DOI:10.14132/j.cnki.1673-5439.2012.04.018

基于图像感知哈希技术的钓鱼网页检测

周国强[1],田先桃[1],张卫丰[1,2],张迎周[1]

(1. 南京邮电大学 计算机学院,江苏 南京 210046
2. 苏州大学 江苏省计算机信息处理技术重点实验室,江苏 苏州 215006)

摘 要:文中提出一种基于图像感知哈希技术的钓鱼检测方法,简称Phash——将网页以图像格式保存,提取图像的主要可视信息的像素点,由这些像素点组成感知哈希序列,再进行哈希序列的相似度匹配。该方法既克服了钓鱼网页存活时间短的问题,又能快速地与特征库进行匹配。实验表明,该方法是有效的,在保证一定误判率和召回率的情况下大大提高了匹配速度。

关键词:图像哈希;钓鱼检测;网页相似性;分类器
中图分类号:TP393.08 **文献标识码**:A **文章编号**:1673-5439(2012)04-0059-05

Detecting Phishing Web Pages Based on Image Perceptual Hashing Technology

ZHOU Guo-qiang[1], TIAN Xian-tao[1], ZHANG Wei-feng[1,2], ZHANG Ying-zhou[1]

(1. College of Computer, Nanjing University of Posts and Telecommunication, Nanjing 210046, China
2. Jiangsu Key Laboratory for Computer Information Processing Technology, Soochow University, Suzhou 215006, China)

Abstract: This paper uses the image perceptual hashing technology, named Phash, to detect phishing web pages. The web pages are first converted into the images, and then the primary visual image pixels are extracted to form the perception hash sequence. After that, we make the hash sequence similarity match. This method not only overcomes the problem of phishing's short survival time, but also makes the image similarity match quickly. The experiments show that the algorithm improves the matching speed while ensures the false positive rate and recall rate.

Key words: image hashing; phishing detection; web page similarity; classifier

0 引 言

随着人们越来越多地依靠互联网工作、生活和娱乐,互联网诈骗已经成为一个越来越大的威胁。互联网诈骗有多种形式,其中网络钓鱼就是互联网诈骗中发展最快的一种。2007年,美国由于钓鱼袭击损失了30亿美元。2008年反钓鱼攻击工作组已经报道了363 662个不同的钓鱼网站。中国互联网络信息中心联合国家互联网应急中心发布的《2009年中国网民网络信息安全状况调查报告》显示,2009年有超过九成网民遇到过网络钓鱼,在遭遇过网络钓鱼事件的网民中,4 500万网民蒙受了经济损失,占网民总数11.9%,网络钓鱼给网民造成的损失已达76亿元。钓鱼网站欺骗行为正有愈演愈烈的趋势。

网络钓鱼方案已经"日趋完善",犯罪花样不断翻新,即便是知道"点击这里检查账号"电子邮件诈

收稿日期:2011-11-15;修回日期:2011-12-05
基金项目:国家自然科学基金(61103045,60973046)、桂林电子科技大学广西可信软件重点实验室开放基金(TJ211037)、江苏省青蓝工程、武汉大学软件工程国家重点实验室开放基金(BJ211002)和苏州大学江苏省计算机信息处理技术重点实验室基金(KJS0714)资助项目
通讯作者:张卫丰 电话:13776678880 E-mail:zhangwf@njupt.edu.cn

骗这种老旧办法的网络高手也有可能上钩。最典型的网络钓鱼攻击过程如下:首先将用户引诱到一个通过精心设计的,与目标组织的网站非常相似的钓鱼网站上,然后获取用户在该钓鱼网站上输入的个人敏感信息,例如银行帐号、银行密码等。通常这个攻击过程不会让受害者警觉。这些个人信息对钓鱼网站持有者具有非常大的吸引力,通过使用窃取到的个人信息,他们可以假冒受害者进行欺诈性金融交易,获得极大的经济利益,而受害者们却因此而遭受到巨大的经济损失,不仅如此,被窃取的个人信息还可能被用于其他非法活动。

提高钓鱼网页检测的性能,开发高效的检测技术是完全必要的,并且是迫切需要的。本文提出一种基于图像感知哈希技术的钓鱼网页检测,实验表明,该方法是有效的,在保证一定误判率和召回率的情况下大大提高了匹配速度。

