

社交网络服务中一种基于用户上下文的信任度计算方法

乔秀全 杨 春 李晓峰 陈俊亮

(北京邮电大学网络与交换技术国家重点实验室 北京 100876)

摘 要 电子商务中的信任问题一直备受关注. 随着近些年社交网络的流行, 其真实的社交网络及信任关系开始受到关注, 基于社交网络开展各种电子商务活动逐渐成为一种新的发展趋势. 因此, 研究用户间的信任度对基于社交网络的电子商务具有重要意义. 文中通过借鉴社会心理学中人与人之间的信任产生原理, 提出了社交网络中基于用户上下文的信任度计算方法. 该方法将社交网络中用户之间的信任度分为熟悉性产生的信任度和相似性产生的信任度; 同时又根据所起作用的重要程度不同, 把相似性划分为内部相似性和外部相似性, 并给出了信任度计算的具体方法. 最后, 仿真实验验证了该方法的合理性和有效性.

关键词 社交网络服务; 用户上下文; 人际关系; 信任度

中图法分类号 TP393

DOI 号: 10.3724/SP.J.1016.2011.02403

A Trust Calculating Algorithm Based on Social Networking Service Users' Context

QIAO Xiu-Quan YANG Chun LI Xiao-Feng CHEN Jun-Liang

(State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876)

Abstract Trust is a significant problem in e-commerce. As the social networking sites are becoming more and more popular in recent years, its real social network and trust relationship attract people's attentions. The e-commerce based on the social networking sites has become a new trend. Therefore, the research on the users' trust for e-commerce based on SNS is very significant. By learning the trust generation principle in the psychology, this paper presents a trust calculating algorithm according to the SNS users' context. This method divides the user trust among the different SNS users into two parts; one is generated by familiarity and the other one is generated by similarity. At the same time, according to the importance of the role, the similarity part is divided into internal similarity and external similarity. And also the specific trust calculation method proposed by this paper is illustrated. Finally, the simulation results verify that the trust calculating algorithm based on the SNS users' context is reasonable and effective.

Keywords social networking services; user context; interpersonal relationship; trust degree

1 引 言

随着计算机与网络技术的普及与发展, 电子商

务在中国迅速崛起, 如今已具有相当的规模. 但是信任仍然是电子商务中的一个未解难题, 并且备受重视^[1-2]. 买家对于只能通过图片和文字描述来了解商品的购物形式存在一定的怀疑和不信任. 互联网研

收稿日期: 2011-07-28; 最终修改稿收到日期: 2011-10-21. 本课题得到国家自然科学基金(60802034)、国家“九七三”重点基础研究发展规划项目基金(2012CB315802)、新一代宽带无线移动通信网科技重大专项(2011ZX03002-002-01)、北京市科技新星计划项目(2008B50)资助. 乔秀全, 男, 1978年生, 博士, 副教授, 中国计算机学会(CCF)会员, 主要研究方向为网络服务智能化理论及技术. E-mail: qiaoxq@bupt.edu.cn. 杨 春, 男, 1985年生, 硕士研究生, 主要研究方向为移动社交网络服务. 李晓峰, 女, 1950年生, 教授, 主要研究领域为网络智能与通信软件. 陈俊亮, 男, 1933年生, 教授, 博士生导师, 中国科学院院士, 中国工程院院士, 主要研究领域为服务计算、网络服务智能化理论及技术.

