

# 电子商务中的一种潜在信任关系预测方法

马 霄 甘早斌 鲁宏伟 马 尧

(华中科技大学计算机科学与技术学院 武汉 430074)

**摘 要** 信任关系在用户寻找可靠信息方面扮演了重要角色。已有信任关系预测方法主要基于信任的传递性和用户间的相似度。然而,在电子商务应用中,不同声誉度的用户对于商品的评价会对其他用户的购买行为产生不同的影响,用户声誉度的差异性会在较大程度上影响用户间建立信任关系的可能性。因此,针对电子商务应用,给出了用户信任关系子网络 and 用户商品评价关系子网络的形式化描述。根据社会学理论,提出了一种基于用户相似度和全局声誉度的潜在信任关系预测方法,旨在挖掘电子商务应用中陌生用户间潜在的信任与不信任关系,为用户辨别评价信息的可信性,进而选择可信商品提供辅助决策。基于 Epinions 数据集的对比实验结果表明,该方法在信任关系预测的准确度方面有较好的表现。

**关键词** 电子商务,信任关系,用户相似度,全局声誉度

**中图法分类号** TP301 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2014.12.029

## Prediction of Latent Trust Relationships in E-commerce

MA Xiao GAN Zao-bin LU Hong-wei MA Yao

(Department of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China)

**Abstract** Trust relationships play an important role in helping users collect reliable information. Existing trust relationships prediction methods are mainly based on the trust transitivity and user similarity. However, in e-commerce systems, reviews from users with different reputations may have different impacts on other users' purchasing behaviors. Different user reputations may have different impacts on trust relationships prediction. Therefore, this paper formally described a user-trust relationships sub-network and a user-product-review relationships sub-network in e-commerce systems. According to sociological theory, a trust relationships prediction method was proposed based on user similarity and global reputation, which aims at discovering the latent trust relationships between unfamiliar users in e-commerce systems and helping users distinguish the reliability of rating information in order to choose reliable products. Comparative experiments on the trust relationships prediction were performed on Epinions dataset. The experimental results show that the proposed method has better trust prediction accuracy.

**Keywords** E-commerce, Trust relationships, User similarity, Global reputation

## 1 引言

随着 Internet 技术的迅速发展及推广普及,电子商务应用领域不断扩大,网购用户数和交易量呈现爆炸式增长<sup>[1]</sup>。电子商务网站允许用户对所购买商品的优劣进行评论,以帮助其他消费者做出正确的决策。然而,在线商品评价存在评价信息过量、评价信息良莠不齐、可信性较低等问题。针对这些问题,一些在线社会网络允许用户建立自己的信任网络(Web of Trust),用于记录哪些用户是可信的,哪些用户是不可信的。例如,当前最大的商品评价网站 Epinions<sup>[2]</sup>允许用户对他人的评价内容打分,同时,允许用户明确表示自己对于他人的信任与不信任的态度。然而,由于信任网络的稀疏性、

不对称性等特征<sup>[3]</sup>,大多数用户之间并没有明确的信任与不信任关系,如果能够利用网络中已经存在的其他相关信息挖掘出网络中用户间可能存在的潜在的信任关系,则可以帮助用户辨别评价信息的可信性,进而选择可信的商品。

已有的信任关系预测方法往往只考虑由信任的传递性<sup>[4-6]</sup>或用户间的相似度所产生的信任度,用户间的相似度包括用户间偏好的相似度<sup>[7,8]</sup>和用户间评价的相似度<sup>[9,10]</sup>。然而,这些方法并不足以解决电子商务应用中的信任关系预测问题。在电子商务应用中,存在着用户、商品和评价 3 种元素,不同声誉度的用户对于一件商品的评价会对其他用户的购买行为产生不同的影响。用户是否会采纳其他用户的评价取决于用户与评价者之间的信任关系。在挖掘潜在信任关系

到稿日期:2013-12-21 返修日期:2014-03-15 本文受国家自然科学基金资助项目(61272406),中央高校基本科研业务费资助项目(HUST:2013TS101)资助。