1 相关工作

从自动化程度来分,钓鱼网站识别主要分为人工识别和计算机自动识别,人工识别[1]采用黑名单机制,用户对某个网站进行举报,通过人工鉴定是否为钓鱼网站,显然这样速度太慢。计算机自动识别目前主要是基于邮件特征的识别、基于第三方工具的识别[2]和基于相似性的判断。

Abu-Nimeh 从钓鱼网页传播的角度提出了一种特征提取方法[3]。该方法主要提取携带钓鱼网页的邮件特征,然后比较了6种机器学习方法在这些邮件特征分类上的效果。这种方法拓展了钓鱼网页的特征,在一定程度上进一步提高了钓鱼网页检测的精度,但在抽取钓鱼网页特征时仍然只是采用了单个网页的信息,因而容易被钓鱼网页制作者欺骗。

Zhang 在2007年提出 CANTINA[2],该方法通过借助第三方工具(比如搜索引擎)来检测钓鱼网页,它首先统计网页中的 TF-IDF,把 TF-IDF 排序靠前的几个词条利用搜索引擎检索,如果该网页不出现在搜索结果的前30个结果中,则认为是钓鱼网页。该方法具有较高的精度和较小的误判率,但是该方法的效果值得商榷。本文做了对应的实验,结果发现:很多钓鱼网页可以在搜索引擎中搜索到,并且结果比较靠前。这可能跟钓鱼网页制作者做了搜索引擎优化有关。另外,这种方法不具有实际的可行性,一方面,Google 搜索已经不提供用户通过 API 来访问其搜索服务,意味着这样的检测不能通过程序自动实现;另外,Google 对来自同一 IP 的每天的搜索次数进行了限制,意味着不能应付大量的钓鱼网页检测。由此看来,这种寄生于第三方服务的方法正失去其意义。

正如文献[4-6]中提到的,由于人们一般都比较注重自己浏览网页的主要目的,进而忽视了安全性问题的提示,并且视觉欺骗率很高。由此人们想到从视觉相似性角度来检测,基于视觉的检测分为基于 HTML 文本的检测、基于布局[6-7]的检测和基于图像[8]的检测。由于 HTML 语言的灵活性和网页元素的动态性及丰富性,仿冒者可以轻易地做出视觉上一样但是 HTML 结构不同的网页,这样,基于 HTML 的匹配将会失效。基于布局特征和图像特征的网页相似检测方法根据人的视觉原理,对网页的相似性进行计算,是一种通用的检测方法。如2006年 Fu 等人提出了一种基于像素的 EMD 距离的匹配算法[9],这种算法是在像素水平上从视觉的相似性角度来察觉钓鱼网页的。从实验结果可以看出:效果要明显好于基于 HTML 内容的检测,但也有其局限性,该算法只考虑了网页图像中的颜色及其分布特点,没有考虑网页中不同部分之间的位置关系。根据格斯塔视觉原理,相对位置在人的视觉中占主要地位,特别是多个形体间的相对位置关系,相对位置关系的变化必然导致视觉上的区别,而该算法由于没有考虑相对位置因素可能导致相似检测的失效。

由于钓鱼网页的存活时间很短,给直接连网提取网站特征进行检测带来了很大的不便,有些钓鱼网页,可能还没来得及对其进行特征提取,该网页就已经不存在了。本文提出的方法没有提取网页中的文字、图像特征,而是将网页整体以图像格式保存,然后提取图像的感知哈希序列,避免钓鱼网页存活时间短的问题,最后进行相似度匹配。实验表明:该方法是有效的,在保证一定误判率和召回率的情况下大大提高了匹配速度。

2 基于图像感知哈希的钓鱼网页检测

图像感知哈希技术将图像数据转换为几百或几千比特的二值序列,对于大量的图像数据库检索来说,极大地缩短了检索的时间,也降低了存储介质的成本。

现有的图像哈希生成方案基本按照如下框架

进行[10]:

(1) 对图像进行 DCT、小波变换等处理,提取部分 DCT 系数或小波系数,对提取的特征进行处理(如加密、视觉模糊);(2) 对第(1)步得到的哈希序列进行量化处理,考虑上一步的特征具有相当多的冗余,因此必须进行量化处理;(3) 对量化后的序列进行压缩编码处理;(4) 数字签名或图像索引都具有序列长度较短的需求,因此还要进行进一步的压缩处理。

图像感知哈希序列的提取方法很多,本文也将采用类似于传统的图像哈希生成方法的步骤进行,即特征提取、量化、压缩编码,生成流程如图 1 所示。将文中提到的符号的命名见表 1。