究咨询机构艾瑞调研数据显示,有 87.9%的消费者在购物之前,会到网络上寻找他人的体验评价。有鉴于此,很多基于 Web 的电子商务网站推出了点评系统,让购买过某种商品的用户发表其对商品及购买过程的评论,以此增加其他用户对卖家及商品的信任。但是这些电子商务网站用户的注册信息不一定真实,并且用户之间缺少联系,相互来说只是一个个 ID 而已,虽然其评价对于欲购买商品的用户来说具有一定的参考价值,但还是很难得到其他用户的信任;此外,由于用户之间缺乏实际的联系,大部分针对某一商品数量繁多的评价都是按照发表的时间先后顺序排列,用户需要浏览全部或者大部分之后才能判断哪些评价可以信任或者具有实际的参考价值。因此,在电子商务中,如何增强用户之间的信任度,从而提升用户的消费体验,促进电子商务的健康发展具有重要意义^[3]。另一方面,近些年来基于 Web 的社会性网络服务(Social Networking Services, SNS)的数量和规模发展非常迅速^[4], SNS 要求用户以真实身份注册,维护着真实的人际关系,所以其人际关系的真实性具有非常大的利用价值,特别是对用户间的信任关系具有积极的意义。除了基本的信息分享和通信的功能之外, SNS 目前开始大量地向垂直行业应用渗透,如游戏、电子商务领域和即时通信领域等。研究表明,和系统的推荐相比,朋友的推荐成功率更高^[5-6]。而在现实生活中,熟人推荐是促成用户产生消费行为的一个非常重要的手段。而社交网络是现实中人际关系的真实反映,所以大型的电子商务网站若是能够结合社会性网络构建用户之间的真实人际关系,针对某一种商品利用现实社会中真实的人际关系来为用户挑选信任度高的评价,可以最大程度发挥 SNS 中人际信任关系的作用,还可以省去用户浏览和过滤大量冗余信息的时间和精力,并且用户所信任朋友的评价对其是否购买商品起到很大的决定作用,能够大大增加商品交易的成功率。因此,在 SNS 平台上开展电子商务,由于用户之间的朋友关系,更有利于口碑效应的产生和传播,规范商家行为、促进口碑产品的销售都具有更大的优势。基于社交网络开展电子商务,通过 SNS 中的真实社会关系及信任关系让“口碑营销”变成“行为营销”,用户之间信任度的衡量和计算非常重要。

目前国内外已经有一些电子商务网站开始建立 SNS 社区^[7],让用户同自己的好友(朋友、同事、同学、家人等)保持联系,及时了解他们的最新状况与

动态信息,获得更可靠的购物经验与建议,使得原本陌生的用户之间建立联系,产生信任。然而,现在的电子商务网站并没有充分利用人们之间的信任关系,熟人的评价也没有发挥出应有的导向作用。本文的目的就是在结合了 SNS 的电子商务或者点评系统中,对于某一个项目(该项目可以是一件商品、旅馆、旅游景点或是某酒店的一道菜)的所有评价,基于 SNS 中用户之间的真实人际关系,计算出用户对每个给出评价的用户的信任度,按照用户信任度的高低将这些评价呈现出来,供用户进行参考并做出自己的决定。

对于信任的研究, P2P 领域已经有很多优秀的算法和模型,如 Kamvar 等人^[8]提出的 EigenRep 信任模型、Song 等人^[9]提出的 FuzzyTrust 模型、陈海宝等人^[10]提出的基于关系强度的交互信任计算模型、李勇军等人^[11]归纳总结了信任模型和信任推理方法的最新研究成果。然而目前社会性网络中信任的研究还处于起步阶段,研究成果较少。Jennifer 等人^[12]提出利用 FOAF 在基于 Web 的社交网络中计算没有直接联系用户之间的信任度关系,但是其信任度仅有两种取值(0 表示不信任, 1 表示信任)使其很难用于现实的环境中; Wang 等人^[13]基于电子商务中卖家的交易记录并使用社交网络中两个指标 k -core 和 center weights 来预测出有交易风险的卖家,但是这种方法必须要求有交易记录存在,也不适用于计算普通用户之间的信任度; Bhuiyan 等人^[14]提出了在基于地理位置的社会性网络中结合信任与社会声誉来决策推荐的机制,该机制没有给出社会声誉的具体计算方法; 雷环等人^[15]提出结合声誉与主观逻辑来为用户推荐产品或服务,分析了 SNS 中恶意主体信任度的变化,其计算方法来源于 P2P 领域,是建立在 SNS 网络中存在有恶意行为的主体,并且其它主体可以对其恶意行为为评分的基础之上,并不适用于一般的 SNS 中用户间信任度计算的需求。以上提到几种 SNS 中信任度的计算方法都是根据作者的经验和自身对于人际关系信任的认识和理解,或是借鉴 P2P 网络中信任度的计算方式而来,基本上都是从技术的某个角度来考虑信任问题,缺乏相应的社会学和心理学方面的基础理论指导,难以全面体现现实社会中人们之间产生信任的过程,很难代表现实生活中用户之间真实的信任关系。而在社会学领域,已有大量关于人们之间信任问题的认识和研究,彭泗清^[16]认为人们之间的实际信任度是由初始信任度与人际交往所共同决定的。Zucker^[17]将信

