

基于行为模式的电子交易过程中信任评估研究*

王志兵, 李长云

(湖南工业大学 计算机与通信学院, 湖南 株洲 412008)

摘要: 在开放动态的网络环境中, 交易者的行为信任问题已经成为电子商务发展的重要障碍。提出了以交易过程中呈现出的行为模式为证据的信任评估方法, 该方法通过构建行为模式匹配级和可信属性为节点的贝叶斯网可信评估模型, 计算推导出电子交易过程中消费可信指数和销售可信指数的等级。在一个实际的电子交易平台验证了该方法的有效性。

关键词: 信任评估; 电子交易; 交易行为; 行为模式

中图分类号: TP393 **文献标志码:** A **文章编号:** 1001-3695(2010)04-1357-03

doi:10.3969/j.issn.1001-3695.2010.04.040

Research on credit assessment based on trade behavior for e-commerce

WANG Zhi-bing, LI Chang-yun

(School of Computer & Communication, Hunan University of Technology, Zhuzhou Hunan 412008, China)

Abstract: The open dynamic Internet has made traders' credit the significant obstacle to the development of E-commerce. This paper put forward that credit assessing could be based on the trade behavior. With this method, presented a Bayes model for assessment, composed of nodes of behavior pattern matching levels and nodes of credibility attributes, and the grade of the consumption credit index and that of the seller credit index could be inferred. The effectiveness of this method has been verified by an actual electronic trading platform.

Key words: credit assessment; e-commerce; trade behavior; behavior pattern

0 引言

伴随互联网的发展, 电子商务迅速崛起, 2009 年第一季度国内电子交易规模达到 467 亿元, 但仅 29.2% 的网民认可网上电子交易安全^[1]。如何在开放动态的网络中建立一套完善的信用评价模型, 全面客观地反映交易各方的信用表现, 提高网上交易的安全, 已变得极为关键^[2-4]。用户信任不仅包括对用户的身份信任, 也包括对用户的行为信任。传统的以授权和认证为基础的安全机制主要针对用户的身份信任问题, 对用户的行为信任问题并不是特别有效^[5]。例如, 在数字化电子资源的订购方面, 校园网用户可以合法下载文章, 但如果利用下载工具大量下载文章, 其行为可能是不可信的。目前, 淘宝、易趣等大多数的电子交易网站主要采用累加信任模型或平均值信任模型通过累加计算用户的信誉反馈评分或累加后除以评分次数得到用户的信用积分, 来形成用户的信用记录和信用度, 并最终确定用户的信用等级。用户某一时刻的信用总积分为所有交易得分的代数和或平均值, 且不同的信用积分对应不同的信用等级^[6]。这种方式存在评价过于主观、信用积分只与交易次数有关、难以处理诋毁及合谋欺诈等问题^[7]。

针对电子交易中的行为信任问题, 本文提出了一种以交易过程中呈现出的行为模式为证据的可信评估方法。行为模式

是软件可观察行为中可能反复出现的具有某些特征的事件序列, 是软件可信性在软件运行态中的一种外在指标。通过构建以行为模式匹配级为节点的贝叶斯网可信评估模型, 计算由监测服务获得的实际交易行为与由专家经验设定的行为模式的匹配度及匹配级, 以模式匹配级为条件, 通过贝叶斯网的概率计算, 推导出电子交易过程中消费可信指数和销售可信指数的等级。该方法融合了用户信任评估中的专家主观经验与实际交易行为, 有效改善了目前信任评估主观性过强、信任积分只与交易次数有关等不足。

1 基本思想

基于交易行为信息的可信性评估的基本思想是以买家、卖家在交易过程中呈现出的交易行为为证据进行买卖双方的信任评估。它的整体模型如图 1 所示。通过电子交易平台中的监测点, 行为监测服务监测、收集和组织软件在运行过程中的可观察行为; 参照根据领域专家经验建立的行为模式库, 对可观察行为进行模式匹配; 根据匹配结果和可信评估模型, 对可信属性进行评估。

行为监测服务的任务是收集软件在电子交易平台运行过程中的可观察行为。通过包装器方法、AOP 方法等技术监视和收集系统与外界的交互事件, 以及系统内部构件之间的交互

收稿日期: 2009-08-28; **修回日期:** 2009-10-23 **基金项目:** 国家自然科学基金资助项目(60773110); 中国博士后科学基金资助项目(20080440216); 湖南省自然科学基金资助项目(09JJ6087); 湖南省研究生创新基金资助项目(CX2009B200); 湖南省教育厅科研项目(09C325, 07C234)

