,publicKey,rsakeyst,non,pubkey,rsakv=
    get_info()
```

使用以上方法在数据池内随机获取 11764 条数据，其中水军数据 2367 条，非水军数据 9397 条。水军用户以购买方式获取，并标记所爬取的水军用户数据；非水军用户以可靠关系获取，并加以人工验证和标记。

4.2　评价指标

为解决数据不平衡问题，本文建立混淆矩阵(Confusion Matrix)来分析实验结果[26]。该矩阵用来呈现算法性能的可视化效果，正确的检测结果在对角线上，可以直观且快速地观察到角线外的错误。矩阵中，TP(True Positive)、TN(True Negative)分别是样本中的检测水军数和检测非水军数，FN(False Negative)、FP(False Positive)分别是识别错误的实际水军数和非水军数。混淆矩阵如表 3 所列。

表 3　混淆矩阵

Table 3　Confusion matrix

混淆矩阵		实际样本数	
		Positive	Negative
检测样本数	Positive	TP	FP
	Negative	FN	TN

算法结果的评估主要包括准确率(Precision Rate，PR)、召回率(Recall Rate，RR)、调和平均值 F1 等评价指标，本文采用此 3 种指标来评估算法的准确性，其中准确率和召回率用来评价实验准确性，调和平均值是对实验综合表现的评价指标。各指标的定义如下。

(1)准确率 PR 的定义如式(9)所示：

$$PR=\sqrt{class^{+}\times class^{-}} \tag{9}$$

其中，$class^{+}=TP/(TP+FP)$和 $class^{-}=TN/(TN+FN)$分别表示分类器对水军类样本和非水军类样本的分类准确率，PR 表示分类器平均准确率。平均准确率 PR 的高低由 $class^{+}$和 $class^{-}$两者和值的高低决定。

(2)召回率 RR 的定义如式(10)所示：

$$RR=TP/(TP+FN) \tag{10}$$

其中，RR 为水军的召回率。

(3)调和平均值 $F1$ 的定义如式(11)所示：

$$F1=(2*PR*RR)/(PR+RR) \tag{11}$$

4.3　结果分析

实验所用抗原是从数据集中随机抽取的 1000，2000，5000 和 10000 条用户数据样本，分别对这 4 种实验情景进行 10 次实验，对实验结果平均值向上取整并进行分析。水军检测结果如表 4 所列。

表 4　新浪微博水军检测结果

Table 4　Detection results of Sina Weibo spammers

实验情景	水军数/样本数	识别出水军数	识别正确水军数
1	223/1000	216	196
2	442/2000	452	409
3	1337/5000	1204	1143
4	2192/10000	2231	2027

通过以上实验检测识别结果，分别针对 4 种实验情景建立混淆矩阵，计算对应的准确率 PR、召回率 RR 以及调和平均值 $F1$，最终取 4 种实验情景下各类结果的平均值作为对实验的最终评价指标。算法结果分析如表 5 所列。

表 5　结果分析

Table 5　Results analysis

(单位：%)

实验情景	准确率 PR	召回率 RR	调和平均值 $F1$
1	93.60	87.89	90.66
2	94.10	92.53	93.31
3	94.91	85.49	89.95
4	94.30	92.47	93.38
平均值	94.23	89.60	91.83

由表 5 可以得出，树突状细胞算法对于微博网络水军的识别平均准确率可达到 94.23%，对应的漏报率和误报率均较低，因此基于危险理论的树突状细胞 DCA 算法可以有效地识别微博社交网络中的水军用户。

为验证实验的有效性，本文选取前文所提及的 3 种水军识别方法进行对比实验，比较实验结果中的准确率、召回率及调和平均值 3 个指标。由于缺乏 IP、TCP 及路由等信息数据，暂未与基于环境特征的识别方法进行比较。实验对比方法主要包括文献[6]提出的基于结构与内容的识别方法、文献[11]提出的基于贝叶斯模型的识别方法和本文提出的基于

DCA 算法的模型识别方法。

实验随机抽取数据集中的 1000,2000,5000 和 10000 条用户数据作为 4 组数据样本,使用 3 种识别方法分别对此 4 组数据进行 10 次实验,取实验结果的平均值进行比较,对比结果如图 2 所示。实验结果表明,基于结构与内容的识别算法对新型网络水军识别的准确率较低;基于贝叶斯模型的识别方法有较高的准确率,但对高量级的数据处理时召回率较低;DCA 算法对于新型网络水军的识别检测有较好的适用性和较高的准确率。

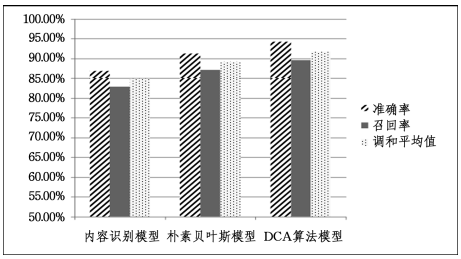

图 2 实验结果对比

Fig. 2 Comparison of experimental results

结束语 本文借鉴生物免疫系统的思想,提出了运用危险理论中的 DCA 算法进行微博网络水军用户的检测。通过分析新浪微博中水军用户的行为特征,根据微博正常用户与水军用户在转发、评论、阳光信用等特征表现上的差异,判断是否存在水军行为。实验结果表明,该方法的检测准确率较高,是检测微博水军用户的一种有效方法。