任的产生机制划分为 3 种: (1) 声誉产生的信任: 通过其过去的行为以及声誉的了解给予一定的信任, 可以理解为对他人的熟悉性产生的信任; (2) 由社会相似性产生的信任: 根据其家庭背景、种族、价值观念的相似程度给予一定的信任; (3) 由法制产生的信任. 实际上, 社交网络是基于真实的人际关系来构建起来的社会关系网络, 社会学领域的这些研究理论和方法为社交网络中用户之间的信任度研究提供了很好的理论指导和借鉴的意义, 然而, 在目前的社交网络系统中尚缺乏具体的方法将二者紧密结合起来, 提供一套行之有效的适用于社交网络的用户信任度计算方法.

本文通过借鉴社会学和心理学中人与人之间的信任产生原理, 提出了社交网络中基于用户上下文的信任度计算方法. 该方法将信任度分为熟悉性产生的信任度和相似性产生的信任度; 同时又根据所起作用的重要程度不同, 把相似性划分为内部相似性和外部相似性, 并给出了信任度计算的具体方法.

本文第 2 节提出社交网络中基于用户上下文的信任度计算方法和计算流程, 并详细阐述计算过程; 第 3 节设计移动 SNS 中基于地理位置的兴趣点社区的点评推荐服务作为本文提出的基于用户上下文信任度计算方法的应用场景, 实验并分析本文所提出的信任度计算方法的合理性和有效性; 最后给出结论和未来的工作.

2 基于用户上下文的信任度计算方法

本文根据社会学与心理学中人们之间信任产生的过程以及在情景模拟实验^[18]的基础上, 将人们之间的信任度划分为熟悉性产生的信任度以及相似性产生的信任度两部分: 一般说来, 在社会关系网络中, 人们总是对越熟悉的人越信任, 所以可以说对于某一个用户越熟悉信任度就越大; 另外在熟悉性相同的朋友中, 人们常常倾向于更喜欢在态度、兴趣、价值观、背景和人格上和自己相似的人^[19], 因而对这部分用户的信任度更高, 若是来自这部分人的推荐, 用户往往更容易接受. 此外, 本文根据信任产生过程中所起的不同作用, 将相似性分为外部相似性(例如年龄、家庭所在地、专业等)和内部相似性(性格、兴趣爱好等): 在熟悉性相同且均为陌生的情况下, 外部相似性对于用户信任其他用户所起的作用更大, 例如对于一群刚认识的朋友, 人们总是和同龄人、老乡或者校友更谈得来, 这种情况下这些用户更

容易被信任; 在熟悉性相同且熟悉性较高的情况下, 内部相似性对于用户信任其他用户所起的作用更大, 比如在好友中(除亲人外), 人们经常保持联系的一般都是兴趣爱好相同的朋友, 此种情况下人们对这部分好友更为信任. 综上所述, 通过熟悉性以及相似性来计算用户之间的信任度, 更能体现出现实社会中信任产生的过程, 也更贴近真实世界中用户之间的信任关系.

本文在计算信任度的过程中引入了用户上下文, 通常来说, 上下文是指任何能够表征一个实体, 如人、位置或用户与应用交互相关的其它对象所处环境特征的信息. 现在随着移动智能终端能力的增强以及社会化网络不断发展壮大, 移动社交网络(Mobile SNS)可以获取的用户上下文相当丰富: 用户的位置信息(通过手机 GPS 或者移动网络定位获得)、用户的状态(Presence)信息(如在线、离线、开会等)、手机日历/日程信息、用户的偏好信息、用户的社会关系网络、好友的评价信息等. 本文用于计算用户间的信任度而引入的用户上下文包括: 用户之间的消息交互、用户的属性、用户的偏好信息、用户的社会关系网络、好友的评价信息等上下文信息. 本文的关键在于根据这些用户上下文计算用户对于某一个项目(Item)所有给出评价用户的信任度.

