基于相似度的信任推荐模型

董晓华[1] 周彦晖[2]
(重庆大学经济与工商管理学院 重庆 400030)[1] (西南大学计算机与信息科学学院 重庆 400715)[2]

摘 要 针对信任推荐过程中难以鉴别恶意推荐的问题,提出一种基于相似性的信任推荐模型。模型中,借鉴社会心理学的研究成果,若两个用户在行为上具有较高的相似性,则表明他们更容易相信对方。把用户评分相似性作为信任推荐时的权重系数,理论分析和仿真结果表明:该模型可有效防范信任推荐中的恶意推荐行为,降低信任度的计算误差,从而提高信任评估的准确性。

关键词 信任推荐,相似性,信任模型
中图法分类号 TP301 **文献标识码** A

Similarity-based Trust Recommended Model

DONG Xiao-hua[1] ZHOU Yan-hui[2]
(School of Economics and Business Administration, Chongqing University, Chongqing 400030, China)[1]
(College of Computer and Information Science, Southwest University, Chongqing 400715, China)[2]

Abstract It is difficult to distinguish the cheating and other malicious behaviours in the trust recommendation. A similarity-based trust recommended model was proposed. With social psychology research, two users are more likely to trust each other, when they have higher similarity in behaviours. Using similarity as trust recommendation weight, theoretical analyses and simulation results show that the model can effectively avoid malicious recommendation behaviour in trust recommendation. The model can reduce the calculation error of the trust remarkably compared with the existing model, and the preciseness of the trust evaluation model is enhanced greatly.

Keywords Trust recommended, Similarity, Trust model

1 引言

在开放网络环境下,电子商务交易的匿名性、随机性和动态性特征,使得交易双方失去了传统商务环境下的信任基础。交易双方往往互不认识,这样滋生了身份欺诈、虚假信息发布、拒绝或延期交货、质量和售后服务差等现象[1],从而为电子商务的健康发展造成了较大的影响。因此,电子商务的信任机制成为了业内人士研究的热点之一。

近年来,国内外学者针对不同的应用环境,在信任机制和信任度量方面进行了大量的研究。总体来说,目前国内外的研究主要涉及4个方向:基于策略和凭证的信任、基于声誉的信任、通用信任模型、Web和信息资源中的信任[2-4]。信任管理的概念最初由M. Blaze等人于1996年提出[5];文献[6]中按照建模方法,又分为基于社会学理论、基于统计、基于概率、基于语义以及基于不确定理论和模糊集等信任模型。根据Donovan Artz[3]等的观点,信任获取的途径主要有两种:基于策略及凭证的信任和基于声誉的信任。基于策略及凭证的信任模型包括基于公钥体系的信任机制和基于凭证的信任管理系统,其具有基于统一的信任管理语言、统一的授权决策引擎(CCA)及分布凭证管理等特点。目前典型的应用有PolicyMaker[5]、KeyNote[7]、REFEREE[8]等。基于声誉的信任管理是基于本地经验和其它实体的反馈(包括了交互历史),使用精确的数值定义信任等级,然后对实体进行信任等级划分,信任程度与活动相关,随着活动不断修正。基于声誉的信任管理在P2P、Ad-Hoc、Sensor Networks和普适计算等领域都有相应研究。目前典型的有F. Azzedin和M. Maheswaran的直接信任和推荐信任模型[9-11],Beth[12]、Abual-Rahman[13]、Stanford的EigenRep[14]、Zhang[15]、Yuan[16]、Matthew Richardson[17]和Christian Bizer[18]等的信任评估模型。Li[19]等人结合安全凭证和反馈信息,提出了一种基于多代理的解决策略。Dong[20]根据博弈理论,提出了基于赔偿的信任评估模型。文献[21]中提出了一种全局的信任评估模型GTM-CPCR,模型中综合了凭证、策略、服务能力和声誉。

在上述的研究中,信任推荐是信任评估的核心内容之一,但这些信任评估模型都没有很好地解决信任推荐的权重问题。目前主要的度量方式是根据推荐者自身的声誉高低来表示其推荐的可靠性,从而作为推荐的权值。但在实际应用环境中,存在以下两方面的问题:

到稿日期:2012-12-03 返修日期:2013-03-10 本文受国家自然科学基金项目(71272086),中央高校基本科研业务项目(CDJSK100076),国家固态酿造工程技术研究中心项目(2011k162280),重庆市教委人文社科项目(08jwsk141)资助。
董晓华(1972-),男,博士,副教授,主要研究方向为电子商务、信任计算,E-mail:dxh.cn@163.com;周彦晖(1972-),男,副教授,主要研究方向为信息系统安全、大数据与统计分析。

①共盟用户声誉虚高。在欺诈团伙中，共盟成员为了互相提高自己的声誉度而相互进行好评。

②声誉的计算复杂性及时间衰减性。声誉的高低由多种因素共同决定，具有时间衰减性，并受到交易情景（上下文相关）的影响，因此在某个应用情景中，声誉能否真实代表可信度也具有不确定性。

社会心理学的研究成果以及现有研究表明，信任与用户行为相似性之间存在联系，如果两个用户在行为上具有较高的相似性，则表明他们更容易相信对方[22]。因此在基于声誉的信任研究中，把相似性计算用于信任推荐已成为目前研究的热点[23]。本文把相似性计算用于信任推荐，把用户相似性作为信任推荐的权重系数。实验结果表明，该方法在一定程度上防范声誉诋毁的欺诈行为，可有效降低声誉计算误差，提高信任评估的准确性。本文第 2 节介绍相关理论研究；第 3 节介绍基于相似度的信任推荐模型，侧重介绍了直接信任推荐模型和间接信任推荐模型；第 4 节通过实验对模型进行有效性验证；最后对本文工作进行了总结，并提出了下一步工作计划。

2 相关理论研究

相似性的研究主要有两个方面：

①用户兴趣相关性：在推荐系统中，如果两个用户共同评价过的项目越多，就说明用户的关注点和兴趣点越相似，这种相似性体现了用户兴趣之间的关系。由于类信任只关心了评价项目而未考虑具体的评分，因此不能真实地反映两者的关系。用户兴趣相关性虽然在基于推荐的系统中很少被使用，但对推荐是非常重要的，在基于信任的推荐方法中常被作为用户之间的基础关系之一。

②评分相似性（用户相似性）：用户对共同评价过的项目的评分可能不同，说明用户对资源项本身的质量或能产生效果的评论不同，如果两个用户之间的评分相近，说明他们具有较高的相似性。用户相似性是推荐系统中最常用的用户关系。

本文主要就用户相似性进行讨论。

定义 1 相似度 $\mathrm{sim}(v_i, v_j)$ 是刻画用户 v_i 和 v_j 的评分行为的相似程度。v_j 和 v_i 的相似程度越高，说明 v_j 和 v_i 对应用环境中其他共同交易伙伴的看法越一致。

相似度可以由多种函数刻画[22]，相似度计算方法主要有余弦相似度函数和皮尔森相关系数法。本文仅介绍这两种方法。

2.1 余弦相似度函数

余弦相似度（Cosine Similarity）计算，是指由向量 x 和 y 之间的夹角的余弦进行计算的。当夹角越小时，相似性越高；而当夹角越大时，则相似性越小，如图 1 所示。

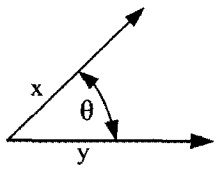

图 1 余弦相似性图示

定义 2 余弦相似度 *Sim_Cos* 的计算方式定义如下：

$$Sim\_Cos(v_i, v_j) = \frac{\sum_{k\in D'}(f_{i,k} \cdot f_{j,k})}{(\sum_{k\in D'} f_{i,k}^2 \cdot \sum_{k\in D'} f_{j,k}^2)^{\frac{1}{2}}} \tag{1}$$

式中，v_i, v_j 表示用户，$Sim\_Cos(v_i, v_j)$ 表示余弦相似度，$D_i = \{f_{i,1}, f_{i,2}, \cdots, f_{i,m}\}$ 表示用户 v_i 的评分项集合，$D_j = \{f_{j,1}, f_{j,2}, \cdots, f_{j,n}\}$ 表示用户 v_j 的评分项集合，$D' = D_i \cap D_j$ 表示 v_i 和 v_j 评分项的交集，$k \in D'$ 表示用户 v_i, v_j 的共同交易伙伴。