由于新浪微博等社交平台的实名认证限制,微博水军无法依靠批量注册新用户来生成水军账号,转而通过购买或盗取微博中综合评价较高的正常用户账号,并将其加入微博水军的行动中,因此水军用户的表现特征日趋接近正常用户。为了进一步提高微博水军检测的准确率,未来可使用文本分析、聚类分析等方法对用户发布的文字信息进行筛选、捕捉和分类,以此作为微博水军识别算法中的重要特征指标,从而提升检测准确率。

参 考 文 献

[1] Beijing Internet Information Office. China Weibo Development Report [M]. Beijing: People's Publishing House, 2014: 59-106. (in Chinese)
北京互联网信息办公室. 中国微博发展报告[M]. 北京:人民出版社,2014:59-106.

[2] CNNIC. The 40th "China Internet Development Statistics Report" [EB/OL]. http://www.cnnic.net.cn/hlwfzyj/hlwxzbg/hlwtjbg/201708/t20170803_69444.htm. (in Chinese)
中国互联网络信息中心. 第 40 次《中国互联网络发展状况统计报告》[EB/OL]. http://www.cnnic.net.cn/hlwfzyj/hlwxzbg/hlwtjbg/201708/t20170803_69444.htm.

[3] SRIRAM B, FUHRY D, DEMIR E, et al. Short text classification in twitter to improve information filtering[C]// International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2010: 841-842.

[4] ZHAO Y Y, QIN B, LIU T. Sentiment Analysis[J]. Journal of Software, 2010, 21(8): 1834-1848. (in Chinese)
赵妍妍,秦兵,刘挺. 文本情感分析[J]. 软件学报,2010,21(8):1834-1848.

[5] LIUB. Sentiment analysis and subjectivity[J/OL]. https://www.researchgate.net/publication/228667268_Sentiment_analysis_and_subjectivity.

[6] JIN L R. Structure and Content-based SpammerDection in Social Networks[D]. Nanjing: Nanjing University of Posts, 2016. (in Chinese)
金礼仁. 基于结构与内容的社交网络水军团体识别[D]. 南京:南京邮电大学,2016.

[7] MO Q, YANG K. Overview of Web Spammer Detection[J]. Journal of Software, 2014, 25(7): 1505-1526. (in Chinese)
莫倩,杨珂. 网络水军识别研究[J]. 软件学报,2014,25(7):1505-1526.

[8] RAMACHANDRAN A, FEAMSTER N. Understanding the network-level behavior of spammers[C]// ACM. 2006: 291-302.

[9] BHAT V H, MALKANI V R, SHENOY P D, et al. Classification of email using BeaKS: Behavior and keyword stemming[J/OL]. https://ieeexplore.ieee.org/document/6129290.

[10] BRENDEL R, KRAWCZYK H. Application of social relation graphs for early detection of transient spammers[M]. World Scientific and Engineering Academy and Society (WSEAS), 2008: 267-276.

[11] ZHANG Y M, HUANG Y Y, GAN S J, et al. Weibo spammers' identification algorithm based on Bayesian model[J]. Journal on Communications, 2017, 38(1): 44-53. (in Chinese)
张艳梅,黄莹莹,甘世杰,等. 基于贝叶斯模型的微博网络水军识别算法研究[J]. 通信学报,2017,38(1):44-53.