马 霄(1989-),女,博士生,CCF 学生会员,主要研究方向为推荐系统、信任计算,E-mail:cindyma@hust.edu.cn;甘早斌(1967-),男,副教授,主要研究方向为电子商务、信任计算;鲁宏伟(1964-),男,教授,主要研究方向为信任计算、网络安全;马 尧(1986-),男,博士,主要研究方向为信任计算。

时,不能仅考虑由用户间相似度产生的信任度,还要考虑用户间全局声誉度的差异性,声誉度低的用户倾向于信任声誉度高的用户,声誉度高的用户倾向于不信任声誉度低的用户,这充分体现了信任关系的不对称性。基于以上分析,根据社会学理论,综合考虑电子商务应用中由用户相似度和全局声誉度的差异性产生的信任度,提出一种潜在信任关系预测算法(PALTR算法),以为用户辨别评价信息的可信性进而选择可信商品提供辅助决策。

本文第2节介绍了已有相关工作的研究进展;第3节对问题进行了详细的描述;第4节提出潜在信任关系预测方法;第5节基于Epinions数据集验证该方法的准确性与有效性;最后,总结全文并对下一步工作进行展望。

## 2 相关工作

近年来,信任关系吸引了越来越多研究者的关注,信任关系在用户寻找可靠信息方面扮演了重要角色。针对信任关系预测问题,国内外众多学者尝试了许多不同的方法。文献[5]利用信任的传递性来预测用户间的信任关系并构建不同上下文中的信任网络。文献[4]首次讨论了信任与不信任的传递问题,将信任矩阵与不信任矩阵区分开来,通过对这些矩阵的操作来获得两节点间的传递信任度。文献[11]提出了Tidal-Trust信任预测算法,通过深度优先遍历信任网络中源节点到其他用户,根据这些用户间的聚合评价得出两个用户间的信任度。文献[7]验证了信任关系中确实存在同质性:相似的用户倾向于建立信任关系,而相互信任的用户倾向于表现出更多的相似性。文献[8]利用社会学理论来研究在线信任演化,并提出了一种信任演化模型—eTrust。在在线商品评价的背景下,通过研究用户偏好的动态性来研究在线信任的发展演化规律。文献[9]从理论和经验两方面分析了信任和评价之间的双向影响,给出了一种在线电子商务网络的理论模型,并提出了一种潜在的在线社交网络挖掘方法。文献[10]指出在通常情况下,即使随着时间推移互相信任的用户间在商品的评价方面也会有较小的不同,从而提出了一种改进的矩阵因子分解模型来预测用户间的信任关系强度。文献[12]引入了商品流行度概念,用二分图来代表网络中用户与商品之间的联系,利用社会关系二分图的拓扑和结构信息预测网络中潜在的信任关系,而并不考虑社会关系图中节点和边的属性。在文献[13]中,作者用一个有向符号图来表示一个信任或不信任网络,文中利用边 $(u_1, u_2, +)$ 来表示它们之间的信任关系,用边 $(u_1, u_2, -)$ 来表示它们之间的不信任关系,利用结构平衡理论来预测任意给定边的符号。

以上信任关系预测方法仅仅考虑了由信任的传递性或用户间的相似度产生的信任度,并没有考虑电子商务应用中用户间声誉度的差异性对用户间建立潜在信任关系的影响。信任关系预测问题实质上是一种特殊的链接预测问题,信任预测的结果有许多实际的应用,比如推荐系统。文献[14-16]的研究结果表明,基于信任的推荐系统能够有效提高推荐系统的准确度和可信度。因此,研究电子商务应用的潜在信任关系预测问题具有重要的理论意义和实用价值。

## 3 问题描述

电子商务应用中主要包含3种元素:用户、商品和评价,

图1表示的是这3种元素之间关联关系的一个示例。

这里,用 $U$ 代表电子商务应用中的用户集合, $u_i$ 代表第 $i$ 个用户, $U = \{u_1, u_2, u_3, u_4, u_5\}$ , $I$ 代表电子商务应用中的商品集合,某件商品可用 $I_i$ 表示, $I = \{I_1, I_2, I_3, I_4\}$ 。如图1所示,电子商务应用中包含两种子网络:

1) 用户信任关系子网络 $G_t = (V_t, E_t)$ ,其中 $V_t = U$ ,即每一个节点代表一个用户, $E_t$ 代表用户间存在明确的信任与不信任关系,且 $E_t \subseteq U \times U$ ;

2) 用户商品评价关系子网络 $G_r = (V_r, E_r)$ ,其中 $V_r = U \cup I$ ,即每一个节点代表一个用户或一件商品, $E_r$ 代表存在用户对商品的评价,且 $E_r \subseteq U \times I$ 。

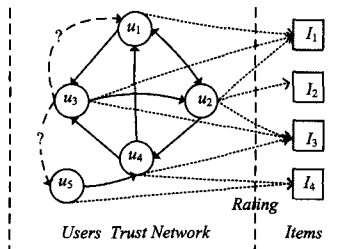


图1 用户信任关系与用户商品评价关系示意图

每一个用户对于一件商品的评价都会对其他用户的购买行为产生影响,用户是否会采纳其他用户的评价取决于该用户和评价者之间的信任关系。图1中,用户 $u_5$ 和 $u_3$ 之间并没有明确的信任关系,当用户 $u_3$ 看到用户商品评价关系子网络中 $u_5$ 提供的商品评价时,用户 $u_3$ 是否应该信任 $u_5$ 所作的评价,本文将这一问题归结为节点之间的潜在信任关系预测问题。由于信任网络中信任关系的不对称性、稀疏性等特征,大多数用户间可能没有直接明确的信任与不信任关系,因此必须找到有效的方法来挖掘潜在的信任关系,为用户辨别评价信息的可信性进而选择可信商品提供辅助决策。

## 4 潜在信任关系预测

现实生活中,人们更倾向于信任那些与自己有相同特征或共同偏好的人。在社会学中,这一现象被称作同质性<sup>[17]</sup>,受同质性影响,相似的用户间倾向于建立信任关系<sup>[7]</sup>。比如说在年龄、兴趣、学历、家乡等方面有相同特点的人们更容易成为朋友并互相信任。因此,用户相似度是本文进行信任关系预测时需要考虑的第一个要素。

假设节点 $u_i, u_j$ 是用户商品评价关系子网络 $G_r$ 中的任意两个用户节点, $u_i, u_j \in U$ 。借助杰卡德系数<sup>[18]</sup>给出由用户间相似度产生的信任度 $T_s(u_i, u_j)$ 的计算公式:

$$T_s(u_i, u_j) = \frac{|N_{u_i} \cap N_{u_j}|}{|N_{u_i} \cup N_{u_j}|} \quad (1)$$

即在用户商品评价关系子网络中,由相似度产生的节点 $u_i$ 与 $u_j$ 的信任度 $T_s(u_i, u_j)$ 等于这两个节点共同邻居节点集的总数与其全体邻居节点集总数的比, $T_s(u_i, u_j) \in (0, 1)$ 。

$N_{u_i}, N_{u_j}$ 分别表示节点 $u_i$ 和节点 $u_j$ 在用户商品评价关系子网络 $G_r = (V_r, E_r)$ 中的邻居节点集。由于在该网络中 $E_r \subseteq U \times I$ ,即任意用户节点 $u_i \in U$ 与集合 $U$ 中其他节点没有直接关联关系,所以,节点 $u_i$ 的邻居节点集 $N_{u_i}$ 等于与节点 $u_i$ 直接关联的集合 $I$ 中的所有元素 $I_i$ 直接关联的集合 $U$ 中的除节点 $u_i$ 以外的节点集。其实际意义即为与节点 $u_i$ 有相同评价记录的节点集。在图1中,各节点的邻居节点集如下所示:

$$\begin{aligned} N_{u_1} &= \{u_2, u_3\}; \\ N_{u_2} &= \{u_1, u_3, u_4\}; \\ N_{u_3} &= \{u_1, u_2, u_4\}; \\ N_{u_4} &= \{u_2, u_3, u_5\}; \\ N_{u_5} &= \{u_4\}. \end{aligned}$$