该方法根据用户间人际关系的熟悉性以及相似性来计算信任度, 得到用户 A (根用户)对于用户 N (欲计算用户 A 对其信任度的用户, 目标用户)的信任度计算公式:

$$tr(A, N) = Ftr(A, N) + Str(A, N) \quad (1)$$

式中, $tr(A, N)$ 为用户 A 对目标用户 N 的信任度, $Ftr(A, N)$ 为用户 A 对目标用户 N 由熟悉性产生的信任度, 在对熟悉性的研究中, 人们常常用交往频率或者曝光效应来控制熟悉性, 即一般情况下用户交往的次数或者见面的次数越多则越熟悉, 所以用户之间的交往频率越高, 两人越熟悉, 因而对彼此的信任度越高, 根据文献^[18]本文选取用户之间的交流次数作为上下文来计算用户之间熟悉性产生信任度; $Str(A, N)$ 为用户 A 对目标用户 N 由相似性产生的信任度, 这部分又分为外部相似性和内部相似性: 本文选取用户的属性(年龄、家庭所在地等)作为上下文来计算用户之间的外部相似度, 另外根据用户的偏好信息作为上下文来计算用户之间内部相似度, 最后根据熟悉性的不同, 将外部相似性和内部相似性调整权重计算出用户之间由相似性产生的信任度.

2.1 熟悉性产生的信任

图论中的图是由若干给定的点及连接两点的线所构成的图形,这种图形通常用来描述某些事物之间的某种特定关系,用点代表事物,用连接两点的线表示相应两个事物间具有这种关系,由于可以清晰地表示出事物之间的拓扑结构而引起广泛的注意,被越来越多地用于科学研究中.本文参考了一种基于图的计算用户推荐分数的算法^[20],进行了适当的改进,提出根据用户熟悉性产生的信任度计算方法如下:

考虑 SNS 网络可以表示为有向带权图 $G(N, E, W)$,其中 N 为网络中的节点集合,代表所有用户; E 为网络中节点之间的有向边,代表用户间的交流关系,节点之间有边则说明这两个用户之间存在交流关系,弧头指向的节点为消息的接收方; W 为边的权重,表示用户之间单向交流的消息数量,如图 1 所示.一般说来,用户之间交流越多则越熟悉;另一方面由于交流是双向的,若在交流中仅有一方发送消息,而另一方并无回复消息,则有可能是骚扰,这样也不能称这两个用户为熟悉,所以本文选取双方交互中最小消息数作为熟悉性的度量.将图 1 中两节点之间的两条边中权重较大的一边去掉,保留较小的一条边,得到用户间的最小消息数网络,如图 2 所示.

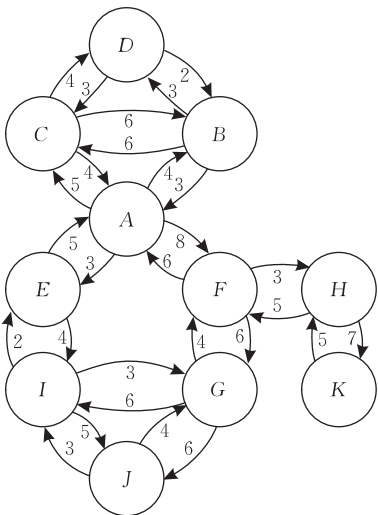


图 1 SNS 网络初始消息交互示意图

在图 2 中以用户 A 作为根节点为例,将与用户 A 有联系(直接以及间接)的所有用户重新排列形成以 A 为圆心的一个同心圆,如图 3 所示:第 1 层节点为用户 A 的直接朋友,第 2 层节点为用户 A 直接朋友的朋友,以此类推.根据六度空间理论(又称六度分割理论或小世界理论等),一个人和世界上任何

