doi:10.3969/j.issn.1000-436x.2012.12.002

基于社会信任的恶意网页协防机制

刘昕[1,3]，贾春福[2,3]，刘国友[2]，胡志超[2]，王冬[2]

（1. 中国石油大学（华东） 计算机与通信工程学院，山东 青岛 266580；
2. 南开大学 信息技术科学学院，天津 300071；3. 南京大学 计算机软件新技术国家重点实验室，江苏 南京 210093）

摘 要：针对恶意网页的威胁，提出了一种基于社会信任的分布式恶意网页协作防御机制：结合第三方专业服务机构提供的恶意网址列表，并利用社会网络中好友间的直接信任和间接信任获取好友对网页的评价信息，集成好友的安全浏览经验形成网页综合评价；每个用户都与其好友进行协作，形成一个网状的防御体系。实验结果表明，该机制能够有效减少恶意网页的访问量，提高社会网络防御恶意网页的能力。

关键词：社会网络；社会信任；恶意网页；网页评价；间接信任

中图分类号：TP393.08 **文献标识码：**A **文章编号：**1000-436X(2012)12-0011-08

Collaborative defending scheme against malicious Web pages based on social trust

LIU Xin[1,3], JIA Chun-fu[2,3], LIU Guo-you[2], HU Zhi-chao[2], WANG Dong[2]

(1. College of Computer & Communication Engineering, China Univers ty of Petroleum, Qingdao 266580, China;
2. Department of Computer and Information Security, Nankai University, Tianjin 300071, China;
3. State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China)

Abstract: To deal with the threat of malicious Web pages, a distributed defending scheme against malicious Web pages based on social trust was proposed. Besides the malicious URL list from third-party professional organizations, the direct and indirect trust relations between friends in social network were used to obtain evaluations of Web pages. The experiences about Web surfing from a user's friends were collected to result in synthetical evaluations on his computer. Each user cooperated with his friends, so that a defending system was formed on the overall perspective, which can improve the defending ability of the social network against malicious Web pages. The experiment results indicate that the visits of malicious Web pages under the scheme decrease obviously than the methods without social trust.

Key words: social network; social trust; malicious Web page; Web page evaluation; indirect trust

1 引言

随着 Web 的广泛应用，包含恶意代码和网上诱骗的网站也日益泛滥，对 Web 的安全应用构成了极大的威胁。网页中的恶意代码是指在用户不知情的情况下自动安装到用户计算机中的代码，其目的是危害用户系统和数据，或从其中窃取信息。网上诱骗是指冒充他人骗取用户个人信息或其他敏感信

收稿日期：2011-08-24；修回日期：2012-02-29

基金项目：国家自然科学基金资助项目（60973141, 61272423）；国家重点基础研究发展计划（“973”计划）基金资助项目（2013CB834204）；高等学校博士学科点专项科研基金资助项目（20100031110030）

Foundation Items: The National Natural Science Foundation of China (60973141, 61272423); The National Basic Research Program of China (973 Program)(2013CB834204); The Specialized Research Fund for the Doctoral Program of Higher Education of China (20100031110030)

息的行为，通常采用虚假网站的方式。钓鱼攻击(phishing attack)是一种常用的网络欺骗手段，主要通过电子邮件、网页等途径散布虚假信息，诱骗不知情的网络用户登录仿冒的网站，如假冒的网上银行、在线购物等，从而骗取用户的网银账号、信用卡号和密码等敏感信息。据国际反钓鱼工作组(APWG)统计，2010 年上半年全球钓鱼攻击为 48 244 次，下半年则达 67 677 次，增长的主要原因是中国钓鱼攻击的增多[1]。

为了保护用户的计算机不受恶意网站攻击，当前各种浏览器普遍采用的解决方案是附加安全功能，在用户浏览恶意网站之前，向用户发出警告信息，比如谷歌浏览器的安全浏览[2]和 IE 浏览器的 SmartScreen 筛选器[3]等，用以帮助用户防御网上诱骗和恶意代码的攻击。这些浏览器安全功能通常通过一种类似黑名单的恶意网址列表机制来实现，将给定的URL与恶意网址列表进行匹配，判断该URL是否为恶意网站。

恶意网址列表通过自动检测和人工举报结合的方式获得，其原理简单且易于被浏览器和其他应用程序实现。其不足是第三方专业服务机构为用户提供的恶意网址列表的更新速度远远跟不上恶意网页的更新速度。同时，随着网络犯罪技术的不断提高，发现恶意网页的难度也在加大。若要及时发现新出现的恶意网站，必须尽可能应用多方面的力量，而不能仅仅依赖第三方专业机构。实际上，能够最早发现新出现恶意网页的正是广大网络用户自己。