2.2 皮尔森相关系数法

在计算用户相似度时，有两种情况：

①两个用户的评分向量相似，若同时好评或同时差评，则两个用户的相似度高。此类相关性我们称为正相关。

②在两个用户的相同评分向量中，若两个用户的评分相反，如一个好评时另外一个差评，那这两用户也同样具有较高的相似性。此类相关性我们称之为负相关。

余弦相似度函数可以解决正相关问题，而不能解决负相关问题。皮尔森相关系数法，也称为线性相关系数（linear correlation coefficient），可以很好地解决正相关性和负相关性问题。

定义 3 皮尔森相关系数 r 可表示为：

$$r = Sim\_pr(v_i, v_j) = \frac{\sum_{k\in D'}(f_{i,k} - \overline{f_i}) \cdot (f_{j,k} - \overline{f_j})}{(\sum_{k\in D'}(f_{i,k} - \overline{f_i})^2 \cdot \sum_{k\in D'}(f_{j,k} - \overline{f_j})^2)^{\frac{1}{2}}} \tag{2}$$

式中，f, D' 定义同式(1)，$\overline{f_i}, \overline{f_j}$ 分别表示用户 i, j 的评分平均值。相关系数 $r \in [-1, 1]$，r 的绝对值越大，表示用户之间的相关程度越高。r 为正表示正相关，为负表示负相关，若 $r = 0$，表明两个变量间不是线性相关，但有可能是其他方式的相关（比如曲线方式）。

3 基于相似度的信任推荐模型

在交易关系 s, x 中，用户 x 在对交易伙伴 s 进行信任度评估。信任度评估权重系数目前常采用评估者（如 x）的信任度作为权重。为了便于分析比较，用 c_r_pr 表示信任度并作为权重系数，基于权重系数 c_r_pr 的信任度推荐对信任度的影响用 τ_t_pr 表示。使用评估双方的评分相似性（用户相似性，即交易用户同时对其他用户的信任度评估行为）作为推荐用户的信任度权重 c_r，用 c_r_cos 表示，其对信任度的影响用 τ_t_cos 表示。表 1 中列出了两种权重系数对信任度影响的关系。

表 1 相似度计算方法与声誉度的关系

s	x	c_r_cos	c_r_pr	τ_t_cos	τ_t_pr
诚信	诚信	高	高（正相关）	++	++
诚信	欺诈	低	高（负相关）	−	−−
欺诈	诚信	低	高（负相关）	−	−−
共盟	共盟	高	高（正相关）	++	++
非共盟	非共盟	高	高（正相关）	−−	−−

注：+少量增加，++大量增加；−少量减少，−−大量减少。

表中，共盟是指两个属于恶意用户共盟成员，非共盟是指两个都是恶意用户，但不是共盟成员。相似度方法对信任度计算的影响效果分析如下：

①交易双方 s, x 同为诚信用户。两者对共同交易过的用户都会给予诚信信任度评估，因此两者的信任度评分相近（一个好评，另外一个差评的可能性不大），这样在信任度评分时

两者相似度高，从而信任度权重大。因为是诚信评估，在相同的评分状况下 s 信任度值增加更快，这也表示对诚信行为进行了奖励。余弦相似度函数 Sim_Cos 和皮尔森相关系数法 Sim_r 这两种相似性计算方法都能满足此种需求。

②交易伙伴 s 为诚信用户，x 为恶意用户。两者对共同评估过的行为存在差异，由于恶意用户对诚信用户的评价分数会较低，因此为了不影响诚信用户的信任度，需对恶意用户的评估给予较低的信任度权值，此时余弦相关满足这种需求。因皮尔森相关系统具有负相关性，所以其系数的绝对值较高。