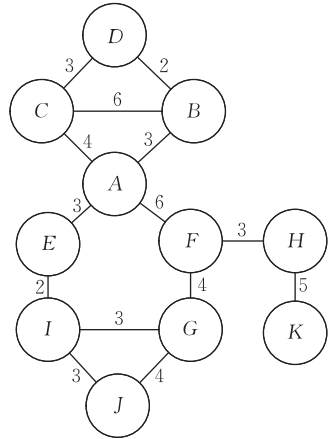


图 2 用户交互最小消息数网络

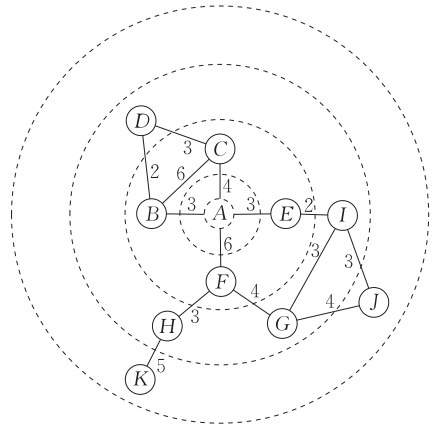


图 3 目标节点网络示意图

一个陌生人之间所间隔的人不会超过 6 个,也就是说,最多通过 6 个人就能够认识任何一个陌生人,所以世界上所有人都在以 A 为圆心的 7 层同心圆的某一层中.

但是由于现实的社交网络系统不能覆盖世界上所有人、联系某位用户必经的关键用户节点缺失等原因,造成有一些用户在同心圆中处于第 7 层之外的某一层,或者成为孤立节点,这种情况不可避免,所以我们对于超过 6 个人才能连接到的用户予以保留,根据以下方法继续计算根用户对目标用户由熟悉性产生的信任度;对于孤立节点则将根用户对其由熟悉性产生的信任度置为零.在图 3 中 B、C 同为 A 的朋友,D 为 B 和 C 的共同朋友,在计算由熟悉性产生的信任度时,由于 B 和 C 互为朋友,从 A 到 D 的路径有 4 条 $\{(A, B, D), (A, C, D), (A, B, C, D), (A, C, B, D)\}$,众所周知,对于间接认识的朋友,所经过的中间节点越少其信任度越高,所以本文对于图 3 中同层节点之间的相互联系删除,如图 4 所示.

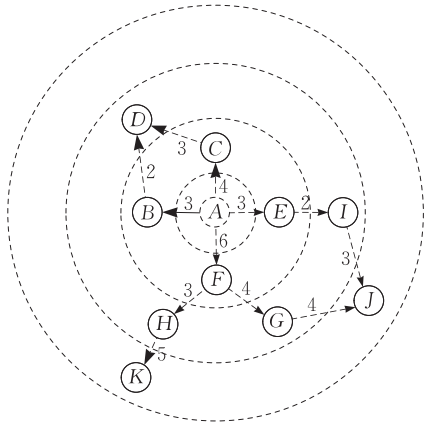


图 4 无同层联系的目标节点网络示意图

从图 4 可以根据式(2)计算得到用户 A 对于每个节点的信任度:

$$Ftr(A, N) = \omega_N \cdot \sum_{i=1}^n \left[\prod_{j=1}^m \frac{N(S_{j-1}, S_j)}{L_j} \right] \quad (2)$$

其中 A 为根节点, N 为除了根节点的其它目标节点; B、C、E、F 4 个节点处于第 1 层, 依次类推; 公式中 i 表示节点 A 到节点 N 的 n 条路径中的第 i 条, j 表示该路径中的第 j 层, m 为目标节点所处的层; L_j 为第 j 层的所有节点与第 $j-1$ 层中所有关联节点之间的消息数量之和, 例如 $L_1 = 16, L_2 = 14$; $N(S_{j-1}, S_j)$ 为节点 S_{j-1} 和 S_j 之间的消息数, S_j 是该路径上第 j 层的节点. 另外, 为了表示目标节点 N 相对于根节点 A 的远近层次关系, 利用 ω_N 作为目标节点 N 对于根节点 A 的层次权重, 离得 A 越近, 权重越高, 因此 ω_N 可以用 $\left(1 - \frac{layer_N}{layer_{SUM}}\right)$ 来计算, 其中 $layer_N$ 代表节点 N 所在的层数, $layer_{SUM}$ 代表层数的总和.

