

# 威客平台信誉能力评价机制研究

刘 寅<sup>1</sup> 吴毅坚<sup>1</sup> 彭 鑫<sup>1,2</sup> 赵文耘<sup>2</sup>

(复旦大学软件学院 上海 201203)<sup>1</sup> (复旦大学计算机科学技术学院 上海 201203)<sup>2</sup>

**摘要** Witkey(威客)模式鼓励网民将自身专业知识、技能、经验、能力通过社区转换成具有经济价值的知识产品并与其他人分享。但目前的威客网站的信誉评价机制都比较简单,无法很好地反映威客的真实信誉能力,不利于项目雇主选择合适的合作伙伴。针对软件项目领域开发者的技术和信誉能力特点,提出了一套结合客观度量、主观评价的威客信誉能力评价模型。实验中,采用典型的场景模拟了威客在平台上的成长历程,分析了使用该模型的评价结果对威客成长的引导作用,从而验证了该评价模型的可行性和可用性。

**关键词** 威客,评价模型,项目外包,信誉

**中图分类号** TP311.5 **文献标识码** A

## Research on Witkey Credit Ability Evaluation Mechanism

LIU Yin<sup>1</sup> WU Yi-jian<sup>1</sup> PEN Xin<sup>1,2</sup> ZHAO Wen-yun<sup>2</sup>

(Software School, Fudan University, Shanghai 201203, China)<sup>1</sup>

(College of Computer Science and Technology, Fudan University, Shanghai 201203, China)<sup>2</sup>

**Abstract** Witkey's idea aims at encouraging people to transform their specialist knowledge, skills, experience, and ability to valuable products by the help of the online community and share them with other people. But nowadays, many Witkey community websites' credit ability evaluation systems are not only too simple to reflect their users' real credit level, but also go against with the project publishers to choose their partners easily. After analyzing developers' technology and credit merits and demerits in software development filed, we raised an objective estimate and subjective evaluation model and specifications. The experiments use classic scenes to simulate the Witkey users' growing-up experience, then analyze the guiding effect to the users' development if applying such evaluation system, finally prove its feasibility and usability.

**Keywords** Witkey, Evaluation model, Project outsourcing, Credit

## 1 引言

威客(Witkey)泛指那些通过互联网社区把自身专业知识、技能、经验通过社区平台转换成具有经济价值的产物并与其他人分享,最终获得酬金的人<sup>[1]</sup>。威客模式的主要应用在于解决科学、技术、工作、生活、学习等领域的问题,体现了互联网按劳取酬和以人为中心的新理念<sup>[2]</sup>。该模式在软件开发领域里也有着广泛的应用。例如,企业通过委托威客发送各种软件工具,来降低软件系统开发成本,使企业能够将资源集中在其核心业务以提高竞争力上。

随着互联网的蓬勃发展,越来越多的网民开始选择尝试使用威客社区来承接一些额外的项目工作,这样不仅能够锻炼自身技能,分享知识,还可以从中获得可观的金钱报酬。不少企业和个人也愿意将一些疑问或者任务通过威客平台悬赏发布,并最终得到满意的解答。

2004年创办的“K68创意产业平台”是我国最早的商业化威客网站。经过7年多的发展,国内已有威客网站<sup>[3]</sup>超过100家。国内其他知名的威客网站还包括猪八戒网、威客中国等<sup>[4]</sup>。全球范围内规模和知名度较高的3家威客网站分别是Eance<sup>[5]</sup>、oDesk<sup>[6]</sup>和FreElancer<sup>[7]</sup>。其项目业务涵盖范围包含软件开发、图形logo设计、写作等。此外国外的威客平台上还有不少公司团队级别的企业用户,他们不仅发布项目,还积极参与到项目竞标中去。

虽然威客模式已经逐渐流行,但现有的威客网站里普遍存在:作弊以及知识产权侵犯问题;威客评价反馈率低问题;对于不同类型的外包任务类型,威客平台没有相应的评价类型对应等问题。可见,在威客模式中,需要有一套客观公正的评价机制来支持日常业务的开展。

本文对威客平台的信誉能力数据深入挖掘分析后,选取软件外包项目领域建立评价模型,该模型考虑了软件外包的

到稿日期:2012-02-28 返修日期:2012-05-29 本文受国家自然科学基金(60903013),国家高技术研究发展计划(863)(2012AA011202)资助。

刘 寅(1986—),男,硕士生,主要研究方向为软件产品线、网络化开发技术,E-mail:09212010013@fudan.edu.cn;吴毅坚(1979—),男,讲师,主要研究方向为软件体系结构、软件复用和产品线;彭 鑫(1979—),男,副教授,CCF会员,主要研究方向为构件技术与软件体系结构、软件产品线、软件维护与再工程等;赵文耘(1964—),男,教授,博士生导师,CCF高级会员,主要研究方向为软件工程、电子商务以及企业应用集成(EAI)。