通常，用户访问恶意欺诈网页或者被恶意代码攻击之后，会获得关于防御这些恶意网页和恶意代码的经验。一用户访问某个网站之前，若其好友恰好浏览过该网站，就可以根据好友提供的安全浏览经验，来决定是否访问该网站，从而避免登录恶意网站，减少损失。这种方式实际上是借助多个用户的经验来处理同一问题。如果能够将网络用户的经验充分利用起来，将非常有利于对恶意网页的防御。

用户只愿意向其信任的好友提供其浏览网页的经验，也只愿意接收来自其信任的好友推荐的信息。许多用户都已加入到各种不同的社会交互应用，如 Facebook、QQ、MSN、校内网、人人网和新浪微博等。截至 2011 年 6 月，Facebook 的用户数量已经突破 7.5 亿[4]。这些在线社会网络将人们之间的信任关系从现实生活中转移到网络中。

信任是所有社会交互的基础。随着 Web 的演化，社会信任关系的使用是实用的而且是必须的[5]。社会网络应用越来越广泛，成为一个良好的通信平台，在信任的基础上用户可以与好友共享资源、分享经验。

本文提出了一种基于社会信任的分布式网页安全评价方案。除应用第三方提供的恶意网址列表之外，还利用用户的经验和好友间的信任，获得更为可信的网页评价结果，及时避免用户访问恶意网页。从一个用户角度看，以用户个体为中心，与好友之间形成协作防御；从整个社会网络角度看，基于社会网络平台，每个用户都可以与其好友协作，充分利用整个网络上的资源、能力和经验，从而在宏观上形成一个网状的防御体系。

2 相关研究

社会信任在很多方面都有应用。比如，文献[6]将信任应用到推荐系统以提高推荐的正确性；文献[7,8]将社会信任用于垃圾邮件过滤；文献[9]将社会网络中用户之间的信任用于检测 Sybil 攻击等。我们希望将社会信任引入用户主机安全防御系统，以减缓恶意代码传播。

Sirivianos[7]提出了一个基于信任的协作式垃圾邮件过滤系统 SocialFilter。每一个 SocialFilter 的节点都发送报告到一个中央数据库，中央数据库形成信任图进行信任推导，并计算一个节点到其他所有节点的最大信任路径。SocialFilter 利用社会网络中的信任关系来评价对发送报告者的信任程度，其目的是聚集多种垃圾邮件检测器的经验，使每一个不具备邮件分类功能的用户主机通过查询，就可以确定一个主机是否为垃圾邮件制造者。

评级、信誉和推荐系统已经成为很多 Web 在线系统的核心组成部分，这些系统通过综合用户对产品和服务的评论，从而形成有价值的参考信息。比较成功的商业案例包括 Amazon 和 eBay 的推荐系统[10]和 Google 网页评级系统[11]等。

有些学者研究识别恶意网页的安全机制。如，Jean Camp 提出了一个 NetTrust 系统[12]，它结合用户的浏览习惯和朋友的意见来识别恶意网页，将好友的浏览行为信息放到工具栏上，供用户参考。但该系统缺乏社会化网络中信任传递的应用，只利用直接好友的意见，没有对评价信息进行综合，所得到的网页评价信息有限。

WOT 是一个基于社区的、免费的浏览器插件[13]。当用户在网上浏览和购物时，WOT 可以保护用户的安全。WOT 用颜色标记网页上的链接来提醒用户：红色代表可疑网站，绿色代表健康网站。每个网络用户都可以注册成为 WOT 用户，并对网站进行评价。WOT 将用户的所有评价综合起来形成对网站的评价值，但是该插件没有应用社会信任，用户不能根据好友的建议访问网站。

很多研究者提出了社会网络中信任传递的计算方法[14,15]。其中文献[15]提出的方法非常简单，是一个轻量级算法。本文将该方法应用于分布式防御系统中计算信任的传递，因而不会明显增加用户主机运行开销。

3　恶意网页协作防御系统

3.1　系统简介

在网页协作防御系统中，网页评价分布式地内嵌于每个用户的浏览器中，用户可以对访问的网页进行评价。依据“本地存储、定期交换、及时通知”的原则存储用户网页评价，获取好友评价信息，警告好友防范恶意网页。每个用户主机设置一个评价表，用于存储用户及其好友对所浏览网页的安全程度的评价；用户通过社会网络平台与好友定期交换网页评价信息；发现恶意程度高的网页，用户应及时通知其好友。