[12] LING Z, BAI Z Y, LUO S S, et al. Integrated intrusion detection model based on rough set and artificial immune[J]. Journal on Communications, 2013, 34(9): 166-176. (in Chinese)
张玲,白中英,罗守山,等. 基于粗糙集和人工免疫的集成入侵检测模型[J]. 通信学报,2013,34(9):166-176.

[13] YANG C, LI Z Y. Spam mass sending examination based on dendritic cell algorithm[J]. Transducer & Microsystem Technologies, 2015, 34(10): 133-136. (in Chinese)
杨超,李子怡. 基于树突状细胞算法的垃圾邮件群发检测[J]. 传感器与微系统,2015,34(10):133-136.

[14] WANG X X, LIANG Y W, et al. Application of dendritic cell algorithm on Web server anomaly detection[J]. Computer Engineering & Applications, 2016, 52(24): 148-152. (in Chinese)
王新新,梁意文. 树突状细胞算法在 Web 服务器异常检测中的应用[J]. 计算机工程与应用,2016,52(24):148-152.

[15] MATZINGER P. Tolerance, Danger and the Extended Family [J]. Annual Review of Immunology, 1994, 12(1): 991-1045.

[16] AICKELIN U, BENTLEY P, CAYZER S, et al. Danger Theory: The Link between AIS and IDS? [C]// Artificial Immune Systems, Second International Conference (ICARIS 2003). Edinburgh, UK, Proceedings. DBLP, 2003: 147-155.

[17] AICKELIN U, GREENSMITH J, TWYCROSS J. Immune System Approaches to Intrusion Detection—A Review[C]// International Conference on Artificial Immune Systems. Springer Berlin Heidelberg, 2004: 316-329.

(下转第 159 页)

参 考 文 献

[1] MA X, LI J, ZHANG F. Outsourcing computation of modular exponentiations in cloud computing[J]. Cluster Computing, 2013, 16(4): 787-796.

[2] HOHENBERGER S, LYSYANSKAYA A. How to Securely Outsource Cryptographic Computations[C]// International Conference on Theory of Cryptography. 2005: 264-282.

[3] CHEN X, LI J, MA J, et al. New Algorithms for Secure Outsourcing of Modular Exponentiations[C]// European Symposium on Research in Computer Security. Springer Berlin Heidelberg, 2012: 541-556.

[4] GOLLE P, MIRONOV I. Uncheatable Distributed Computations [C]// Topics in Cryptology-CT-RSA 2001, The Cryptographer's Track at RSA Conference 2001. DBLP, 2001: 425-440.

[5] DING Y, XU Z, YE J, et al. Secure outsourcing of modular exponentiations under single untrusedprogrammemodel[J]. Journal of Computer & System Sciences, 2017, 90(1): 1-13.

[6] SU Q, YU J, TIAN C, et al. How to securely outsource the inversion modulo a large composite number[J]. Journal of Systems & Software, 2017, 129(C): 26-34.

[7] YE J, XU Z, DING Y. Secure outsourcing of modular exponentiations in cloud and cluster computing[J]. Cluster Computing, 2016, 19(2): 811-820.

[8] PAILLIER P. Public-Key Cryptosystems Based on Composite Degree Residuosity Classes[J]. Lecture Notes in Computer Science, 1999, 547(1): 223-238.

[9] ATALLAH M J, FRIKKEN K B. Securely outsourcing linear algebra computations[C]// ACM Symposium on Information, Computer and Communications Security. ACM, 2010: 48-59.

[10] BENJAMIN D, ATALLAH M J. Private and Cheating-Free Outsourcing of Algebraic Computations[C]// Sixth Conference on Privacy, Security and Trust. IEEE Computer Society, 2008: 240-245.

[11] REN Y, DING N, ZHANG X, et al. Verifiable Outsourcing Algorithms for Modular Exponentiations with Improved Checkability[C]// ACM on Asia Conference on Computer and Communications Security. 2016: 293-303.

[12] ZHAO L, ZHANG M, SHEN H, et al. Privacy-preserving Outsourcing Schemes of Modular Exponentiations Using Single Untrusted Cloud Server[J]. Ksii Transactions on Internet & Information Systems, 2017, 11(2): 826-845.