一些特有的客观因素指标,结合了威客平台上雇主的主观评价因素指标,从而能够更加客观、全面、真实地评定每一个威客用户,使得项目雇主能够更好地认清参与项目竞标威客们的优劣,最终选定合适的项目合作伙伴;并为威客们提供一个利于成长的良性社区氛围。本文最后通过一个实验案例来说明该评价模型机制对于威客的健康成长以及客观评价是有效的。

## 2 相关工作

威客平台上的项目涉及到大量金钱和知识产品交换的交易,而且平台上人与人之间不熟悉、信息不对称,在这种情况下,为了尽可能降低风险,人们会考虑第三方(威客平台)所提供的推荐参考信息<sup>[8]</sup>。

而 Resnick 等<sup>[9]</sup>认为基于在线平台的评价推荐系统在网络环境下,通过收集、合计、发布用户的历史数据来激励陌生人之间的相互沟通和合作,有助于促进网络社区内的相互信任,从而营造一个良好的社区环境氛围。

数学方法里有 Analytic Hierarchy Process(AHP,结合层次分析法)和模糊理论,其采用 AHP 对外包软件威客的各项评价指标权重进行分析,将问题分解为若干部分,构成若干元素,各个元素按照属性和权重归为不同组,形成不同的结构层次<sup>[10]</sup>,从而对威客进行客观的综合评价,以便于软件发布方(甲方)选择合适的合作伙伴<sup>[11]</sup>。虽然这种方式有良好的数学理论背景,但其参数的设定与确认十分依赖于专家的评定,并不适合于威客平台的自动化评价机制。

另外,纪淑娴等的研究指出诸如简单累加信用度或者均值的计算模型仍是在线信誉管理系统的两种主要计算模式,但这两种模型都忽略了一些影响信用度的重要因素<sup>[12]</sup>。为了得到有效、实用、客观的评价价值,必须考虑到威客社区平台用户的需求和特点。

JinSonghe 等人虽然针对威客模式提出了一套基于任务和威客的推荐匹配机制<sup>[13]</sup>,但并未考虑到平台上有着各种不同种类的项目,因而将会有着不同的项目特性与之对应,所以它仍有待改进。

正如软件产品的质量取决于软件的开发过程,通过建立一致、标准的软件开发过程,可以提高对软件项目开发的控制性并提高开发效率<sup>[14]</sup>;同样,良好氛围的威客社区平台取决于平台是否能够提供一个客观、公正的评价机制,从而鼓励诚信交易、打击作弊违规行为并最终营造出一个活跃、诚信、开放的威客社区。

而对于作弊用户以及虚假项目交易欺诈检测机制,虽然威客社区平台可以借鉴一些 C2C 电子商务平台的策略,应用社会网络的方法去自动甄别那些可疑度较高的用户,继而交给管理员进行最终的人工审核<sup>[15]</sup>,但我们也应认识到威客社区和电子商务社区无论在用户规模还是业务数量上都不是一个数量级,因而有必要针对威客社区的用户和项目数据进行深入分析并制定相应对策。

## 3 威客平台的评价机制分析

威客平台方除了作为项目信息和用户信息的发布者之外,还通过对发布的项目进行验收管理等方面的工作,介入到

双方的交易活动中来。

### 3.1 运作模式

现有威客平台的主要运作流程模式主要分为两种:现金悬赏模式以及任务招标模式<sup>[3]</sup>。

由于现金悬赏模式会导致参与竞标的威客数量较多(其中不乏一些恶意竞标者)而且水平参差不齐,难以保证项目雇主能获得满意的作品,因此对于一些大型复杂、涉及金额较多的项目,推荐使用任务招标模式。目前,全球大多数威客网站采取的是任务招标模式。

### 3.2 Elance 的评价案例研究

#### 3.2.1 威客评价机制基础数据分析

Elance 上每当一个项目被其雇主确认完结后,平台会将之前代为保管的酬金支付到威客的账户里,之后雇主对威客在项目中的表现进行评分;其中项目数和项目得分分别是威客的历史项目总数及其平均得分;雇主人数是威客在历史项目中接触过的项目雇主总人数;雇主推荐率是每个项目的发布者在项目结束后根据威客在项目中的表现决定是否向公众推荐这位威客;雇主回访率则是衡量一位威客的历史项目记录里有多少项目的雇主是以前的“老东家”;项目节点是雇主和威客约定的应当提交某项产物的时间节点。

为了能够对现有威客平台的现状有一个清晰的认识,通过程序抓取 Elance 上的用户信誉能力评价数据进行定量的分析并对 Elance 上有过项目经验的个人威客数据(20019 条记录<sup>1)</sup>)进行深入分析后,根据威客的平台等级和其它一些评价因素进行归类整理分析得到表 1(Elance 平台上还有大量的企业级威客用户,由于他们大多是团队、单位而非个人,故不在本文的讨论范围之内)。