③交易伙伴 s 为恶意用户，x 为诚信用户。诚信用户会给予恶意用户较低的信任度评分，此时余弦相似性低，皮尔森相关系数高，因此后者作为权重比前者减少得更多。

④如果两者均是恶意用户，不管同盟与否，两者的相似性都较高。如果为共盟用户，则信任度增加得更快，如果为非共盟用户，则声誉度减少得更快。

⑤若用户以前为诚信用户，现在转化为恶意用户，则不管是哪种情况，信任度的减少会变得更快；相应地，如果以前为恶意用户，现在转换为诚信用户，不管什么情况，信任度的增加都会变得很慢。因此，此方法还具有奖励长期诚信行为、惩罚恶意行为的特点。

在实际应用中，我们希望对于 s 的信任度计算，不管是诚信的还是恶意的，越接近真实情况越好。余弦相似对诚信用户保护较好（恶意用户评分时，声誉减少得少），皮尔森相关系数法对恶意用户惩罚效果更佳（诚实用户对欺诈用户评价时其声誉度下降得更快）。两者尽管对恶意同盟用户没有较好的防范措施，但对声誉欺诈（即上述情况 5）具有较好的效果。因基于余弦相似性和基于皮尔森相关系数法的定义及使用方法类似，所以下面仅介绍基于余弦相似性的信任度推荐。

在式(1)的基础上，基于余弦的相似性信任度推荐可定义如下：

定义 4 设 $N(s)$、$N(x)$ 为时域 $[t-1,t]$ 内 s,x 的交易伙伴，$N'=N(s)\cap N(x)$ 表示 s,x 的共同交易伙伴，则交易用户 s,x 基于余弦相似性的评分用户 x 的信任度权重 c_r_cos 可表示为：

$$c_r\_cos=Sim\_Cos(x,s)=\frac{\sum_{k\in N'}\frac{\sum_{i=1}^{|C|}f_{c_i}(x,k)\cdot f_{c_i}(s,k)}{(\sum_{i=1}^{|C|}f_{c_i}^2(x,k)\cdot\sum_{i=1}^{|C|}f_{c_i}^2(s,k))^{\frac{1}{2}}}}{|N'|} \tag{3}$$

式中，$k\in N'$ 表示用户 s,x 的共同交易伙伴；$f_{c_i}(s,k)$，$f_{c_i}(x,k)$ 表示 s,x 对共同交易伙伴 k 在关键因素 c_i 上的评分。相似性越强，则评分用户 x 的信任度 $\tau_t(x)$ 的权重 $C_r(\tau_t(x))$ 也越大。

4 仿真实验及结果分析

为了描述方便，我们借鉴文献[21]中的信任评估模型 RCM-MKF。根据 RCM-MKF 中信任度权重 $c_r[\tau_{t-1}(x)]$ 的计算方法不同，将其分为基于用户信任度的计算方法（简称为 RCM-MKF_UT）和基于余弦相似度的计算方法（简称 RCM-MKF_Cos）。

下面在相同的实验环境[21]和相同的参数设置下，对 RCM-MKF_UT、RCM-MKF_Cos 模型进行对比实验。实验结果如图 2—图 5 所示。

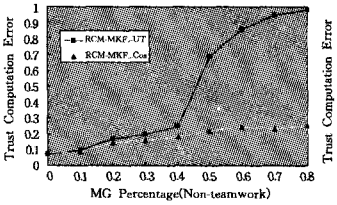

图 2 无共盟稳定欺诈时的信任计算误差

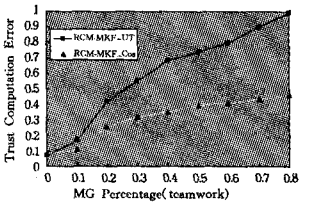

图 3 共盟稳定欺诈时的计算误差

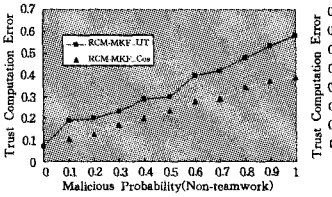

图 4 无共盟波动欺诈时的计算误差图

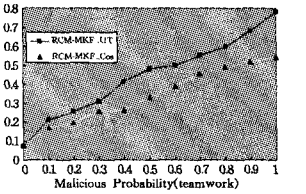

图 5 共盟波动欺诈时的计算误差

实验结果表明：

①在稳定欺诈情况下，无共盟比共盟用户的相似度低，在声誉榨取时，无共盟用户增加的声誉较少，从而无共盟用户的信任误差较低。

②在波动欺诈情况下，因无共盟用户之间未相互抬升声誉评分，所以信任误差率比共盟情况低。因欺诈的波动性降低了相似度，所以对声誉计算的调整效果降低，从而比稳定欺诈的信任误差要高。