例如对于节点 B, A 对其由熟悉性产生的信任度为 $[3/(3+4+3+6)] \times [1-1/6]$; 而对于节点 D, A 可以通过两条路径 $\{(A, B, D), (A, C, D)\}$ 到达 D 节点, 所以 A 对 D 由熟悉性产生的信任度为 $\{[3/(3+4+3+6)] \times [2/(2+3+2+4+3)] + [4/(3+4+3+6)] \times [3/(2+3+2+4+3)]\} \times [1-2/6]$, 其它节点的信任度计算过程类似.

根据式(2)计算图 4 中用户 A 对所有节点由熟悉性产生信任度的结果如表 1 所示. 对于所有目标节点总是先挑选信任度最高的节点, 若是存在信任度值相同的目标节点, 则先选取由根节点到每个目标节点路径中所经过节点最少的一条, 若经过的节点数量相同, 则选取整条路径上各边权值之和最大的目标节点, 以此类推. 最终得到根节点对其它节点

按照由熟悉性产生信任度的节点顺序列表: F, C, B, E, G, D, H, J, K, I. 得到的排序结果与以上分析相符.

表 1 熟悉性产生的信任度结果举例

节点	信任度
B	0.1563
C	0.2083
D	0.0557
E	0.1563
F	0.3125
G	0.0714
H	0.0536
I	0.0119
J	0.0212
K	0.0167

2.2 相似性产生的信任

本文在熟悉性产生的信任度的基础上, 将目标用户分为熟悉和不熟悉两类, 对这两类用户分别计算其相似性产生的信任度, 用户的相似性又分为外部相似性和内部相似性, 其计算方法如下:

$$Str(A, N) = \alpha S_o(A, N) + (1-\alpha) S_i(A, N) \quad (3)$$

其中 $Str(A, N)$ 表示用户 A 对于用户 N 由相似性产生的信任度, $S_o(A, N)$ 为用户 A 与用户 N 的外部相似度, $S_i(A, N)$ 表示用户 A 与用户 N 的内部相似度, α 为调整系数, 当用户 A 与用户 N 为陌生人时, 外部相似度对于用户 A 决定是否信任用户 N 所起的作用更显著, 此时根据文献[21]中情景实验的结果取 $\alpha = 0.6107$; 当用户 A 与用户 N 为熟人时, 内部相似度对于用户 A 决定是否信任用户 N 所起的作用更显著, 此时 $\alpha = 0.3983$.

外部相似性的计算选取年龄、家庭所在地这两种属性, 鉴于目前还没有用户属性在相似度中所起作用的研究, 本文视这两种属性同等重要, 进而分配相同的权重, 得到外部相似性的计算式:

$$S_o(A, N) = S_a(A, N) + S_h(A, N) \quad (4)$$

式(4)中, $S_a(A, N)$ 为年龄相似性, 且 $S_a(A, N) \in [0, 1]$, 其值越高说明二者的年龄越相似, 计算方法如下:

$$\begin{cases} 1 - \frac{|Age(A) - Age(N)|}{10}, & |Age(A) - Age(N)| \leq 10 \\ 0, & |Age(A) - Age(N)| > 10 \end{cases} \quad (5)$$

$S_h(A, N)$ 为家庭所在地相似性, 一般说来, 用户对于其它用户在相同的熟悉性下, 家庭所在地相同并且所在地的行政区域单位越小则信任度越高,

具体计算方法如下:若用户 N 与用户 A 的家庭所在地位于同一小区则为 1,位于相同街道则为 0.9,位于相同县则为 0.8,位于相同地区则为 0.7,位于相同市则为 0.5,位于相同省则为 0.3,位于相同国家则为 0.1,其余均为 0.