图 1　感知哈希序列生成流程图

表 1　命名表

M: 图片个数,实验中等于 100
$i,j=1,2,\cdots,M$
Ip(i): phishing 中第 i 张图片
Ir(i): reference 中第 i 张图片
Il(j): legal 中第 j 张图片
Sim_p: $1\times M$ 的矩阵,存放 phishing 中每张图片与特征库的相似度
Sim_r: $1\times M$ 的矩阵,存放 legal 中每张图片与特征库的相似度
C(·): 图片的大小
S(i,j): 图片 i 与图片 j 的相似度

图 1 中图像预处理部分就是对图像进行规格化,将图像统一变为具有 255 阶的灰度图像,并用双线性插值的方法将分辨率统一变为 $m\times m$(实验中 m 取 32),目的是使得最后生成的哈希序列长度统一。

图像感知提取部分首先对图像进行 DCT 变换,将 $m\times m$ 的图像分成 $n\times n$(实验中 n 取 8)的小块,对每一小块进行 DCT 变换,最后对于每一小块,保留 1 个 DC 系数,9 个 AC 系数,其余的将其置为 0。然后用 Watson 视觉模型对新生成的 DCT 系数矩阵进行处理,能很好地去掉信息中的冗余数据,提高了图像压缩的效率,极大地减少了传输的数据量和时间。

DCT 变换是一种实数域变换,其变换核为实数的余弦函数。利用 Fourier 变换的对称性,采用图像边界褶翻操作将图像变换为偶函数形式,然后对这样的图像进行二维离散 Fourier 变换,变换后的结果将仅包含余弦项,变换之后有关图像的主要的可视信息都集中在 DCT 变换的一部分系数中。因此,DCT 是有损图像压缩 JPEG 的核心,同时也是所谓"变换域信息隐藏算法"的主要"变换域(DCT 域)"之一[11]。另外,Watson 提出的基于 DCT 的视觉模型是一个经典的综合敏感度、掩蔽和误差合并的感知模型,本文中只用了敏感度。该模型用 DCT 取代了多通道分解,与 JPEG 和一些水印算法能很好的结合[11]。

加密处理部分是对矩阵进行标准化处理,处理后的矩阵更具有鲁棒性,根据混沌区数据的迭代不重复性和初值敏感性用 Logistic 方程作为混沌序列发生器进行加密[12],由一个密钥生成一个加密矩阵,用此矩阵对 DCT 系数矩阵进行加密,保证哈希函数的安全性。

Logistic 方程如下:

$$x_{n+1}=\mu x_n(1-x_n),\ 3.569\ 9<\mu<4 \qquad (1)$$

量化就是将浮点型数据变为二值数据,减少冗余,便于存储。最后用哈夫曼压缩编码进行压缩,得到最终的哈希序列,尽可能地减小哈希序列的长度,以便于应用。

得到哈希序列后用海明码距离计算公式进行图像匹配,设 h_1 和 h_2 为两个哈希序列,L 为哈希序列的长度,则:

$$\text{Distance}=\sum_{i=1}^{L}|h_1(i)-h_2(i)| \qquad (2)$$

根据图 1 所示的流程图提取每张图片的感知哈希序列,然后计算特征库中每一张图片与样本库中图片的相似度。由式(2)可以计算两个哈希序列的距离,取其距离最小的作为与特征库中图片的相似度,完成匹配,匹配流程如下:

输入: 特征库和样本库中的图片　输出: Sim_p Sim_r
伪代码:

```
1: 提取每张图片的感知哈希序列
2: 计算特征库与样本库中图片的相似度
/*  计算 phishing 中图片与特征库中的相似度 */
for i = 1: M /*  i 表示 phishing 中第 i 张图片的哈希序列 */
    for j = 1: M /*  j 表示 legal 中第 j 张图片的哈希序列 */
        d(i,j) = Distance(Ip(i),Il(j)); /*  海明距离 */
    end     Sim_p(i) = min(d(i,1),d(i,2),… d(i,M));
end
/*  计算 legal 中图片与特征库中的相似度 */
for i = 1: M /*  i 表示 reference 中第 i 张图片的哈希序列 */
    for j = 1: M /*  j 表示 legal 中第 j 张图片的哈希序列 */
        d(i,j) = Distance(Ir(i),Il(j));
    end
    Sim_r(i) = min(d(i,1),d(i,2),… d(i,M));
end
```