社会网络采用 MSN 用户构成的即时通信网络，通过信任传递的计算获得对间接好友的间接信任程度，即间接信任值。根据用户直接好友和间接好友的信任值及其对网页的评价值，计算用户评价表中每个网页的综合评价值。由于每个用户信任的好友以及信任的程度都是不同的，因此针对同一个网页，在不同用户主机上，计算所得的网页综合评价值通常也不相同。

用户访问一个先前没有浏览过的网页时，浏览器插件会过滤网页：首先检查第三方提供的恶意网址列表，然后检查用户评价表，根据计算得到的综合评价值所处的区间，给用户不同的提示。若继续访问该页面，用户可以评价该页面的安全级别，给网页打分。如果用户评价表中没有该网址，直接显示该网页，用户即可打分。

若用户访问一个恶意程度很高的网页，对其评分很低，则需要立即发送其评价给好友，及时提醒好友避免打开该网页。好友可以继续向其好友传递该评价，从而降低恶意网页的访问量，减少恶意网页带来的危害。

3.2　社会信任及其传递

通过对信任及其传递进行计算，获得达到一定信任程度的好友，根据这些好友的信任值计算网页综合评价值。

3.2.1　信任的概念

根据文献[16]对信任的定义以及本文的应用环境，对社会信任定义如下。

定义 1　社会信任是一种建立在已有知识上的主观判断，是用户 A 根据所处环境，对用户 B 能够按照用户 A 的意愿提供特定服务的度量，以下简称信任。

定义 2　直接信任(direct trust)是用户 A 根据与用户 B 的直接联系历史记录而得出的对用户 B 的信任。

定义 3　间接信任(indirect trust)是用户间根据第三个用户的推荐形成的信任，也称推荐信任。

如图 1 所示，将 3 个用户的信任关系进行了抽象，其中，T_{AB} 表示用户 A 对用户 B 的信任值。用户 A 与 C 之间不存在直接关系，用户 A 为获得对用户 C 的间接信任值，需要求助于用户 B。根据用户 B 的推荐，利用信任传递公式计算，可以得到 A 对 C 的间接信任值。

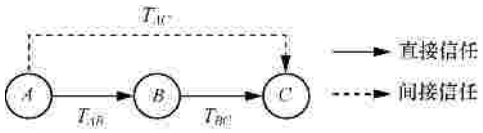

图 1　用户信任关系传递

3.2.2　信任传递机制

间接信任通过信任传递机制来计算。

本文使用 Golbeck[15]提出的信任计算模型，其中，2 个用户之间的间接信任值计算公式如式(1)所示。

$$T_{is}=\frac{\sum_{j\in N(i)}\begin{bmatrix}(T_{js}\times T_{ij}) & , & T_{ij}\geq T_{js}\\ (T_{ij}^{2}) & , & T_{ij}<T_{js}\end{bmatrix}}{\sum_{j\in N(i)}T_{ij}} \tag{1}$$

其中，T_{ij} 表示用户 i 对用户 j 的信任值，$N(i)$表示 i 的所有好友集合，包括直接好友和间接好友。

当用户计算对间接好友的信任值时，需要获得好友对间接好友的信任。通过与好友互换信息的方式可以获得这些信任值，然后进行信任传递计算。对于信任值很低的好友，用户不愿意与其交换信

息。为了忽略这些好友的评价，设置一个可信度阈值 $T_{\min}$。当计算得到的间接信任值大于 $T_{\min}$ 时，则记录该好友信息，同时与该好友交换网页评价信息；否则不记录该用户及其评价信息。

3.3 个人评价

用户访问一个网页时，可以对该网页进行评价，称为直接评价。但要求用户评价所访问的每一个网页比较困难，为了补充评价信息，根据用户的浏览行为增加一种评价方式：行为评价。

3.3.1 直接评价

WOT 将网页分为 5 个级别：极差、较差、令人不满意、较好、极佳，这种分级方式比较粗略。NetTrust 允许用户在区间[1, 5]和[-1, -5]指定一个值为网页评级。为了精确地对网页分级且便于计算网页综合评价值，-1、-2、-3、-4、-5 分别表示网页的危险程度：有点低、较低、低、很低、严重。相应地，正值代表网站的健康程度。用户可以指定一个[-5, 5]之间的数值来评价一个网页的安全程度。负值代表用户认为该网页存在危险性，正值代表用户认为该网页为健康网页。用户根据自己的浏览经验为网页打分，将其评价结果存放到本地网页评价表。最终网页上显示给用户的综合评价值是[-5, 5]区间内的一个值，可以使用户获得更精确的网页安全信息。