③波动欺诈时的相似度比稳定欺诈时的相似度低，因此模型不能很好地区分其恶意行为，所以波动欺诈时的信任计算误差率比稳定欺诈时的误差率高。

④基于相似性的信任度权重 $c_r[\tau_{t-1}(x)]$ 计算方法的准确性比基于信任度的权重高了很多，从而验证了基于相似度信任推荐模型的有效性及第 3 节中对模型分析的正确性。

综上所述，本文提出的基于相似度的声誉评估模型对可靠性信任的评估是有效的。

结束语 在信任推荐中，如何有效防范恶意推荐是提高信任推荐真实性和准确性的核心研究内容。目前采用的以用户信任度作为推荐权重系数的方法不能有效解决此问题。因此，本文提出了基于相似性的信任推荐，并就此方法及对信任推荐的影响进行了分析。分析表明，基于相似度的信任推荐可有效降低信任度的计算误差，从而提高了信任评估的准确性，最后通过实验对基于相似度的信任推荐模型进行了有效性验证。如何把相似性计算应用于间接推荐，是下一步需要继续进行的研究工作。

参 考 文 献

[1] 甘早斌，曾灿，等. 电子商务下的信任网络构造与优化[J]. 计算机学报，2012，35(1)：27-37

[2] 马礼，郑纬民. 网格环境下的信任机制研究综述[J]. 小型微型机系统，2008，29(5)：825-830

[3] Artz D，Gil Y. A Survey of Trust in Computer Science and the Semantic Web[J]. Journal of Web Semantics，2007，5(2)：58-71

（下转第 158 页）

然后，将时间分区的概念具体化，加入阶段集，形成具体的改进 RBAC 模型。该模型将传统的“角色-权限”二元关系改进为“角色-阶段-权限”三元关系，增强了 RBAC 在实际运用中的灵活性。由于改进模型与传统模型是相容的，因此可同时使用两种权限控制方法进行系统的权限分配，从而大大提高权限管理的使用效率。

该改进模型的有效性[9]在实际项目中得到了验证，但是其通用性还有待进一步的研究。

参 考 文 献

[1] Xing Tian-yang, Cao Min. Research and application of algorithm for generating authority-tree based on TP-RBAC model [J]. Computer Engineering and Design, 2010, 31(5): 950-953

[2] 钟华，冯玉琳，姜洪安. 扩充角色层次关系模型及其应用[J]. 软件学报，2000，11(6)：779-784

[3] 信科，杨峰，杨光旭，等. 基于 RBAC 权限管理系统的优化设计与实现[J]. 计算机技术与发展，2011，21(7)：172-174

[4] Sandhu R, Coyne E J, Feinstein H, et al. Role-based access control models [J]. IEEE Computer, 1996, 29(2): 38-47

[5] Zhou Wei, Meinel C. Team and task based RBAC access control model[C]// Network Operations and Management Symposium, 2007, LANOMS 2007. Latin American, IEEE, 2007: 84-94

[6] Ferraiolo D, Kuhn R. Role-Based Access Controls [C]// Proceedings of the 15th NIST-NCSC National Computer Security Conference. 1992: 554-563

[7] Yu Su, Wang Yin, Hua Kun. The research of information security based on RBAC with SOD [J]. International Journal of Advancements in Computing Technology, 2012, 4(14): 482-490

[8] 杨彩侠，王小慧，曹旻. OF_RBAC 权限控制模型的研究及应用[C]// Proceedings of 2010 International Conference on Management Science and Engineering. 2010: 65-69