虽然相似性在信任关系预测中发挥了重要作用,但并不全面推断用户间的信任关系,还需要考虑第二个重要因素,即用户的全局声誉度,充分考虑其他用户对该用户的综合评价。

当用户间声誉度的差异较大时,由于信任的不对称性,用户  $u_3$  对  $u_5$  和用户  $u_5$  对  $u_3$  建立信任关系的可能性是不同的。声誉度低的用户倾向于信任声誉度高的用户,反之,倾向于不信任。为了定量描述这一现象,下面将给出电子商务应用中用户全局声誉度的计算方法,并将其引入由用户间全局声誉度差异性产生的信任度的计算公式中。

任意节点  $u_i$  的全局声誉度  $R_{u_i}$  是综合节点  $u_j$  的所有邻居节点对  $u_j$  的评价而得出的  $u_j$  的可信程度。

如果用  $r_{u_i}^j$  表示节点  $u_i$  的第  $j$  个邻居节点  $u_j$  对它的评价,那么节点  $u_i$  的全局声誉度  $R_{u_i}$  的定义如下:

$$R_{u_i} = \frac{\sum r_{u_i}^j}{m}, i, j=1, 2, \dots, m \quad (2)$$

式中,如果节点  $u_i$  的第  $j$  个邻居节点  $u_j$  信任节点  $u_j$ ,则  $r_{u_i}^j$  的取值为 1,否则为 -1; $m$  表示节点  $u_j$  的邻居节点总个数,所以,  $R_{u_i} \in [-1, 1]$ ;如果  $R_{u_i}, R_{u_j} \in [0, 0.5]$  或  $(0.5, 1]$ ,则称节点  $u_i$  和节点  $u_j$  处于相同的声誉度区间;否则,称节点  $u_i$  和节点  $u_j$  处于不同的声誉度区间。

考虑 3 种不同的节点声誉度分布情形:

- 1)  $R_{u_i} \in [-1, 1], R_{u_j} < 0$ ;
- 2)  $R_{u_i} = 0.3, R_{u_j} = 0.4, R_{u_i} - R_{u_j} = -0.1, R_{u_j} - R_{u_i} = 0.1$ ;  
 $R_{u_i} = 0.8, R_{u_j} = 0.9, R_{u_i} - R_{u_j} = -0.1, R_{u_j} - R_{u_i} = 0.1$ ;
- 3)  $R_{u_i} = -0.8, R_{u_j} = 0.9, R_{u_i} - R_{u_j} = -1.7, R_{u_j} - R_{u_i} = 1.7$ 。

如果节点  $u_i$  和节点  $u_j$  的全局声誉度如 1) 所述,无论节点  $u_i$  的全局声誉度取值范围是什么,只要节点  $u_j$  的全局声誉度取值为负值,就认为节点  $u_i$  与节点  $u_j$  建立信任关系的可能性为 0;

如果节点  $u_i$  和节点  $u_j$  的全局声誉度如 2) 所述,即节点  $u_i$  和节点  $u_j$  处于相同的声誉度区间,那么,用  $R_{u_i} * R_{u_j}$  来表示两节点间建立信任关系的可能性;

如果节点  $u_i$  和节点  $u_j$  的全局声誉度如 3) 所述,即节点  $u_i$  和节点  $u_j$  处于不同的声誉度区间,那么,由前文分析可知,两节点间声誉度差异性越大,两节点间建立信任关系的可能性越小,反之,可能性越大。此时,用  $f(R_{u_i} - R_{u_j}) = \theta e^{-(R_{u_i} - R_{u_j})}$  来表示两节点间建立信任关系的可能性,实验过程中将  $\theta$  取值为 2/15。

综合以上 3 种不同情形,给出由用户间全局声誉度差异性产生的信任度  $T_R(u_i, u_j)$  的计算公式:

$$T_R(u_i, u_j) = \begin{cases} 0, & R_{u_j} < 0 \\ R_{u_i} * R_{u_j}, & R_{u_i}, R_{u_j} \in [0, 0.5] \text{ 或 } R_{u_i}, R_{u_j} \in (0.5, 1] \\ \theta e^{-(R_{u_i} - R_{u_j})}, & \text{其他} \end{cases} \quad (3)$$