表 1 Elance 平台威客数据分析

威客等级	人数	项目数	项目得分	雇主人数	雇主推荐率	雇主回访率	工作时间(小时)/项目节点
11 级	8	9.875	4.975	6.875	97%	18%	1197.3/10.5
10 级	59	66.32	4.896	37.474	98%	33%	924/103.5
9 级	100	66.27	4.894	33.93	98.4%	32%	500.8/96.26
8 级	360	63.98	4.892	37.54	99%	28%	312/84.51
7 级	607	51.67	4.894	30.29	99%	26%	173.7/66.1
6 级	1570	36.08	4.888	22.53	99%	25%	86/42.59
5 级	2809	21.71	4.849	14.1	98%	25%	57.4/23.5
4 级	3477	7.77	4.805	5.4	93.6%	17%	37.9/8.98
3 级	7604	4.16	4.704	2.97	96%	14%	13.02/4.04
2 级	3425	2.48	4.845	1.79	96.3	4%	14.67/2.77

通过表 1 的数据分析,可以清楚地得出以下两个结论:

1)平台威客等级呈现金字塔分布:不同等级之间的人数区别较明显,高等级人数在数量级上远少于中低等级用户,其中 5 级以下(含 5 级)威客用户人数占到了总数的 86.5%。

2)威客等级和能力总体是正相关的:威客所承接的项目数量、雇主人数随着威客等级的下降而下滑,可见只有大多数具有丰富历史项目经验背景的威客才能够在平台上获得较高的等级;参考 Elance 上设置的评价因子:雇主回访率以及项目所涉及的工作时间/项目节点,根据图表很容易看出其随着威客等级的下降而下滑,因而只有认真对待项目赢得雇主的高评价才能获得较高的平台等级。

总体来看,Elance 的威客评价机制能够较好地管理评价现有威客群体的能力等级划分,但通过对数据进行深入剖析

<sup>1)</sup> Elance 平台截至 2012 年 1 月 6 日的威客用户数据

后,发现了几个潜在的问题。

### 3.2.2 评价机制存在的问题

评价机制存在的问题包括:评价机制简单、作弊问题、反馈机制利用率较低和不利于新手起步。

#### 1) 评价机制简单

现有的评价机制大多没有充分利用挖掘平台上现有的历史数据(诸如提交项目的代码行数、文件数量、项目涉及金额、项目持续时间等一些客观数据),因而无法起到客观地评价用户信誉及能力的作用;而且 Elance 目前的评价体系结构不够细化,对于不同类型的项目采用相同的评价机制,不利于项目雇主有针对性地选择合适的合作伙伴;通过分析 Elance 上雇主给威客的项目评分和推荐率这两项指标(见图 1),52% 威客的历史项目评价分是满分 5 分,44% 在 4.9~4 分这个区间段,可见这样对于威客优劣的区分度不够明显。

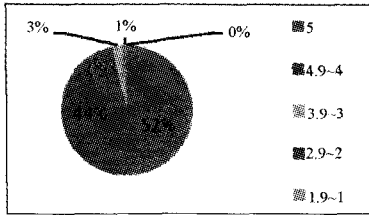


图 1 威客项目平均得分

#### 2) 作弊问题

由于威客平台里开放、自由的双向选择机制,因此在项目实施的过程里常会出现各种违规、作弊的情况。现有平台往往缺乏自动化的作弊检测方案而且依赖人工审核或者用户举报,无法起到有效的监督职能,继而不利于营造良好的平台社区氛围<sup>[16]</sup>。比如一个项目经验丰富的高等级威客,仔细检查其历史记录后发现其所获得的经验和酬金都来自于同一个项目雇主,而该雇主也只和这位威客有过合作记录,可见该个案是一个十分可疑的场景;此外雇主骗取威客“智慧产物”也是一个常见的作弊情况<sup>[17]</sup>。

Elance 平台中,排名前两万的威客个人用户里,有 71 位威客所承接的历史外包任务数少于 3 个(部分数据参考表 2),但其威客等级都已经达到或者超过 6 级,而平台里 6 级及以上用户仅占 13.5%。

进一步分析后发现:a)其客户数都只有 1~2 人;b)雇主回访率大多是 0% 或者 100%;c)没有获得过雇主推荐(雇主推荐率这项都是 N/A)。这些威客的历史项目大多持续时间较短,但涉及金额较大;而且这些项目的项目雇主也很少与其它平台用户合作。因此,我们认为这些用户有重大涉嫌作弊的嫌疑。