3 实 验

3.1 实验准备

由于大多数钓鱼网页的存活时间很短,所以没有直接从网上采集钓鱼网页,而是从 http://www.phishtank.com(一个免费的反钓鱼网站,用户可以提交可疑的钓鱼网页)上采集样本。本文采集其中的部分网页,通过手工检查,去除合法网页,并将其以图像格式保存,然后找到与钓鱼网页对应的合法网页。共收集了 100 张钓鱼网页图片,命名为 phishing;100 张与钓鱼网页对应的合法网页图片,命名为 legal;100 张其它的正常网页图片,命名为 reference。将 legal 图片作为特征库,与之对应的 phishing 图片和 reference 作为样本库。

通过阅读大量国内外文献以及经过前期的实验研究,获取了多种形式的评价指标,包括:TPR(召回率 Recall)、FPR、精确率(Precision)、F-measure、AUC 值以及准确率。前 3 个比较重要,反钓鱼检测比较看重召回率,认为召回率越高越好,而用户看重的是 FPR,认为 FPR 越低越好。FPR 表示把真实的正常网页错误地预测为钓鱼网页的概率;Recall,即 TPR,表示把真实的钓鱼网页正确地预测为钓鱼网页的概率,描述了总的钓鱼网页中被检测出的比例,而精确率 Precision 描述了所有被预测为钓鱼的网页中真正是钓鱼网页的比例。计算式如下:

$$\mathrm{TPR(Recall)}=\frac{|\mathrm{TP}|}{|\mathrm{TP}|+|\mathrm{FN}|} \tag{3}$$

$$\mathrm{FPR}=\frac{|\mathrm{FP}|}{|\mathrm{FP}|+|\mathrm{TN}|} \tag{4}$$

$$\mathrm{Precision}=\frac{|\mathrm{TP}|}{|\mathrm{TP}|+|\mathrm{FP}|} \tag{5}$$

其中,$|\mathrm{TP}|$ 表示把真实的钓鱼网页正确地预测为钓鱼网页的数量,$|\mathrm{FP}|$ 表示把真实的正常网页错误地预测为钓鱼网页的数量,$|\mathrm{TN}|$ 表示把真实的正常网页正确地预测为正常网页的数量,$|\mathrm{FN}|$ 表示把真实的钓鱼网页错误地预测为正常网页的数量。

实验代码分别采用 matlab7.0.1 和 visual c++6.0 编写,笔记本硬件配置为 Intel 奔腾 1.7 Ghz 处理器;1.25 G 内存;Windows XP Professional SP2。

3.2 实验过程

Phash 的主要思想是提取图像的主要可视信息的像素点,减少冗余,然后进行相似度的匹配。与基于图像全部像素点的方法比较,简称 Baseline,该方法是将两张图片的对应像素点进行相减运算,然后由式(6)进行相似度的计算:

$$S(i,j)=\frac{C(|I_i-I_j|)}{C(I_0)} \tag{6}$$

其中,I_i 表示样本库 phishing 或 reference 中的图片,I_j 表示特征库 legal 中的图片,I_0 表示灰度值全为 0 的 32×32 的灰度图片,$C(\cdot)$ 表示图片的大小,$C(|I_i-I_j|)$ 表示两个图片对应像素点的灰度值相减后的图片大小。

下面简单介绍一下实验过程:首先对特征库和样本库中的每一张图片提取感知哈希序列,具体实施已经在第 2 节中介绍;其次计算 phishing 中的每一张图片与特征库中所有图片的海明距离,取其距离最小的作为与特征库中图片的相似度 Sim_p,同样计算 reference 中每一张图片与特征库中的相似度 Sim_r,匹配流程见第 2 节;最后采用机器学习方法进行训练和预测。Baseline 方法根据式(6)计算相似度,实验过程同上。本文用 Simple Logistic、Support Vector Machines(支持向量机)和 Random Tree 3 种分类器对这两种方法的实验结果进行分类,结果见表 2。

表 2 不同分类器下两种实验方法的评价指标比较

(a) Simple Logistic 分类器分类效果

分类方法	评价指标				
	TPR(Recall)	FPR	Precision	F-measure	AUC
Baseline	0.556	0.444	0.564	0.541	0.599
Phash	0.894	0.106	0.896	0.894	0.895

(b) Support Vector Machines 分类器分类效果

分类方法	评价指标				
	TPR(Recall)	FPR	Precision	F-measure	AUC
Baseline	0.756	0.244	0.756	0.756	0.756
Phash	0.894	0.106	0.896	0.894	0.894