[9] 董理君，胜生，杜敏，等. 一种基于环境安全的角色访问控制模型研究[J]. 计算机科学，2009，6(1)：1-54

(上接第 134 页)

[4] 曲向丽. 网格环境下互信机制关键技术研究[D]. 长沙：国防科学技术大学计算机学院，2008

[5] Blaze M, Feigenbaum J, Lacy J. Decentralized trust management [C]// Dale J, Dinolt G, eds. Proceedings of the 17th Symposium on Security and Privacy. Oakland, CA: IEEE Computer Society Press, 1996: 164-173

[6] 朱艳春，刘鲁，张巍. 在线声誉系统中的信任模型构建研究[J]. 控制与决策，2007，22(4)：413-417

[7] Blaze M, Feigenbaum J, Keromytis A D. Keynote: Trust management for public-key infrastructures [C] // Christianson B, Crispo B, William S, et al. , eds. Cambridge 1998 Security Protocols International Workshop. Berlin: Springer-Verglag, 1999: 59-63

[8] Chu Y H, Feigenbaum J, Lamacchia B. REFEREE: trust management for Web applications[J]. WordWideWeb Journal, 1997, 2(2): 127-139

[9] Azzedin F, Maheswaran M. Evolving and Managing Trust in Grid Computing Systems[C]// Proceedings of the IEEE Canadian Conference on Electrical&Computer Engineering. 2002: 1424-1429

[10] Azzedin F, Maheswaran M. Towards Trust-Aware Resource Management in Grid Computing Systems [C] // Proceedings of the 2nd IEEE/ACM International Symposium on Cluster Computing and the Grid. 2002: 452-452

[11] Azzedin F, Maheswaran M. A Trust Brokering System and Its Application to Resource Management in Pubic-Resource Grids [C]// Proceedings of the 18th International Parallel and Distributed Processing Symposium. 2004: 289-298

[12] Beth T, Borcherding M, Klein B. Valuation of trust in open system[C] // Collmann D, ed. Computer Security, ESORICS'94. volume 875 of Lecture Notes in Computer Science, Berlin: Springer Verlag, 1994: 3-18

[13] Abdul-Rahman A, Hailers S. A distributed trust model[C]// Proceeding of the 1997 New Security Paradigms Workshop. Cumbia, UK: ACM Press, 1997: 48-60

[14] Kamvar S D, Schlosser M T. EigenRep: Reputation management in P2P networks[C] // Lawrence S, ed. Proc. of the 12th Int'l World Wide Web Conf. Budapest: ACM Press, 2003: 123-134

[15] Zhang Q, Zhang X, Wen X Z, et al. Construction of peer-to-peer multiple-grain trust model[J]. Journal of Software, 2006, 17(1): 96-107

[16] Yuan L L, Zeng G S, Jiang L L, et al. Dynamic Level Scheduling Based on Trust Model in Grid Computing[J]. Chinese Journal of Computers, 2006(7): 1217-1224

[17] Richardson M, Agrawal R, Domingos P. Trust management for the semantic web[C]// Proceedings of the Second International Semantic Web Conference. 2003: 351-368

[18] Christian B, Radoslaw O. Using context-and content based trust policies on the semantic web[C]// Proceedings of the 13th international World Wide Web Conference on Alternate track papers & Posters. 2004: 228-239

[19] 李海华，杜小勇，田萱. 一种能力属性增强的 Web 服务信任评估模型[J]. 计算机学报，2008，31(8)：1471-1477

[20] 董晓华，吴中福. 网格服务信任的赔偿评估模型[J]. 重庆大学学报，2010，33(6)：121-127

[21] 董晓华. 网格服务的信任机制研究[D]. 重庆：重庆大学，2010

[22] Ziegler C N, Golbeck J. Investigating interactions of trust and interest similarity[J]. Decision Support Systems, 2007, 43(2): 460-475

[23] 李景涛，荆一楠，肖晓春，等. 基于相似度加权推荐的 P2P 环境下的信任模型[J]. 软件学报，2007，18(1)：157-167

[24] 袁传思. 基于用户信任的攻击检测防御模型[J]. 重庆理工大学学报：自然科学版，2010，24(6)：72-77