表 2 部分涉嫌作弊威客数据

用户 id	项目数	时间节点	工作时间	雇主推荐率	雇主人数	雇主回访率	威客等级	项目得分
923	1	0	37	N/A	1	0%	11	N/A
3889	1	0	420	N/A	1	0%	11	5.0
7234	1	0	726	N/A	1	0%	11	5.0
9011	1	0	2898	N/A	1	0%	11	4.8
10551	2	102	0	N/A	1	100%	10	5.0
13781	1	1	0	N/A	1	0%	10	N/A
17681	2	0	12	N/A	1	100%	9	5.0
19076	1	0	67	N/A	1	0%	9	N/A

#### 3) 信誉评价机制利用率低

Elance 主要通过项目评分、评语和雇主推荐率来衡量一个威客的信誉和能力。

但我们发现平台上至少有过一次项目经历的威客用户里有超 52% 的人从未获得过来自于雇主的评语或者推荐。这对于其它雇主在进行项目合作伙伴选择时是不利的。另一个令人担忧的现象是:很多中级用户(3 级至 5 级)在有了多项项目经验后,仍然未获得雇主的评语和推荐,这一点值得平台加以改进,以鼓励雇主多对威客进行评价。

此外,平台还缺乏项目合作双方的互评鼓励机制;很多项目在佣金结算结束后就一拍两散;对排名前两万的个人威客用户数据分析显示(见图 2),只有 37% 的威客得到了推荐;这样长久以往对于平台发展是不利的。

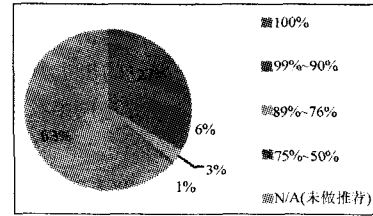


图 2 项目雇主推荐率

#### 4) 不利于新手起步

对于新用户群体,其资质能力及信誉评价尚处于起步阶段,难以在激烈的竞标中得到项目雇主的认可从而赚到威客生涯的“第一桶金”。通过对现有威客平台的大量用户数据分析,不难发现用户等级普遍呈现金字塔形,低等级用户占平台用户总数的绝大多数,Elance 上注册的有效个人威客用户共有近 48 万,而有过至少 1 次项目经验的用户却只有不到两万,仅占 Elance 全体威客的 4%(见图 3)。

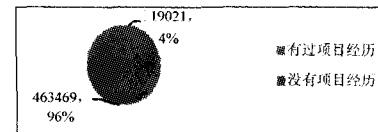


图 3 有过项目经验的威客比例

平台对于这部分用户必须给予足够的关注,从而使得社区逐步向稳定的橄榄型发展(从社会学意义上说:中间阶层的壮大,可以使得每一个社会成员看到晋升的希望,有助于缓解差距过大产生的对立情绪等一系列由此衍生的社会矛盾),使得他们能够尽快获得第一次雇主的亲睐,顺利踏上威客成长的历程。

## 4 面向软件研发任务的信誉能力评价体系

由于在软件外包领域,很多评价机制标准的打分和权重都是基于一个动态而且模糊的评分<sup>[9]</sup>,因此将一个基于威客各项数据信息的评价系统用作平台的核心机制来支持项目雇主选择项目合作伙伴是十分重要的。

### 4.1 评价体系概述

威客信誉能力评价体系主要由技术与信誉评价指标和可信度组成。

技术与信誉评价指标是综合考虑了在软件项目中一些会影响项目雇主评价的客观因素诸如代码、文档等,继而建立起的一套评价指标。可信度则是在基于平台的历史数据的基

基础上,综合分析威客与其他用户的平台数据后给出的威客在平台上的可信程度。其目的主要是消除雇主对陌生威客的不信任感。

可信程度指标里虽有部分参数(威客所获的酬金总额  $P$  等)在技术与信誉评价指标里也有出现,但其主要职能是与平台的基准值进行对比,从而生成一个全新的数据以备平台生成威客的可信程度值。所以重复出现的参数在不同的评价指标里职责也不同。

项目雇主可以结合威客的技术能力和可信程度这两项指

标来评定该威客是否值得作为项目的合作伙伴。建立评价体制的目标在于对交易双方的行为产生约束作用,从而限制作弊行为,鼓励诚信交易,降低信用风险,最终营造一个良性的社区氛围。

#### 4.2 技术与信誉评价指标

技术与信誉指标包含技术能力评价指标和历史项目雇主评价指标,分别对威客的个人项目技术实力和历史项目的信誉度进行定量的度量。首先威客的技术与信誉值  $E$  由其历史项目评价价值  $A$  和历史雇主评价价值  $B$  综合决定(见表 3)。