(c) Random Tree 分类器分类效果

分类方法	评价指标				
	TPR(Recall)	FPR	Precision	F-measure	AUC
Baseline	0.633	0.367	0.788	0.576	0.727
Phash	0.883	0.117	0.884	0.883	0.927

3.3 实验结果分析

如何快速有效地计算网页的相似度是检测钓鱼网页的关键。由表 2 可以很明显看出:在不同分类器下,Phash 的各项评价指标相差不大,并且 Phash 的各项评价指标都要比 Baseline 的好,这说明 Phash

方法进行钓鱼检测是有效的。另外,Phash 考虑了图像的主要可视信息的像素点,而 Baseline 方法考虑了图像中所有的像素点,时间复杂度很高,不利于在一个大量的特征库中查找最相似的图片。实验中,Baseline 方法要对每两张图片的对应像素点进行一次相减运算,即 100 张phishing图片中的每一张都要跟 100 张 legal 中的图片计算一次对应像素点相减运算,那么就要计算10 000次,同样 reference 中图片也要计算 10 000 次,该方法的总运行时间大约为 7 561.108 s,见表 3。而 Phash 的运算时间却很短,因为考虑的是每张图片的主要可视信息的像素点,其核心算法总运行时间只有 22.593 s,见表 4。

表 3　Baseline 方法运行时间

实验核心步骤		所需时间/s
特征库与样本库两两图片对应像素点做相减运算	phishing 中的图片与特征库中图片做差值	3 660.344
	reference 中的图片与特征库中图片做差值	3 900.530
计算样本库中图片与特征库中的相似度		0.234
总计		7 561.108

表 4　Phash 方法运行时间

实验核心步骤		所需时间/s
提取图像的哈希序列	提取 phishing 中图像哈希序列	7.140
	提取 reference 中图像哈希序列	7.125
	提取 legal 中图像哈希序列	7.140
计算样本库中图片与特征库中的相似度		1.188
总计		22.593

Baseline 方法运行所需要的时间很长,只适用于小数据的样本。而 Phash 可以快速匹配,适用于大数据样本库,符合实际应用的要求。考虑到快速与有效,Phash 方法更适用于实际的反钓鱼检测工作。

4　结束语

本文主要提出一种基于图像感知哈希技术的钓鱼检测方法,该方法提取图像的主要可视信息的像素点,大大提高了在特征库中匹配的速度,既避免了钓鱼网页存活时间短的问题,又能快速地与特征库进行匹配。相比而言,Baseline 方法的匹配时间很长,只适用于小样本的分类模型,不能被实际中大样本数据所应用。未来的工作将会继续优化图像感知哈希技术的方法以提高其召回率降低误判率。同时,也有了新的启发,考虑先对数据集进行聚类,然后用 Baseline 方法来进行钓鱼检测,改善其实用性。

参考文献:

[1] JACKSON C,SIMON D R,TAN D S,et al. An evaluation of extended validation and picture-in-picture phishing attacks [C] // Proceedings of Usable Security(USEC'07) . 2007.

[2] ZHANG Y,HONG J,CRANOR L. CANTINA: A content-based approach to detecting phishing web sites [C] // Proceedings of the 16th International Conference on World Wide Web. 2007: 639 - 648.

[3] ABU-NIMEH S,NAPPA D,WANG X,et al. A comparison of machine learning techniques for phishing detection [C] // Proceedings of the e-Crime Researchers Summit. 2007.

[4] DHAMIJA R,TYGAR J D,HEARST M. Why phishing works [C] // Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. NewYork: ACM Press,2006.

[5] DONG X,CLARK J A,JACOB J L. Defending the weakest link: Phishing websites detection by analysing user behaviours [J] . Telecommunication Systems,2010,45: 215 - 226.

[6] AFROZ S,GREENSTADT R. Phishzoo: An automated web phishing detection approach based on profiling and fuzzy matching [R] . Philadelphia: Drexel University,2009.

[7] LIU W,HUANG G,LIU X. Detection of phishing web-pages based on visual similarity [C] // International World Wide Web Conference. 2005: 1060 - 1061.

[8] CHEN K T,CHEN J Y,HUANG C R,et al. Fighting phishing with discriminative keypoint features of webpages [C] // IEEE Internet Computing. 2009.