作者简介: 王志兵(1974-), 男, 湖南邵东人, 实验师, 硕士研究生, 主要研究方向为可信软件(hnpuzb@163.com); 李长云(1971-), 男, 湖南耒阳人, 教授, 博士后, 主要研究方向为可信软件和软件动态演化。

事件。监测服务的另一个任务是如何组织好这些交互事件。组织的方法根据应用目的的不同,可以以任务为线索,也可以以事件发生源构件或事件目标构件为线索,最终形成一个事件序列。行为模式由应用领域专家根据以往的经验知识预先设计并使用模式语言描述,存放在行为模式库中。以监测服务收集和组织的交互事件序列为源,以模式库中的行为模式为参照,进行模式匹配,计算事件序列匹配到哪些行为模式及其匹配度和匹配级。可信属性评估模型是贝叶斯网模型,它的一部分节点为模式匹配级,另外一部分节点为用户可信属性。以模式匹配级为条件,通过贝叶斯网的概率计算,可得出可信属性各等级的概率,取概率最大的等级为可信属性的评估值^[8,9]。

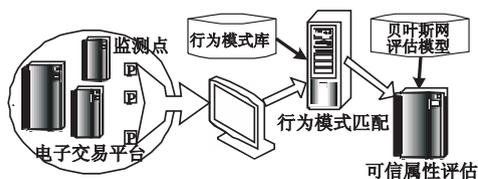


图1 基于行为模式的软件可信性评估过程

2 评估模型及可信值计算

2.1 设定行为模式

行为模式是软件可观察行为中可能反复出现的具有某些特征的事件序列,是软件可信性在软件运行态中的一种外在指标。行为模式主要由迹模式构成。所谓迹指的是事件序列片断,迹模式则是对迹的抽象和概括。为表示相似迹的共性和统计特性,迹模式使用以事件标志为字母的正则表达式来描述。例如,迹模式“(a * b)”表示以事件 a 开始,中间接若干任意事件并以事件 b 结束的迹,如“ab”“acdeab”和“afcb”等;迹模式“(ab) {2,6}”表示事件 a 和 b 交替出现 2 ~ 6 次的迹,如“ab-ab”“ababab”等。

行为模式由应用领域专家预先设计并使用模式语言描述,存放在行为模式库中。通过对大量电子交易过程的观察和分析,设定如下五种行为模式:

模式 a(Pa): 买家搜索商品(search) - 买家下订单(order) - 买家预付款 - 买家确认收货(accept), 迹模式表示为 (soi) {1,} * a。

模式 b(Pb): 买家下订单 - 卖家确认, 修改价格(modification) - 买家预付款 - 卖家发货(consignment) - 买家收货、付款(payment), 迹模式表示为 (o * m * i * c * p) {1,}。

模式 c(Pc): 卖家查看订单(examine) - 卖家修改价格 - 查看付款信息(inquire) - 发货, 迹模式表示为 em * qc。

模式 d(Pd): 买家多次(2 次以上)下订单不付款, 迹模式表示为 o {2,} * (p̂)。

模式 e(Pe): 买家多次(2 次以上)下订单、预付款, 而卖家不发货, 迹模式表示为 (op) {2,} * (ĉ)。

2.2 计算模式匹配度及模式匹配级

行为模式匹配是指以某个行为模式为参照, 搜索事件迹中符合该行为模式的所有不交叉的子迹。为表达事件迹匹配行为模式的程度, 引进匹配度的概念。对于指定的事件迹, 事件迹匹配某行为模式的所有子迹的事件数相对事件迹中事件总

数的比例, 称为事件迹在该模式下的匹配度。设事件迹为 T, 模式为 P, T 匹配 P 的子迹分别为 T₁, T₂, ..., T_k, 迹中事件个数使用函数 Eventsum 统计, 则 T 在 P 下的匹配度 M 的计算为

$$M = \frac{\sum_{i=1}^k \text{Eventsum}(T_i, *)}{\text{Eventsum}(T)} \quad (1)$$

行为模式匹配的目的之一就是计算匹配度, 算法总体过程如下: 首先查找并保存事件迹中符合行为模式的迹模式的所有不交叉的子迹; 然后按照式(1)计算匹配度。

模式匹配级是将软件可观察行为对行为模式的匹配度划分为若干个等级 N, 并将这些等级从高到低进行顺序编号为整型变量 i (i ≤ N), 它们所代表的匹配度区间范围从低到高的顺序分别是 [0, 1/N], [1/N, 2/N], ..., [(N-1)/N, 1], 匹配度落在哪一个范围, 匹配级就是哪一级。假设所有模式匹配级分为 4 级, 分别为完全匹配(匹配级为 1)、比较匹配(匹配级为 2)、基本匹配(匹配级为 3)和不匹配(匹配级为 4)。