表 3 威客技术与信誉评价指标

评价项目	评价指标	评价因子及定性描述	参数描述	定量计算公式
A. 技术能力评价指标	A1 软件技术能力评价	A 软件研发能力:结合历史项目的代码行数 and 项目质量,项目雇主和平台给予威客的评分	$L_n$ :项目代码行数(不包含注释) $F_n$ :文件个数	$a=(L_n/200+F_n)$
		b 文档撰写能力:威客在开发和交付项目的同时,相关的需求,开发,测试,使用文档的质量和完整度,平台和项目雇主给予威客的评分	$SF_n$ :威客提交的项目文档数量 $RF_n$ :雇主在项目中要求的文档数量 $DocScore$ :雇主对文档的评价分数(1~10分)	$b=(SF_n)/(RF_n) * DocScore$
	c 沟通管理能力:每一个软件项目结束后,由软件外包雇主根据威客的表现给出一个评价	主观评价:根据好坏分为四档	c:优秀(9~10),良好(7~9)一般(4~7),较差(4分以下)	
	A2 项目开发成功率	威客的历史项目开发成功率	$WP_n$ :威客获得奖金的项目数 $P_n$ :威客完成的项目总数	$A2=WP_n/P_n$
A3 历史项目违约率	威客在历史项目中的负面表现率	$I_n$ :威客被平台裁定违规次数	$A3=I_n/P_n$	
A4 项目酬金总额评价	威客在项目中获得酬金总额评价	$P$ :威客所获的酬金总额	$A4=P/10$	
B. 历史项目雇主评价指标	B1 历史项目雇主评价	威客在历史项目中表现的评定值	$HPP$ :项目雇主对于威客在项目中的表现以及提交作品的综合评价分数	$B2=HPP$
	B2 项目合作伙伴总数	威客的在历史项目历程中与多少雇主有过合作关系	$SumPon$ :与威客有过项目合作经历的项目雇主总数	$B3=SumPon$
	B3 历史雇主回访率	威客历史项目中有多少雇主多次与之合作的评定值	$XPOn$ :和威客有过复数次合作经历的项目雇主人数	$B1=XPOn/SumPon$

每当威客完成了一个项目后,平台首先结算该项目中威客所获得的技术能力评价价值  $A$ ,在对威客平台猪八戒网<sup>[20]</sup>上的项目酬金和威客技术能力数据分析统计后,将酬金( $A4$ )和技术能力( $A1$ )在整个技术能力评价指标里的比重设定为 3:7。

$$A = f(A1, A2, A3, A4) = \sqrt{(0.7 * A1 + 0.3 * A4) * A2 * (1 - A3)}$$

对 Elance 平台上的软件项目相关的威客数据分析后,本文将威客的软件研发能力  $a$  设定为 65%,文档撰写能力  $b$  设定为 20%,项目沟通管理能力  $c$  设定为 15%。

$$A1 = (0.65 * a) + (0.2 * b) + (0.15 * c)$$

之后综合历史项目数据得出威客的项目雇主评价指标  $B$ ;

$$B = g(B1, B2, B3) = \sqrt{(B2 * B1)^{(1+B3)}}$$

$$最后得出威客的经验值 E = h(A, B) = \sum_{i=1}^n (A_i) + B$$

#### 4.3 可信程度指标

在平台上注册时间长、有着较多项目经验的资深威客能够得到项目雇主的信赖。参与涉及金额越大、持续时间较长的项目比小项目更有价值,因而对其所获得的经验值和能力值也应当量化地区别对待;最后,违约违规被投诉的记录对于分值的影响必须考虑在内,以起到公正公平的评价作用,也是对用户一个警示。综合上面提出的多项参数指标,平台评价机制给出一个基于历史数据的对于每个软件的威客可信程度分值  $R$ (见表 4)。

表 4 威客可信程度度量因子表

威客可信程度度量指标	参数描述	详细计算公式
R1:注册时间相关的可信度	$RT$ :用户在平台的注册时间; $ART$ :平台注册时间基准值	$(RT / ART) \times 20\%$
R2:威客成功的项目数	$P_n$ :威客完成的项目总数; $AP_n$ :平台项目数基准值	$(P_n / AP_n) \times 20\%$
R3:威客的开发能力值	$Exp$ :项目开发经验; $AExp$ :平台威客经验基准值	$(Exp / AExp) \times 30\%$
R4:历史项目信誉评价价值	$P$ :威客所获的酬金总额 $HPP$ :项目雇主对于威客在项目中的表现以及提交作品的综合评价分数	$(P/1000 + HPP) \times 30\%$
Q:历史违约记录评价价值	$I_n$ :威客被平台裁定违规次数; $P_n$ :威客完成项目总数	$1 - (I_n / P_n)$

威客可信程度分值  $R = (R1 + R2 + R3 + R4) \times Q$ ,作为一个度量指标,和威客经验值一起作为两个评价威客信誉能力等级的度量值。