3.3.2 行为评价

用户只浏览网页而不对网页的安全性进行主动评价时，可以利用用户的浏览行为对网页形成评价。当一个用户多次重复访问一个网站时，说明该用户已经足够信任该网站，默认该网站为健康网站。NetTrust 系统记录用户对每个网站的访问频率，以此作为参数获得对该网站的评价[12]。

行为评价使用用户访问一个网站的次数来评价网站的健康程度。对于一个给定的网页，行为评价用一个[0, 5]区间内的整数值来评价网页的健康程度，与直接评价的正数区间保持一致。根据用户的浏览行为和综合评价值的计算需要，为行为评价选取 4 个影响因素：最大约束值、开始计数延迟、最小访问间隔和计数衰减期限，用以限制计数的最大值和有效性，可以较精确地间接评价网页的健康程度。

1) 最大约束值

用来设置行为评价值的上限。最大约束值将行为评价值限定在 5 以内，计数达到 5 之后就不再增加，使得行为评价值和直接评价值在数值上统一，便于综合评价值的计算。

2) 开始计数延迟

当用户第一次访问一个网站时，不能立即将这个网站的行为评价值记为 1，要经历一段时间即开始计数延迟之后才会将该网站的评价值设为 1。开始计数延迟设置为 5 天。根据 APWG 的统计[17]，钓鱼网站的平均存活时间通常不超过 5 天，设置开始计数延迟能够避免对钓鱼网站的错误计数。

3) 最小访问间隔

行为评价随着用户的访问次数不断变化。对网页的重复访问必须达到一定的时间间隔，访问次数才具有一定的意义。如果不设置该参数，用户可以通过不断刷新某一个页面来达到提升该页面行为评价值的目的。若用户长期重复访问某个网站，每天访问一次该网站是多数用户的习惯，因此将最小访问间隔设置为一天。

4) 计数衰减期限

如果用户在一段很长时间内（即衰减期限）都没再访问某个网页，该网页的行为评价值缩减到原来的一半。计数衰减期限设置为 15 天。若用户很久没有访问一个网页，之后重新开始访问该网页，则根据前面的计数规则继续对其进行行为评价。

对于同一个网页，如果用户给出了直接评价，则不再记录行为评价。除了对一个网页进行数值评价外，用户还可以进行文字描述。这些文字描述对网页的数值评价没有影响，但当用户访问该网页时，相关描述会被显示，用户可以选择查看。

3.4 基于信任的网页综合评价

每个用户主机上设置一个网页评价表，如表 1 所示。用于存储用户及其好友（包括直接好友与间接好友）对访问过网页的安全程度评价，以及用户对其好友的信任值。用户 ID 使用用户的 MSN 账号，是全局唯一的。该表同时也保存本地用户做出的评价，并以 100%作为其信任值。表 1 中最后一行是综合评价值，在综合其他用户对网页评价和其信任值的基础上进行计算得到。这些评价具有个体性，每个用户根据自己的网页评价表计算网页综合评价值。任何 2 个用户的网页评价表是不同的。

好友之间交换网页评价信息可以帮助用户利用好友的经验丰富自己的网页评价表。经验丰富和值得信赖的好友信任值会比较高，会提供更有价值的信息。

式(2)给出了计算一个网页的综合评价值的方

法，其中 R_u 表示对网址 u 的综合评价值，r_{ju} 表示好友 j 对网址 u 的评价值。

$$R_u = \frac{\sum_{j \in N(i)} \left(T_{ij} \times r_{ju}\right)}{\sum_{r_{ju} \neq 0,\, j \in N(i)} T_{ij}} \tag{2}$$

它综合所有访问过该网页且达到可信度阈值的好友给出的评价值，以对好友的信任值作为权重，进行加权评价，计算出的综合评价值如表 1 所示。一个信任程度高的好友对网页的评价值在综合评价值中的贡献会多一些。

表 1　　网页评价

用户 ID	信任值	URL1	URL2	URL3	URL4
ID1	100%	3			
ID2	90%	2	- 2	3	4
ID3	70%	1	- 2	3	
ID4	70%	2			3
ID5	90%		- 2	2	4
ID6	80%		- 1		
综合评价值		2.09	- 1.76	2.76	3.72