[13] REN K, WANG C, WANG Q. Security Challenges for the Public Cloud[J]. IEEE Internet Computing, 2012, 16(1): 69-73.

[14] WANG C, CAO N, REN K, et al. Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data[J]. IEEE Transactions on Parallel & Distributed Systems, 2012, 23(8): 1467-1479.

[15] CHUNG K M, KALAI Y, VADHAN S. Improved delegation of computation using fully homomorphic encryption[M]// Advances in Cryptology — CRYPTO 2010. Berlin: Springer-Verlag, 2010: 483-501.

[16] GENNARO R, GENTRY C, PARNO B. Non-interactive Verifiable Computing: Outsourcing Computation to Untrusted Workers [M]// Advances in Cryptology - CRYPTO 2010. Berlin: Springer-Verlag, 2010: 465-482.

[17] BOYKO V, PEINADO M, VENKATESAN R. Speeding up Discrete Log and Factoring Based chemes via Precomputations[M]// Advances in Cryptology — EUROCRYPT'98. Berlin: Springer-Verlag, 1998: 221-235. S

[18] COSTER M J, JOUX A, LAMACCHIA B A, et al. Improved low-density subset sum algorithms[J]. Computational Complexity, 1992, 2(2): 111-128.

[19] HOROWITZ E, SAHNI S. Computing Partitions with Applications to the Knapsack Problem[M]. New York: Cornell University, 1972.

（上接第 142 页）

[18] GREENSMITH J, AICKELIN U, CAYZER S. Introducing Dendritic Cells as a Novel Immune-Inspired Algorithm for Anomaly Detection[M]// Artificial Immune Systems. Springer Berlin Heidelberg, 2005: 153-167.

[19] WANG Y Q, LIANG Y W, LIU S. Application-layer DDoS attack detection based on dendritic cell algorithm[J]. Computer Engineering and Design, 2015, 36(4): 841-845. (in Chinese)
王亚芹，梁意文，刘赛. 基于树突状细胞算法的应用层 DDoS 攻击检测[J]. 计算机工程与设计，2015，36(4)：841-845.

[20] GREENSMITH J, AICKELIN U. The Dendritic Cell Algorithm [J]. RevistaClínica Española, 2007, 202(10): 552-554.

[21] YI L L. Research on Statistical Characteristic Analysis and Modeling for User Behavior in Micro-blog Community Based on Human Dynamics[D]. Beijing: Beijing University of Posts and Telecommunications, 2012. (in Chinese)
易兰丽. 基于人类动力学的微博用户行为统计特征分析与建模研究[D]. 北京：北京邮电大学，2012.

[22] HE L, HE Y, HUO Y Q. Micro-blog user characteristics analysis and core user mining [J]. Intelligence Theory and Practice, 2011, 34(11): 121-125. (in Chinese)
何黎，何跃，霍叶青. 微博用户特征分析和核心用户挖掘[J]. 情报理论与实践，2011，34(11)：121-125.

[23] WANG X G. Empirical Analysison Behavior Characteristics and Relation Characteristics of Micro-blog Users — Take "SinaMicro-blog" for Example[J]. Library and Information Service, 2010, 54(14): 66-70. (in Chinese)
王晓光. 微博客用户行为特征与关系特征实证分析——以"新浪微博"为例[J]. 图书情报工作，2010，54(14)：66-70.

[24] Sina Weibo. Sunshine credit[EB/OL]. http://service.account.weibo.com/sunshine/guize.

[25] LIAN J, ZHOU X, CAO W, et al. SINA microblog data retrieval [J]. Journal of Tsinghua University, 2011, 51(10): 1300-1305. (in Chinese)
廉捷，周欣，曹伟，等. 新浪微博数据挖掘方案[J]. 清华大学学报(自然科学版)，2011，51(10)：1300-1305.

[26] YANG M, YIN J M, JI G L. Classification Methods on Imbalanced Data: a Survey[J]. Journal of Nanjing Normal University, 2008, 8(4): 7-12. (in Chinese)
杨明，尹军梅，吉根林. 不平衡数据分类方法综述[J]. 南京师范大学学报(工程技术版)，2008，8(4)：7-12.