平台根据上述指标,赋予每个威客对应的信誉能力和历史项目数值;外包雇主可以通过对各项平台参数(威客等级、历史项目数、项目评价分等)进行搜索、对比,选择出合适的威客进行进一步交流。

## 5 数据实验和评价

### 5.1 实验目标系统

为了验证本文提出的威客信誉能力评价体系,最理想的目标实验数据应当是带有项目酬金信息和项目雇主评价数据的软件项目。但由于此类数据对象涉及到知识产权和机密性

等原因很难收集到,因此为了验证本文提出的信誉能力机制模型的有效性和可行性,实验分别选取两个开源项目 Openfire 和 Arale 作为实验数据;Openfire 是一个基于 XMPP 协议的高效开源即时通信软件服务器项目<sup>[18]</sup>,而 Arale 是一款高效的网络爬虫下载工具<sup>[19]</sup>,两者都是基于 Java 的开源项目。项目的基本信息参见表 5。

表 5 实验目标项目基本数据

项目名称	项目文件个数	项目代码行数
Openfire	1089	299746(275.25 行/文件)
Arale	7	1193(170.43 行/文件)

## 5.2 数据假定

对于一个新加入平台的威客用户来说,Arale 这样的小型外包软件项目是一个理想的上手项目,有过多次成功的同等规模的项目经验后,该用户将累计其可观的技术能力经验值,他也会有与不少项目雇主合作的经验。

但是由于这两个项目并非商业项目,所能获取的项目数据并不完全适用于本文提出的这套威客技术能力评价指标。为了将威客技术能力评价指标用于这两个项目应用,本文做如下假定:

1. 为了方便实验的进行,实验中两个项目涉及到的文档撰写能力、项目沟通能力、历史项目雇主评价分别恒定为 9 分,9.5 分和 9 分。

2. 雇主回访率:在实验的前半部分,保持雇主回访率恒定不变以查看威客的不同雇主策略对于其所获得的经验值的影响;而实验后半部分则改变回访率,查看不同回访率对于威客所能获得经验值数值的影响。

3. 项目酬金:通过分析国内大型威客平台猪八戒网<sup>[20]</sup>的一些相关软件项目酬金信息后,实验设定 Openfire 项目的酬金为 80000 元,Arale 的项目酬金为 1500 元。

## 5.3 数据实验过程和结果分析

为了分析技术与信誉指标与威客成长过程的关系,我们的模拟采用了 11 个项目(以 Arale 和 Openfire 为假想项目),分析了威客经历这些项目的成长过程。

该模拟成长过程分为两个阶段:第一阶段是项目 1-10,都是小型项目(以 Arale 为假设代表);第二阶段是项目 11,是一个大型项目(以 Openfire 为假设代表)。

由于实际开发项目分配情况的多样性,我们仅采用两种极端情况(模式 A 和模式 B)来模拟威客和雇主的双向选择关系。

模式 A:本文设计了 3 个场景(包含两个模式)来验证我们的信誉能力评价体系,其中图 4 模拟了威客信誉能力值经历各种不同的项目模式后的成长过程,该图横坐标为经历的项目数量,纵坐标为威客所积累的技术与信誉值。

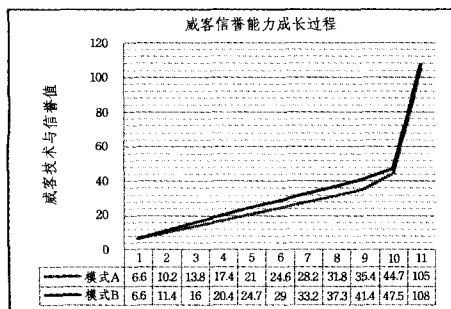


图 4 威客信誉能力成长过程

模式 A 线图在第 1 个项目至第 9 个项目期间模拟了一位威客与同一位项目雇主合作的情况(其历史雇主回访率从第二个项目开始恒定为 100%)。在进行第 10 个项目时,选择了新的一位雇主(雇主回访率由 100%变为 50%)。

模式 B:在第 1 个项目至第 9 个项目期间模拟了一位威客与多个互不重复的项目雇主进行合作的情况(其回访率从第一个项目开始恒定为 0%)。在进行第 10 个项目时,选择了之前有过项目经历的雇主(雇主回访率由 0%变为 11.1%)。在此情况下威客技术与信誉指标值的增长曲线如图 4 所示。

从图 4 中可发现本评价模型的如下特点:

### 1. 鼓励多元合作

在第一阶段,对于只和同一位项目雇主有项目往来的威客(模式 A),其获得的经验值将略少于与多位雇主有过项目经历的威客(模式 B)。本文提出的评价机制对于多元的项目合作采取了鼓励,希望威客多与其他陌生的项目雇主进行合作,从而激励用户之间的相互沟通与接触。