当用户访问一个网页时，根据综合评价值给用户不同的提示。若综合评价值不大于- 4，直接终止网页的显示，给出严重警告并显示给出评价值的好友及其信任值，向用户提示该网站可能会损害用户计算机；若综合评价值在(- 4,0)范围内，显示网页之前给用户一个提示，显示出综合评价值，提示用户是否继续显示；如果综合评价值是正值，直接显示网页，不给用户提示，同时仍然显示综合评价值；若用户评价表中没有该网页，则直接显示。若只有一个好友访问过某个恶意网页，对该网页评价值不大于- 4 时，他会通知其所有好友包括该用户。若该用户所有其他好友都没有浏览过该网页，即其他好友对该网页的评价值都为 0。最后计算得到的综合评价值与好友的评价值相同。若该好友是一个恶意用户，用虚假的网页评价阻止其好友访问该网页，多个用户会受其影响不能访问健康网页。为了避免这种情况，将评价该网页的好友显示给用户，由用户决定是否继续访问该网页。

对于用户以前浏览过的网页，根据网页评价表中用户自己的评价进行过滤。

3.5　好友之间的协作方式

本系统通过使用 MSN 提供的开放接口来应用社会网络。MSN 是微软公司推出的一款即时通信软件，可以与亲人、朋友和同事等进行即时交流。现实生活中的朋友关系在 MSN 中得到很好的体现，出现在 MSN 好友列表中的用户大多数是用户生活中真实的同学、朋友和亲人。

用户可以通过 MSN 账号登录我们的系统，系统自动记录用户账号信息，并将 MSN 的好友信息集成到系统中，实现好友之间的协作。

系统提供接口允许用户对自己的好友设定一个信任值，用百分数来描述。如用户非常信任某个好友，可以用 90%以上的数值表示。这些信任值用于信任传递和网页综合评价值的计算。

每个用户对网页的评价和对好友的信任值都保存在本地。为了获得好友对网页的评价值，用户与其直接好友会定期交换彼此的网页评价和对好友的信任值信息。交换网页评价信息相当于用户向好友咨询的过程。因为任何浏览行为都无法容忍过长的等待时间，用户不能在访问网页的时候才去向好友咨询，所以，只有获取尽可能多可信任用户的网页评价值存储到本地，才能在访问网页前及时获得准确的网页综合评价信息，对用户访问网页提供建议。

图 2 是一个信息交换的例子。用户 A 与 B_1、B_2 互相交换信息后，A 就获得了 B_1 和 B_2 对 C 的信任值，通过式(1)计算得到 A 对 C 的间接信任值。B_1、B_2 和 C 以及 C 和 D 也会互相交换信息，所以 B_1 和 B_2 也可以计算对 D 的信任值。从而在下一次交换信息时 A 能够计算对 D 的间接信任值。用户通过与好友之间信息的不断交换，会获得尽可能多的可信任用户的网页评价信息。

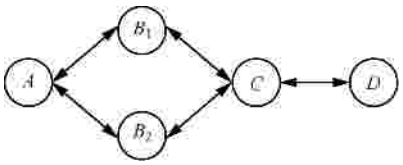

图 2　信息交换举例

为了及时向好友发送恶意网页的警告信息，除了定期交换信息外，系统还将被用户评价为高危网页的网页评价信息（如用户把某个网站评价为- 4 以下的值）迅速分发给自己的好友。收到这个评价信息的好友按同样方法处理。这使得恶意

网页的评价信息能及时发送出去，避免更多的用户访问恶意网页。

3.6 系统实现

图3给出了本文设计的系统结构。虚线框表示防御系统在用户主机上实现的功能。浏览器工具栏模块用于显示网页评价信息和用户进行网页评价的接口，评价信息主要包括好友对该网页的综合评价值以及第三方资源评价。评价引擎主要功能是计算用户对其好友的间接信任值和网页综合评价值。社会网络模块主要用于存储用户的好友信息，以及通过MSN接口与其他好友进行信息交换功能。同步模块主要功能是获取和存储网页评价信息。

用户启动装有该插件的浏览器后，用其MSN账号登录系统。系统启动以后，用户可以通过浏览器工具栏查看网页综合评价值以及用户对网页的相关文字描述，或者评价正在访问的网页。用户评价信息通过同步模块完成存储。在用户浏览网页的同时，系统会定期通过同步模块和社会网络模块与其他用户交换评价信息。