2.3 建立评估模型

贝叶斯网络是目前不确定知识表达和推理领域最有效的理论之一。一个贝叶斯网络是由一些节点及连接这些节点的有向边构成的有向无环图。节点代表随机变量; 节点间的有向边代表了节点间的关系, 用条件概率表达关系强度。贝叶斯网络适用于表达和分析不确定性和概率性的事件, 可以从不完全、不精确或不确定的知识或信息中作出推理。因此, 本文使用贝叶斯网络来构建可信评估模型^[10,11]。

评估模型建立可信证据和可信属性间的映射关系。在可信评估的贝叶斯模型中, 节点变量有两种类型, 即模式匹配级和用户可信属性。设定行为模式后, 分析每一种行为模式和需评估的可信属性是否可能条件相关, 从而对每一个可信属性选择确定其评估证据的模式集合。然后对于每一个可信属性, 以可信属性为根节点, 以相关模式的匹配级作为子节点, 分别建立各个可信属性的贝叶斯网络模型的结构。本文对以上五种模式、消费可信指数及销售可信指数进行分析, Pa、Pb、Pd 与消费可信指数(Ca)相关, Pb、Pc、Pe 与销售可信指数(Cb)相关。贝叶斯网评估模型结构如图 2 所示。

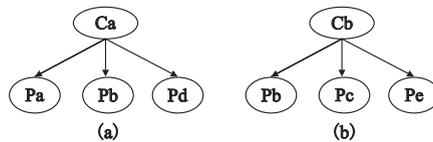


图2 信购网的贝叶斯网评估模型

2.4 计算可信值

单模式匹配级条件下可信属性的评估的贝叶斯计算公式为

$$p(Ca_i/Pa_j) = \frac{p(Pa_j/Ca_i)p(Ca_i)}{p(Pa_j)} \quad (2)$$

p(Ca_i) 表示属性 Ca 在等级 i 的先验概率; 如表 1 中, p(Ca₂) 为 0.4。

p(Pa_j) 表示模式 Pa 的匹配级 j 的先验概率; 如表 1 中, p(Pa₂) 为 0.3。

p(Pa_j/Ca_i) 表示可信属性根节点 Ca 在 i 这个等级的条件下模式 Pa 在 j 等级下的条件概率, 如表 1 中, p(Pa₂/Ca₂)

为 0.25。

在相同的模式匹配级条件下,通过贝叶斯网的概率可算,计算出可信属性各等级的概率,取概率最大的等级为该可信属性的评估等级。

3 实验及分析

本文方法在一个名叫信购网的 C2C 交易平台上得到初步验证。信购网使用 SSH 技术开发,主要包括系统管理、店铺管理、商品管理三大模块,具有对店铺中的商品进行管理,对商品进行分类、开店、租店、店铺审核、店铺信息管理等功能。通过基于动态 AOP 的监测器对系统的接口数据进行监测,过滤器对监测数据进行过滤后得到系统运行迹^[12]。

为确定贝叶斯网中的先验概率参数,在已知买家的消费可信指数和卖家的销售可信指数的前提下,使信购网在不同的典型场景下运行,由监测服务收集软件的可观察行为,计算出与可信属性相关模式的匹配度乃至匹配级。同时,对于每一次运行,由领域专家判定可信属性的等级。在信购网中,将可信属性信任等级也分为四级,分别为非常信任(信任等级为 1)、比较信任(信任等级为 2)、基本信任(信任等级为 3)和不信任(信任等级为 4)。

表 1 是监测的信购网中 10 次电子交易过程的迹描述及行为模式匹配级。

表 1 信购网中交易数据示例

交易过程 事件迹	Pa 匹 配级	Pb 匹 配级	Pc 匹 配级	Pd 匹 配级	Pe 匹 配级	消费可信 指数 Ca	销售可信 指数 Cb
soisoiemeiqcpa	1	3	3	4	4	1	3
oosoisoiasoaisoiaoo	2	4	4	4	4	1	3
sosoisoieamaoioiqcp	2	2	3	4	4	1	1
aiesoioomoiqooicoiai	2	4	3	4	4	2	3
saoiesoimoiqas	3	4	4	4	2	2	4
aioioipsoia	3	4	4	4	2	2	4
iaemsoioiaacs	3	4	2	4	4	2	2
aoemsoiaiqcsp	4	2	2	4	4	3	1
asooaemqoasocampaspas	4	4	3	2	4	4	3
osooosaaas	4	4	4	2	4	4	2