由以上分析可知,电子商务应用中任意两节点间的潜在

信任关系受到用户商品评价关系子网络和信任关系子网络中用户相似度和全局声誉度的共同影响。因此,给出用户间潜在信任关系预测值计算公式:

$$T_P(u_i, u_j) = \alpha T_S(u_i, u_j) + \beta T_R(u_i, u_j) \quad (4)$$

其中,  $\alpha$  和  $\beta$  是信任关系调节因子,  $\alpha + \beta = 1$ ,且  $\alpha$  和  $\beta$  分别控制着节点相似度和全局声誉度对信任关系预测方法的影响程度,  $\alpha$  和  $\beta$  的确定方法将在第 5 节进行详细讨论。

综上所述,电子商务应用中任意两用户  $u_i$  和  $u_j$  的潜在信任关系预测算法(Prediction Algorithm of Latent Trust Relationships, PALTR)可用图 2 来描述。

输入:  $G_r = (V_r, E_r), G_t = (V_t, E_t)$

输出:  $T_p$

```

for all target users  $u_i$  in  $V_r$  do
  for all target users  $u_j$  in  $V_t$  do
    find the neighbor set  $N_{u_i}, N_{u_j}$ 
    calculate  $T_S(u_i, u_j)$ 
  end for
end for
for all target users  $u_i$  in  $V_t$  do
  for all target users  $u_j$  in  $V_t$  do
    calculate  $R_{u_i}, R_{u_j}$ 
    if  $R_{u_j} < 0$ 
      then  $T_R(u_i, u_j) = 0$ 
    else if  $R_{u_i}, R_{u_j} \in [0, 0.5]$  或  $R_{u_i}, R_{u_j} \in (0.5, 1]$ 
      then  $T_R(u_i, u_j) = R_{u_i} * R_{u_j}$ 
    else  $T_R(u_i, u_j) = \theta e^{-(R_{u_i} - R_{u_j})}$ 
    end if
  end for
end for
for all target users  $u_i$  in  $U$  do
  for all target users  $u_j$  in  $U$  do
     $T_P(u_i, u_j) = \alpha T_S(u_i, u_j) + \beta T_R(u_i, u_j)$ 
  end for
end for

```

图 2 PALTR 算法

本算法中主要的步骤是:

首先,在用户商品评价关系子网络  $G_r = (V_r, E_r)$  中分别获得用户  $u_i$  和  $u_j$  的邻居节点集  $N_{u_i}$  和  $N_{u_j}$ ,利用式(1)计算出用户  $u_i$  和  $u_j$  由相似度产生的信任度  $T_S(u_i, u_j)$ 。由于该计算过程中需要嵌套遍历所有节点,因此,由用户间相似度产生的信任度的时间复杂度为  $O(N^2)$ 。

然后,在用户信任关系子网络  $G_t = (V_t, E_t)$  中计算出用户  $u_i$  和  $u_j$  的全局声誉度,并根据  $R_{u_i}$  和  $R_{u_j}$  的分布情况,利用式(3)计算出由用户间全局声誉度的差异性产生的信任度  $T_R(u_i, u_j)$ 。由于该计算过程中需要嵌套遍历所有节点,因此,此部分时间复杂度也为  $O(N^2)$ 。

最后,将上述计算得出的由用户相似度和全局声誉度差异性产生的信任度代入用户间潜在信任关系计算式(4)即可获得用户  $u_i$  和  $u_j$  之间的潜在信任关系预测值。因此, PALTR 算法的时间复杂度为  $O(N^2)$ 。

## 5 实验及结果分析

### 5.1 实验需求与数据集

为了验证本文提出的信任关系预测方法的准确性与合理

性,选择在线社会网络中真实数据集来进行实验验证。根据第4节算法,在实验过程中需要提供以下两类实验数据:

1)用户信任关系数据。在该数据中应该包含用户间明确的信任与不信任评价的数据。这些数据将被作为实验部分预测潜在信任关系的训练集和测试集。

2)用户商品评价关系数据。在该数据中应该包含有用户对不同商品做出评价的数据。这些数据将用于计算用户间相似度。

鉴于以上两点要求,本文选择在线社会网络中的数据集 Epinions。Epinions.com 是一种在线社会网络,用户可以对网站中任意一种商品发表评论,与其他用户分享自己的购物体验。另外,Epinions.com 也提供明确的“Web of Trust”服务,如果用户信任其他用户的评论,那么他可以把该用户标记为自己信任网络中的用户,如果不信任,可以明确地把自己不信任的用户标记出来。Epinions 数据集的特点满足本次实验的需求。

### 5.2 实验设置

在 Windows 平台下,使用 matlab 软件编程实现 PALTR 算法并比较验证其有效性。本次实验中,从 Epinions 数据集中随机选取 1000 个商品节点及其对应的商品评价者节点。将选取的数据作为两个网络子图,用户信任关系子图  $G_t = (V_t, E_t)$  和用户商品评价关系子图  $G_r = (V_r, E_r)$ ,如图 1 所示。在实验过程中,将实验分为两个阶段进行,即训练和验证阶段以及比较实验阶段。根据经典的训练与验证方法,将取样数据分为两部分,80%用于对本文提出的信任关系预测方法进行训练,20%用于对实验结果进行检验。然后,将实验结果与同类型的预测方法中的实验结果进行比较。

### 5.3 实验验证与对比

实验验证阶段主要用于验证本文提出的信任关系预测方法的准确度,以及优化其中的参数。确定参数后,我们可以明确得知在式(4)中,哪一种因素对信任关系预测的贡献较大。在最初开始实验时,我们假设参数  $\alpha$  和  $\beta$  的取值均为 0.5,即假设相似度和全局声誉度对信任关系预测方法的贡献相当。将选取的训练集部分数据用于预测,并将预测结果与测试集数据进行比较,然后不断重复这一过程,优化式(4)中信任关系,调节因子  $\alpha, \beta$  的取值,以提高预测的准确度。经过实验验证发现,当  $\alpha$  取值为 0.25,  $\beta$  取值为 0.75 时,预测准确度最高。由于 Epinions 数据集中只提供了信任与不信任即 1 和 -1 两个离散数值,并没有给出任何有关信任与不信任关系程度的表示,比如,如果预测出的两节点间的信任值为 0.7,在 Epinions 数据集中并没有这种表示程度的信任关系值,因此我们在实验中引入阈值  $TH$  来衡量计算得出的数值属于信任关系或者不信任关系。如果计算出的信任关系值大于该阈值,则认为两节点间的关系是信任关系,否则为不信任关系。由于式(4)中所得  $T_P(u_i, u_j)$  的取值范围为(0,1),因此实验中我们将阈值  $TH$  取值为 0.5。参数选择如表 1 所列,从表中可以看出,全局声誉度因素对信任关系预测的贡献较大。

表 1 参数选取

TH	$\alpha$	$\beta$	$\theta$
0.5	0.25	0.75	2/15

在实验比较阶段,选择与本文研究内容相似的文献[12]中的信任关系预测方法进行比较,该方法也考虑了用户、商品

与评价之间的关系,并引入了用户相似度和商品流行度两种因素来预测网络中潜在的信任关系。

在实验结果比较过程中,主要利用以下几种指标来评判实验结果的准确度:

1)平均绝对误差(Mean Absolute Error, MAE)

$$MAE = \frac{\sum_{(i,j)} |T_{i,j} - P_{i,j}|}{n} \quad (5)$$

式中, $n$ 代表信任关系的总个数; $T_{i,j}$ 代表用户间真实的信任关系评价; $P_{i,j}$ 代表预测出的用户间的信任关系值。

2)为了便于与文献[12]进行比较,利用文献[12]中定义的两个实验验证指标,即信任关系预测准确度(Trust Prediction Rate, TPR)和不信任关系预测准确度(Distrust Prediction Rate, DPR)。

$$TPR = \frac{TP}{TP + FN} \quad (6)$$

$$DPR = \frac{TN}{TN + FP} \quad (7)$$

式(6)、式(7)中, $TP$ (True Positive)表示计算所得两节点间的关系是信任关系,而在相应数据集中也为信任关系的边的总个数; $FP$ (False Positive)表示计算所得两节点间的关系是信任关系,在相应数据集中为不信任关系的边的总个数; $TN$ (True Negative)表示计算所得两节点间的关系是不信任关系,在相应数据集中为不信任关系的边的总个数; $FN$ (False Negative)表示计算所得两节点间的关系是不信任关系,在相应数据集中为信任关系的边的总个数。