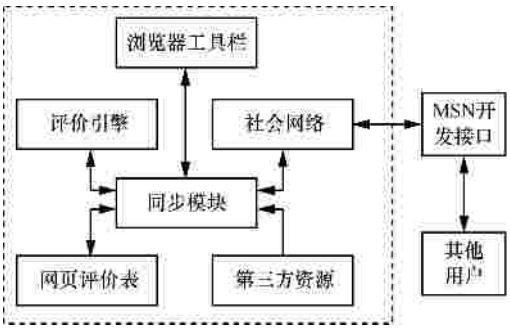

图3 系统结构

第三方专业机构提供的恶意网址列表比较全面，好友之间的经验共享比较及时，两者相互补充能够更加有效地降低恶意网页的访问量。本系统使用Google安全浏览API[18]，可以根据Google不断更新的可疑仿冒网页和恶意软件网页列表来检查网址。恶意网址列表中出现的网页或者好友评价的恶意网页都要被过滤，减少对用户的威胁。

4 Sybil攻击防御

Sybil攻击是指一个恶意节点可以使用多个ID伪装成网络系统中多个不同节点，从而对评价结果产生影响。基于社会信任的网页防御机制一个重要的安全特性就是可以有效抵御Sybil攻击。系统通过MSN开放接口来接入社会网络。若一个恶意用户伪造多个MSN账号进行Sybil攻击，由于缺乏信任，伪造的账号难以与网络中其他用户建立好友关系，因此无法获得其他用户比较高的信任。系统不接收没有达到可信度阈值的用户信息，评价表中不存放这些用户对网页的评价值，使得伪造账号对网页的评价无法影响网页综合评价值的计算。因此本文设计的系统能够有效抵御Sybil攻击。

5 仿真实验

设计的仿真实验中，100个用户随机访问1 000个虚拟网页，其中包括900个健康网页和100个恶意网页。实验分为3种情况：第一种不应用协作防御系统；第二种社会信任不传递，只利用直接好友的浏览行为信息，即NetTrust系统采用的方法；第三种在信任传递的基础上，用户相互协作，共同防御恶意网页。

实验记录所有用户访问的恶意网页的总次数。对信任不传递和信任传递2种情况的实验结果进行比较。这2种情况下每步恶意网页的总访问次数与第一种情况相比，计算得到总访问次数的降低率，如图4所示。其中横坐标是实验步数，每访问一次网页作为一步；纵坐标表示恶意网页总访问次数的减少率，即应用防御系统和未应用防御系统恶意网页访问次数之差与未应用防御系统时访问次数的比值。应用防御系统，30步后，恶意网页访问减少率逐渐达到100%，说明用户最终都不再访问恶意网页。应用信任传递比没有信任传递恶意网页访问减少率明显增大，说明防御系统的防御能力比NetTrust系统有明显增强。

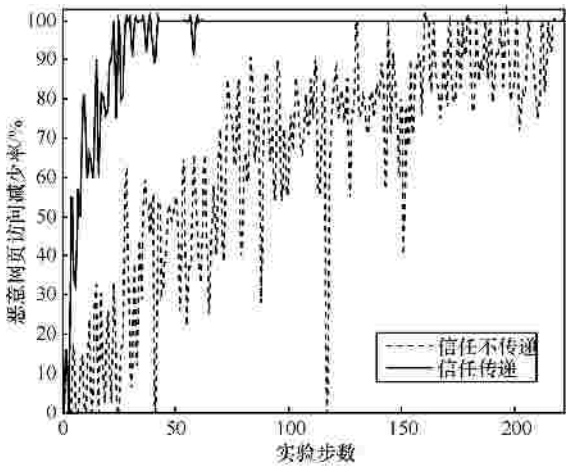

图4 2种情况下访问恶意网页减少率

用户对网页的主观评价有可能是不正确的。改变评价的正确率，在 60%、75%、90%、100%的用户评价是正确的情况下，与第一种情况进行比较，恶意网页的总访问量下降比例如图 5 所示。实验表明用户评价的正确率越高，系统防御能力越强，但是差别并不明显，说明防御系统对用户评价正确率具有弹性，系统可以容忍一定程度的用户错误评价。