### 2. 激励威客赢取稳定的市场(雇主)和寻找新的市场(雇主)

在第一阶段进行第 10 个项目时,两种模式产生了雇主回访率的变化。模式 A 从单一雇主变为找到了新雇主;模式 B 从多个雇主变为有了重复的雇主。从图 4 中可以发现两种模式都在第 10 个项目中取得了比前几个项目更多的经验值。这个变化说明,本文提出的评价机制在鼓励多元合作的同时,也赢取雇主的回访。这要求威客认真对待每一个项目,以赢取雇主的信任和回访,并创出自己的品牌以吸引更多的雇主。

### 3. 鼓励威客挑战困难的或规模较大的项目

第二阶段体现了项目规模的变化对威客所获得经验值的影响。图 4 中体现了经验值呈现一个飞跃的提升。可见本文提出的机制鼓励有能力的威客承担大项目,能够较客观评价一个威客在项目中的表现,如果成功则能赋予其较高的经验值。但同时威客也需承担相应的风险。如果威客在一个项目中失败(由于威客自身原因未能完成任务、被项目雇主投诉、以及被平台判定违规作弊等情况),他不仅在本项目中获得的信誉能力值会大幅减少(历史项目的违约率会降低威客的技术能力评价指标),而且会影响到他在平台上的威客可信度指标(历史违约记录评价值的降低会极大地影响威客可信度分值)。因此这对于一些抱有侥幸心理的投机分子来说能起到警示的作用。

**结束语** 本文对目前的威客平台现状进行了深入分析,针对著名的 Elance 平台的数据进行了充分挖掘,提出了一些值得关注的问题和必须改进之处,并针对上述问题提出了对应解决方案:一套基于现有平台数据的信誉能力评价指标,用以客观地评价每一个威客的技术能力和雇主对其综合评价,并给出一个基于平台基准值的威客信赖度指标,以方便项目雇主选择合作伙伴。本文的实验部分模拟了新人威客在平台上经历项目的成长过程,在经历不同的项目和雇主后,威客的技术与信誉能力值也会有相应的增长,体现出平台能对威客的发展起到良好的引导作用,从而营造一个雇主和威客双方互信的良性互联网生态社区。

本文虽然对 Elance 平台的大量数据用户进行了深入数据分析,但我们也注意到:不同工作性质的威客之间,其信誉

(下转第 49 页)

从表 8 可以看出,基于相同的训练集及测试集,P-Abacus 比 Abacus 的训练时间增加 3.6%,其测试时间只增加 0.7%。

综合识别准确率以及时间开销两方面来考虑,P-Abacus 与 Abacus 相比,其识别精度从 93%增至 98%,尽管训练时间增加了 3.6%,但是识别时间却只增加 0.7%,因此从识别方法的实用性来考虑,P-Abacus 比 Abacus 更优。

**结束语** 本文针对 P2P 流媒体流量的精细化识别法进行了研究,对 Abacus 进行了改进,提出了一种基于 SVM 概率输出的识别法 P-Abacus。该方法能够将待识别实例属于已知应用的可能性大小反映在概率输出上,并结合概率判决对输出结果进行判断。由于 SVM 的概率输出包含更多的信息,同时,其概率判决中的二次判断也对最优超平面附近的实例进行了细致的识别,因此 P-Abacus 具有更好的识别效果。通过对 7 种较流行的 P2P 流媒体应用进行的实验测试表明,P-Abacus 比 Abacus 具有更高的识别率和更低的误判率,并且时间开销增加有限。

本文基于 P2P 流媒体应用的全部流量提取特征,因此 P-Abacus 方法同 KISS、PSD、Abacus 一样,只能部署在网络出口处,针对骨干链路上 P2P 流媒体流量的识别将是下一步的研究重点。

## 参 考 文 献

[1] 张艺瀛,张志斌,赵咏,等. TCP 与 UDP 网络流量对比分析研究[J]. 计算机应用研究,2010,26(6):2192-2197

[2] Perenyi M, Dang T D, Gefferth A, et al. Identification and Analysis of Peer-to-Peer Traffic [J]. Journal of Communications, 2006,1(7):36-46

[3] 雷蕾,沈富可. 基于连接特征的 P2P 流媒体应用的识别[J]. 计算机应用,2007,12:41-43

[4] Moore A, Papagiannaki K. Toward the accurate identification of network applications[C]//Proceedings of the Passive and Active

Measurements Workshop. 2005:41-54

[5] Nguyen T, Armitage G. A survey of techniques for internet traffic classification using machine learning[J]. IEEE Communications Surveys, 2008,10(4):56-76

[6] Liu Chao-bin, Yang Yue-xiang, Tang Chuan. A Classification Method of Unstructured P2P Multicast Video Streaming Based on SVM[C]//2009 IEEE International Conference on Multimedia Information Networking and Security(MINES2009). 2009:11