从实验比较结果可以看出,当采用相同的 1000 个随机样本进行实验时,我们的方法相比于文献[12]有较小的平均绝对误差以及较高的信任关系预测准确度,如表 2 所列。

表 2 准确度评价指标比较

	MAE	TPR	DPR
PALTR 算法	0.3501	0.8424	0.6703
文献[12]	0.4917	0.7628	0.6735

文献[12]综合利用了用户相似度和商品流行度,但他们只是利用用户-商品评价关系构成的二部图网络中节点的出度和入度等结构信息来预测信任关系,没有考虑节点和边的实际属性信息,这势必会降低预测结果的准确度。而本文提出的方法,综合考虑了用户信任关系子网络 and 用户商品评价关系子网络,充分利用了这两种子网络中节点和边的属性信息,从而在一定程度上提高了预测的准确度。同时,文献[12]在预测信任关系时将信任关系作为无向关系进行处理,而真实数据集中,信任关系是有向的、不对称的,这也是本文所提的信任关系预测方法预测结果准确度较高的另一个重要原因。

**结束语** 本文讨论了影响电子商务应用中陌生用户间建立信任关系的两种因素,给出用户相似度和全局声誉度对其信任关系影响的计算公式,进而提出一种基于用户相似度和全局声誉度的潜在信任关系预测算法(PALTR 算法)。对比实验结果表明,本文提出的潜在信任关系预测方法在信任关系预测的准确度方面有更好的表现。

由于应用背景和信任关系的复杂性,本文仅考虑了离散的二值信任关系,信任关系的细化及阈值的选取有待进一步研究。另一方面,将集中研究如何将基于用户相似度和全局声誉度的信任关系预测方法应用于基于信任的商品推荐系