[9] FU A Y,LIU W Y,DENG X T. Detecting phishing web pages with visual similarity assessment based on earth mover's distance(EMD) [J] . IEEE Transactions on Dependable and Secure Computing, 2006,3(4) : 301 - 311.

[10] SWAMINATHAN A,MAO Y N,WU M. Robust and secure image hashing [J] . IEEE Transactions on Information Forensics and Security,2006,1(2) : 215 - 230.

[11] 王丽娜,郭迟,李鹏. 信息隐藏技术 [M] . 武汉: 武汉大学出版社,2004: 10.
WANG Li'na,GUO Chi,LI Peng. Experiments of information hiding technology [M] . Wuhan: Wuhan University Press,2004: 10. (in Chinese)

[12] 刘晓义,王述洋. 一种基于混沌和魔方的数字图像置乱算法 [J] . 中国安全科学学报,2008,18(7) : 111 - 115.
LIU Xiaoyi,WANG Shuyang. An algorithm for digital image scrambling based on chaos and rubik's cube [J] . Journal of China Safety Science,2008,18(7) : 111 - 115. (in Chinese)

(下转第 69 页)

4　结束语

本文设计并实现一个利用可编程仪器组建的天线测试系统,建立计算机与卫星地球站天线套接字通信,以及与可编程仪器的 GPIB 通信,同时利用 MATLAB 强大的数值计算和图形显示能力,以及易用的编程开发环境,采用 C++与 MATLAB 的混合编程,将复杂的数据分析与图形绘制过程交给 MATLAB 来处理。通过实际应用该系统进行实际测试,结果表明该系统确实可以使测试过程更加便捷,测试结果更加准确,对于评价一副卫星天线性能更加可靠。完善计算机与天线之间,与可编程仪器之间的基于多种接口的通信方式将是下一步研究的重点。

参考文献:

[1] WANG Pinglian, YUAN Sumin. Automated measurement system for wireless transmitters [C] // Proc of the 8th International Conference on Electronic Measurement and Instruments. 2007, 1: 975 - 977.

[2] FRED B. Introduction and update to an open standard for instrument control: SCPI standard commands for programmable instruments [C] // Wescon Conference Record. 1992, 11: 600 - 604.

[3] LI Jianmin, ZHENG Bin, HOU Wen. Design of supervision software for programmable power supply based on virtual instrument development environment [C] // World Congress on Computer Science and Information Engineering. 2009, 65: 301 - 305.

[4] BRUECKMANN H. Antenna pattern measurement by satellite [J]. lEEE Transactions on Antennas and Propagation, 1963, 3: 143 - 147.

[5] Agilent Technologies. User's/Programmer's Reference [M]. California: Agilent, 2006.

[6] JIANG Ronghua, WU Xibei. Design and realization of the software of GPIB controller based on USB [J]. China Measurement Technology, 2006, 7, 32(4): 93 - 96.

[7] GUO Zhanshan, WANG Zenghao, MU Neng. The standard of GPIB of instruments and the form of test system [J]. Electronic Instrumentation Customer, 2002, 1: 39 - 41.

[8] JI Xianfa, WU Yifeng. Design of program controlling GPIB instrument based on VC + + 6.0 [J]. Metrology and Measurement Technique, 2004, 7: 26 - 27.

[9] DONG Weiguo. MATLAB 7. x hybrid programming for the Layman [M]. Beijing: China Machine Press, 2005.

[10] LIANG Jun. Communication systems and measurement [M]. Xi'an: Xidian University Press, 2005.

作者简介:

陈佳滨(1986 -),男,广东汕头人。南京邮电大学通信与信息工程学院硕士研究生。主要研究方向为卫星通信系统。

(上接第 63 页)

作者简介:

周国强(1968 -),男,湖北黄石人。南京邮电大学计算机学院副教授,博士。主要研究方向为信息安全、可信计算和服务安全。

田仙桃(1985 -),女,山西文水人。南京邮电大学计算机学院硕士研究生。主要研究方向为 Spam 检测技术、搜索引擎技术和移动电子商务技术。

张卫丰(1975 -),男,江苏南通人。南京邮电大学计算机学院教授,博士。主要研究方向为 Spam 检测技术、移动社交网络、搜索引擎技术和移动电子商务技术。

张迎周(1978 -),男,安徽巢湖人。南京邮电大学计算机学院副教授,博士。主要研究方向为网络应用软件的安全可靠性分析与研究、基于信息流分析的 Web 应用诊断和检测技术和基于数据流分析的软件安全技术。