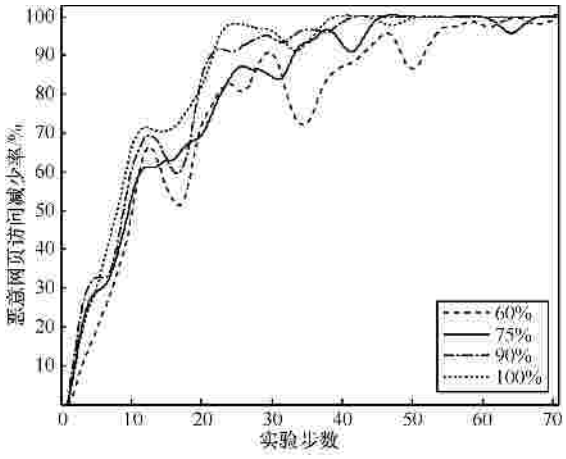

图 5　评价准确率不同时访问恶意网页减少率

恶意网页的数目会不断变化，设置实验中恶意网页的数目分别为：100、200 和 300，恶意网页总访问次数的减少率如图 6 所示，表明网络中恶意网页的数目对系统的性能几乎没有影响。

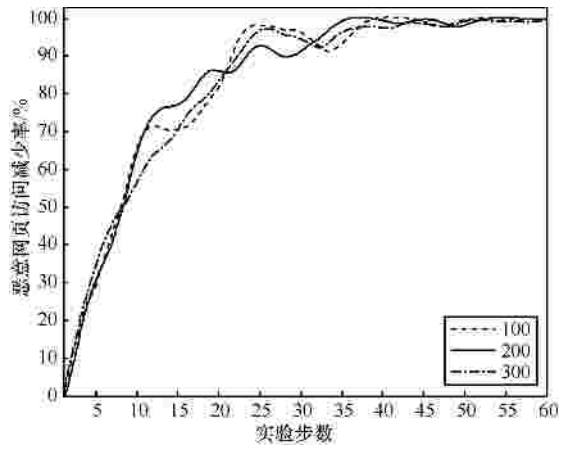

图 6　恶意网页数目不同时访问恶意网页减少率

有些用户不对网页进行评价，做出评价的用户分别为 50%、70%、90%和 100%时，恶意网页总访问次数减少率如图 7 所示。评价的用户越少，系统的性能会越差。只要有 50%的用户参与评价，即可避免相当一部分用户访问恶意网页。如果所有用户都不评价，相当于没有应用协作防御系统。但是在用户不直接对访问的网页做出评价时，系统可以应用行为评价获得网页的评价值，弥补这个不足。

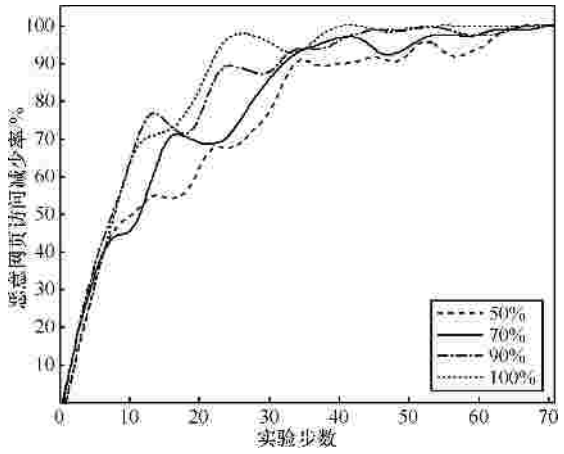

图 7　参与评价的用户的比例不同时访问恶意网页减少率

由于每个用户访问的网页数目有限，而且用户好友大多是其同学、同事或者亲朋好友，用户与其可信任好友通常会有共同的爱好兴趣或者成长经历，网页评价表中存储的信息是有限的。采用的信任传递算法是轻量级的，因此防御系统不会明显影响用户主机系统的开销。

6　结束语

本文提出了一种轻量级的基于社会信任的恶意网页分布式协作防御系统。通过即时通信网络利用好友间的信任，进行信任传递，存储和获取用户对网页的评价值，并且综合多个用户对于访问过的网页的安全经验，协作防御恶意网页，整个社会网络在宏观上形成一个网状的防御体系。由于即时通信的实时性，好友之间的经验共享比较及时，另外结合第三方专业机构提供的恶意网址列表，利用该列表的全面性对系统加以补充，更加有效地降低恶意网页的访问量。实验结果表明该系统显著提高了社会网络防御恶意网页的能力。