收集大量实验数据后,依据本文方法对消费可信指数及销售可信指数进行计算,结果如图 3 所示。

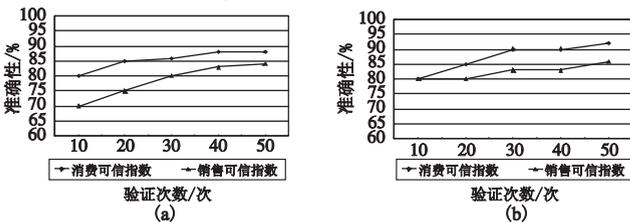


图 3 可信指数验证情况

图 3(a) 是样本数为 20 次时的验证情况,当验证次数为 10 次时,计算得到的消费可信指数和销售可信指数的准确性分别为 80% 和 70%,随着验证次数的增加,其可信指数的准确性也逐步提高,当验证次数达到 50 次时,其准确性分别达到 88% 和 84%;(b) 是样本数为 40 次时的验证情况,当验证次数为 10 次时,计算得到的消费可信指数和销售可信指数准确性都达到 80%,当验证次数达到 50 次时,其准确性分别达到 92% 和 86%。通过验证试验,说明本文方法是有效的。

4 结束语

相比已有工作,本文方法具备如下特点:

a) 主观性和客观性相结合,本文提出的方法是以电子交易过程中的行为信息为评估证据的,因此本质上是客观性的评估方法,但也具备主观性评估方法的特点,如在行为模式的设计、评估模型的构建过程中需要借鉴领域专家的经验知识。

b) 以可信属性和模式匹配级为节点的贝叶斯网评估模型,能够从不确定性的软件运行行为中推导出软件的可信性。评估模型既纳入了领域专家的经验知识,又包含了客观数据的分析结果,蕴涵了可信属性和行为模式间的条件相关关系。

本文方法的不足之处在于需要有相关的专家经验知识作为基础以及对无先验知识的行为难以准确处理。

进一步的研究工作有:依据多维可信属性的评估结果,对电子交易过程进行整体可信性综合评估;以行为模式的匹配和分析为基础,研究电子交易过程中买、卖双方行为预测的模型和方法。

参考文献:

- [1] 中国互联网信息中心. 第 24 次中国互联网络发展状况统计报告 [EB/OL]. (2009-07-16). <http://www.cnnic.cn/uploadfiles/pdf/2009/7/16/125126.pdf>.
- [2] 王怀民,唐扬斌,尹刚,等. 互联网软件的可信机理[J]. 中国科学 E 辑,2006,36(10):1156-1169.
- [3] 王远,吕建,徐锋,等. 一个适用于网购软件的信任度量及演化模型[J]. 软件学报,2006,17(4):682-690.
- [4] 黄海生,王汝传. 基于隶属云理论的主观信任评估模型研究[J]. 通信学报,2008,29(4):13-19.
- [5] 田立勤,林闯. 可信网络中一种基于行为信任预测的博弈控制机制[J]. 计算机学报,2007,30(11):1930-1938.
- [6] LI Qin-fang, LIU Zhong-ying. Research on Chinese C2C e-business institutional trust mechanism: case study on taobao and ebay (cn) [C]//Proc of International Conference on Wireless Communications, Networking and Mobile Computing. 2007: 3787-3790.
- [7] JØSANG A, ISMAIL R, BOYD C. A survey of trust and reputation systems for online service provision [J]. Decision Support Systems, 2007, 43(2):618-644.
- [8] MAN Jun-feng, WEN Zhi-cheng, LI Chang-yun, et al. Research on running time behavior analyzing and trend predicting of modern distributed software [J]. Journal of Computers, 2009, 4(8):747-754.
- [9] MAN Jun-feng, WEN Xiang-bing, WANG Zhi-bing, et al. Distributed software interactive behavior analysis based on knowledge fusion [C]//Proc of International Symposium on Information Processing. 2009:441-444.
- [10] FRIEDMAN N, LINIAL M, NACHMAN I, et al. Using Bayesian networks to analyze expression data [J]. Journal of Computational Biology, 2000,7(3):601-620.
- [11] WANG Yao, VASSILEVA J. Bayesian network-based trust model in peer-to-peer networks [C]//Proc of Agents and Peer-to-Peer Computing Conference. Heidelberg:Springer, 2005:23-34.
- [12] 万灿军,李长云. 动态演化环境中可信软件行为监控研究与进展 [J]. 计算机应用研究,2009,26(4):1201-1204.