[7] Silverston T, Fourmaux O. P2P IPTV measurement: a comparison Study[M]. Preprint, 2006

[8] Alessandria E, Gallo M, Leonardi E, et al. P2P-TV Systems under adverse network conditions: a measurement study [C] // IEEE Infocom. Riode Janeiro, 2009

[9] Finamore A, Mellia M, Meo M, et al. KISS: Stochastic Packet Inspection[C]//Proceedings of the First International Workshop on Traffic Monitoring and Analysis (TMA). Aachen, Germany, 2009

[10] Li Jin, Zhang Xin, Zuo Xiao-liang, et al. Using Packet Size Distribution To Identify P2P-TV Traffic[C]//2010 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery(PSD). Oct. 2010:150-155

[11] Valenti S, Rossi D, Moe M, et al. Accurate, Fine-Grained Classification of P2P-TV Applications by Simply Counting Packets [M]. Papadopouli M, Owezarski P, Pras A, eds., TMA 2009. LNCS 5537, Berlin Heidelberg: Springer-Verlag, 2009:84-92

[12] Bermolen P, Mellia M, Meo M, et al. Abacus: Accurate behavioral classification of P2P-TV traffic[J]. Computer Networks, 2011,55(6):1394-1411

[13] 李国正,王猛,曾华军. 支持向量机导论[M]. 北京:电子工业出版社,2004

[14] 吴朝晖,杨莹春. 说话人识别模型与方法[M]. 北京:清华大学出版社,2009

[15] Weka[OL]. <http://www.cs.waikato.ac.nz/ma/weka>

(上接第 30 页)

能力数据也有着迥异。本文的分析是基于一个平台整体的用户统计分析,如果将用户根据其工作归类后再分析,应该能得到一些更为准确的分析结论。另外对于单个威客平台 Elance 的数据分析有局限性,将全球另外两家平台数据收集起来加以横向比较,也是未来的工作方向之一。

## 参 考 文 献

[1] 王伟军,孙晶. Web2.0 的研究与应用[J]. 情报科学,2007,25(12):1907-1931

[2] 刘锋. 威客(witkey)的商业模式分析[D]. 北京:中国科学院研究生院,2006

[3] AbleSky 国内威客网站调研报告[OL]. <http://www.ablesky.com/viewCourseDetail.do?courseId=8223>,2011-11-25

[4] 2010 中国威客行业白皮书[OL]. <http://www.iresearch.com.cn/report/>,2011-11-25

[5] Elance[OL]. <https://www.Elance.com/>,2012-01-12

[6] oDesk[OL]. <https://www.odesk.com/>,2012-01-12

[7] Freelancer[OL]. <http://www.freelancer.com/>,2012-01-12

[8] 樊丽杰,王素贞,刘卫. 基于人类信任机制的移动电子商务信任评估方法[J]. 计算机科学,2012,39(1):190-192

[9] Paul Resnick AT, Labs-Research T, Hill M, et al. Recommender

Systems[J]. Communications of the ACM,1997,40(3):56-58

[10] 李杨,韦伟,刘永忠. 一种基于 AHP 的信息安全威胁评估模型研究[J]. 计算机科学,2012,39(1):61-64

[11] 吴明晖,曹梦筠,王海涛. 基于 FAHP 的综合评价系统及在软件外包中的应用[J]. 计算机系统应用,2008(11):91-94

[12] 纪淑娴,胡培,程飞. 在线信誉管理中信用度计算模型研究[J]. 预测,27(4):59-64

[13] Jin Song-he, Song Bao-wei, He Lei. Recommendation of Online Tasks Based on Witkey Mode Website[Z]. IFITA. 2009:268-270

[14] 刘瑞芳,谢长生,谭志虎. 基于 CMM 的软件开发过程研究[J]. 计算机应用研究,2004(7):133-134

[15] 尤薇佳,刘鲁,杨俊杰,等. 基于交易记录的欺诈识别[A]//第二届网商及电子商务生态学术研讨会论文集[C]. 2009:178-182

[16] 朴春慧,韩旭芳,陈挥青. 威客电子商务作弊评价模型及算法研究[J]. 科学的实践与认知,2010,40(11):124-130

[17] 朴春慧,韩旭芳,杨春燕. 基于 Web2.0 的威客电子商务作弊处理机制研究[J]. 情报杂志,2009(10):124-128

[18] Openfire 开源项目[OL]. <http://www.igniterealtime.org/projects/openfire/>,2012-01-08

[19] Arale 开源项目[OL]. <http://flavio.tordini.org/arale>,2012-01-09

[20] 猪八戒威客网站[OL]. <http://www.zhubajie.com/>,2012-01-15