由于系统基于社会信任实现，可以保证获得的评价信息是可信的。但是在进一步分辨高水平评价信息的问题上还存在缺陷。在以后的工作中，可以采用有效反馈机制，用户对每个好友做出的评价给出一个反馈。将第三方资源的评价结果作为参考评判好友对网页评价的正确性，如果好友的评价正确，则提高其信任值；如果评价错误，则降低其信任值。随着网络的演化动态改变综合网页评价值，提高网络对恶意网页的实时防

御能力。

防御系统未区分社会网络中不同连接度节点的能力。在未来工作中，可以应用复杂网络理论分析社会网络特征，充分利用优势节点（好友特别多的节点）的作用，更加合理地部署防护机制。

参考文献：

[1] APWG. Anti phishing work group[EB/OL]. http://www.antiphishing.org, 2011.

[2] Google chrome and google safe browsing[EB/OL]. http://www. google.com/chrome/intl/zh-cn/more/security.html, 2011.

[3] SmartScreen filter[EB/OL]. http://windows.microsoft.com/en-US/ internet- explorer/products/ie-9/features/smartscreen-filter, 2011.

[4] Mark zuckerberg officially confirmed that facebook reached 750 million active users milestone[EB/OL]. http://www.techsnapr.com/2011/07/07/mark-zuckerberg-officially-confirmed-that-facebook-reached-750-million-active-users-milestone/, 2011.

[5] GOLBECK J. Weaving a Web of trust[J]. Science Magazine, AAAS, 2008, 321 (5896): 1640-1641.

[6] DONOVAN J O', SMYTH B. Trust in recommender systems[A]. IUI'05: Proceedings of the 10th International Conference on Intelligent User Interfaces[C]. New York, NY, USA, 2005. 167-174.

[7] SIRIVIANOS M, KIM K, YANG X. SocialFilter: introducing social trust to collaborative spam mitigation[A]. INFOCOM[C]. 2011. 2300-2308.

[8] BOYKIN P O, ROYCHOWDHURY V. Personal email networks: an effective anti-spam tool[J]. IEEE Computer, 2005, 38(4):61-68.

[9] YU H, KAMINSKY M, GIBBONS P B, *et al*. Sybilguard: defending against sybil attacks via social networks[A]. Proc ACM SIGCOMM[C]. 2006. 576-589.

[10] RESNICK P, ZECKHAUSER R. Trust among strangers in Internet transactions: empirical analysis of ebay's reputation system[J]. The Economics of the Internet and E-Commerce, Advances in Applied Microeconomics, Elsevier Science, 2002,11:127-157.

[11] PAGE L, BRIN S, MOTWANI R, *et al*. The Pagerank Citation Ranking: Bringing order to the Web[R]. Stanford University, 1998.

[12] JEAN CAMP L. Net trust: signaling malicious Web sites[J]. I/S A Journal of Law and Policy in the Information Society, 2007, 3(2): 211-235.

[13] Safe browsing tool: WOT (Web of trust) [EB/OL]. http://www. mywot.com/,2011.

[14] HANG C, WANG Y, SINGH M P. Operators for propagating trust and their evaluation in social networks[A]. Proceedings of the 8th International Joint Conference on Autonomous Agents and Multiagent Systems[C]. 2009.1025-1032.

[15] GOLBECK J. Computing and Applying Trust in Web-based Social Networks[D]. University of Maryland, College Park, 2005.

[16] **李勇军,代亚非**. **对等网络信任机制研究**[J]. **计算机学报** 2010, (3): 390-405.
LI Y J, DAI Y F. Research on trust mechanism for peer-to-peer network[J]. Chinese Journal of Computers, 2010, (3):390-405.

[17] AARON G, RASMUSSEN R. Global phishing survey 2h2010: trends and domain name use[EB/OL]. http://www.antiphishing.org/reports/APWG_GlobalPhishingSurvey_2H2010.pdf, 2011.

[18] Google safe browsing API[EB/OL]. http://code.google.com/intl/en/apis/safebrowsing/, 2010.

作者简介:

刘昕（1974-），女，山东青州人，南开大学博士生，主要研究方向为网络与信息安全、可信计算。

贾春福（1967-），男，河北文安人，博士，南开大学教授、博士生导师，主要研究方向为信息安全与可信计算、恶意代码发现与分析。

刘国友（1986-），男，天津人，南开大学硕士生，主要研究方向为信息安全。

胡志超（1987-），男，河北沧州人，南开大学硕士生，主要研究方向为信息安全。

王冬（1985-），男，山东济南人，南开大学硕士生，主要研究方向为信息安全。