统,以为用户辨别评价信息的可信性进而选择可信商品提供辅助决策。

## 参 考 文 献

- [1] 阳振坤,张清,王勇,等. 大数据的魔力[J]. 中国计算机学会通讯,2012,8(6):17-21
- [2] Epinions[OL]. <http://www.epinions.com>
- [3] Yuan W, Guan D, Lee Y K, et al. The small-world trust network [J]. Applied Intelligence, 2011, 35(3): 399-410
- [4] Guha R, Kumar R, Raghavan P, et al. Propagation of trust and distrust[C]// Proceedings of the 13th International Conference on World Wide Web. New York, USA, 2004: 403-412
- [5] Liu X. Towards Context-Aware Social Recommendation via Trust Networks[M]// Web Information Systems Engineering-WISE 2013. Springer Berlin Heidelberg, 2013: 121-134
- [6] Ma Y, Lu H W, Gan Z B. Trust Inference Path Search with Minimum Uncertainty for E-commerce [C]// Proceedings of the 10th Web Information Systems and Applications Conference. Yangzhou, China, 2013: 133-137
- [7] Tang J, Gao H, Hu X, et al. Exploiting homophily effect for trust prediction [C]// Proceedings of the 6th ACM International Conference on Web Search and Data Mining. Rome, Italy, 2013: 53-62
- [8] Tang J, Gao H, Liu H, et al. eTrust: Understanding trust evolution in an online world [C]// Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Beijing, China, 2012: 253-261
- [9] Matsuo Y, Yamamoto H. Community gravity: measuring bidirectional effects by trust and rating on online social networks [C]// Proceedings of the 18th International Conference on World Wide Web. Madrid, Spain, 2009: 751-760
- [10] Au C, Iwata T. Strength of social influence in trust networks in product review sites [C]// Proceedings of the 4th ACM International Conference on Web Search and Data Mining. HongKong, 2011: 495-504
- [11] Ziegler C N, Golbeck J. Investigating interactions of trust and interest similarity [J]. Decision Support Systems, 2007, 43(2): 460-475
- [12] O'doherty D, Jouili S, Van P. Towards trust inference from bipartite social networks [C]// Proceedings of the 2nd ACM SIGMOD Workshop on Databases and Social Networks. Scottsdale, AZ, USA, 2012: 13-18
- [13] Leskovec J, Huttenlocher D, Kleinberg J. Predicting positive and negative links in online social networks[C]// Proceedings of the 19th International Conference on World Wide Web. Raleigh, North Carolina, USA, 2010: 641-650
- [14] Ma H, Zhou T C, Lyu M R, et al. Improving recommender systems by incorporating social contextual information[J]. ACM Transactions on Information Systems (TOIS), 2011, 29(2): 9
- [15] O'doherty J, Smyth B. Trust in recommender systems [C]// Proceedings of the 10th International Conference on Intelligent User Interfaces. San Diego, CA, USA, 2005: 167-174
- [16] Yuan W, Lee S, Han Y, et al. Efficient routing on finding recommenders for trust-aware recommender systems [C]// Proceedings of the 6th International Conference on Ubiquitous Information Management and Communication. Kuala Lumpur, Malaysia, 2012: 29
- [17] Wasserman S, Faust K. Social network analysis: Methods and applications [M]. Cambridge university press, 1994
- [18] Liben D, Kleinberg J. The link-prediction problem for social networks [J]. Journal of the American Society for Information Science and Technology, 2007, 58(7): 1019-1031
- 
- (上接第 137 页)
- [9] Moschitti A. A study on convolution kernels for shallow semantic parsing [C]// Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics, 2004
- [10] Moschitti A. Making Tree Kernels Practical for Natural Language Learning [C]// EACL. 2006
- [11] Joachims T. Making large scale SVM learning practical [OL]. <http://www.academic.research.microsoft.com/paper/1361548>
- [12] Wiegand M, Klakow D. Convolution kernels for opinion holder extraction [C]// The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Los Angeles, California; Human Language Technologies. Association for Computational Linguistics, 2010
- [13] Qiu G, Liu B, Bu J, et al. Opinion word expansion and target extraction through double propagation [J]. Comput. Linguist., 2011, 37(1): 9-27
- [14] Kim Y, Kim S, Myaeng S. Extracting Topic-related Opinions and their Targets in NTCIR-7 [C]// Proc. of NTCIR-7 Workshop. 2008
- [15] Agarwal A, Xie B, Vovsha I, et al. Sentiment analysis of Twitter data [C]// Portland, Oregon; Proceedings of the Workshop on Languages in Social Media. Association for Computational Linguistics, 2011
- [16] Wilson T, Wiebe J, Hoffmann P. Recognizing contextual polarity in phrase-level sentiment analysis [C]// Vancouver, British Columbia, Canada; Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2005
- [17] Liu B, Hu M, Cheng J. Opinion observer: analyzing and comparing opinions on the web [C]// Proceedings of the 14th international conference on World Wide Web. ACM, 2005
- [18] Collins M, Duffy N. New ranking algorithms for parsing and tagging; Kernels over discrete structures, and the voted perceptron [C]// Proc. of the ACL 2002. 2002: 28-136
- [19] 赵妍妍, 秦兵, 车万翔, 等. 基于句法路径的情感评价单元识别 [J]. 软件学报, 2011(5)
- [20] 王倩, 何婷婷, 闻彬, 等. 基于依存关系的中文情感要素抽取技术研究 [C]// 第十届全国计算语言学学术会议. 烟台, 2009
- [21] 黄亿华, 濮小佳, 袁春风, 等. 基于句法树结构的情感评价单元抽取算法 [J]. 计算机应用研究, 2011, 28(9): 3229-3234
- [22] 姜德成, 姚天坊. 汉语句子语义极性分析和观点抽取方法的研究 [J]. 计算机应用, 2006, 26(11): 2622-2655
- [23] 黄晨, 钱龙华, 周国栋, 等. 基于卷积核的无指导中文实体关系抽取研究 [J]. 中文信息学报, 2010, 24(4): 11-17