内部相似性选取用户偏好作为评价标准,偏好相似的用户说明他们的兴趣相似,因为在 SNS 中的很多应用有大量用户参与,从中不难采集到用户的兴趣爱好所在,也可以通过这些应用的数据计算出用户之间的偏好相似程度.这里借鉴协作过滤推荐中计算用户相似性的方法,选择了相关相似度量公式,即 Pearson 相关系数 (Pearson Correlation Coefficient)^[21] 并进行适当调整作为内部相似度的度量.设用户 A 和用户 N 共同评分过的项目集合为 $I_{A,N}$,则用户 A 和用户 N 的内部相似度 $S_i(A, N)$ 为

$$S_i(A, N) = 1 + \frac{\sum_{c \in I_{A,N}} (R_{A,c} - \bar{R}_A)(R_{N,c} - \bar{R}_N)}{\sqrt{\sum_{c \in I_{A,N}} (R_{A,c} - \bar{R}_A)^2} \sqrt{\sum_{c \in I_{A,N}} (R_{N,c} - \bar{R}_N)^2}} \quad (6)$$

其中, $R_{A,c}$ 表示用户 A 对项目 c 的评分, \bar{R}_A 和 \bar{R}_N 分别表示用户 A 和用户 N 对项目的平均评分.注意,平均评分是指两个用户有共同评分项目的平均值.求得和相关相似性系数的范围是 $[0, 2]$, 相关系数越大,则表示这两个用户的兴趣爱好越相似.而用于计算评分的项目集合可以是现在 SNS 中流行的电影或者书籍评分系统中的项目.

2.3 算法过程及性能分析

根据上述算法思想,假设 A 为根用户, N 为目标用户群体,整个算法的计算流程如下:

1. ORD \emptyset
2. for all target user n in N do
3. calculate the trust generated by the familiarity of A and n
4. end for
5. for all target user n in N do
6. if n is familiar to A then
7. calculate the trust generated by the similarity of A and n with the inside part of higher weight
8. else { /* if n is not familiar to A */ }
9. calculate the trust generated by the similarity of A and n with the outside part of higher weight
10. end if
11. end for

12. add the trust generated by familiarity and similarity
13. sort all the users of N by their trust
14. return ORD

本算法中,主要的步骤是先计算目标用户 n 与根用户 A 的熟悉性产生的信任度(2~4步),然后在此基础上再根据用户的熟悉性来计算相似性产生的信任度(5~11步),再把熟悉性产生的信任度和相似性产生的信任度相加(12步),最后再根据信任度进行用户排序(13步).在计算熟悉性相关的信任度时,要遍历所有与根用户 A 相关的目标用户,并且根据式(2)来通过遍历多条路径来计算,理论上,在如图 4 所示的 N 个节点的图中,两个节点之间路径数计算时间复杂度最坏为 $O(3^{N/3})$,因此,熟悉性产生信任度的时间复杂度最坏为 $O(N \cdot 3^{N/3})$.但在实际的社交网络中,用户的社会关系网络在系统数据库中都有静态的存储,而且节点之间并没有如此互相密集的关联度,因此,两个用户之间路径数均是常量,因此,基于熟悉性产生的信任度计算时间复杂度为 $O(N)$;在基于用户相似性计算信任度的步骤中,外部相似性是根据用户的属性来计算的,而属性的个数一般都是常量,因此,该步骤的时间复杂度为 $O(N)$;在内部相似性的计算中,要遍历两个用户之间共同评分过的项目集合,假定为 M ,则时间复杂度为 $O(N \cdot M)$,在实际的系统中, M 也是一个常量,因此,内部相似度的时间复杂度为 $O(N)$.因为需要遍历所有的目标用户集 N ,因此,相似性产生的信任度计算复杂度为 $O(N^2)$.另外,一般的排序算法的时间复杂度也为 $O(N^2)$.因此,该算法的总体的时间复杂度为 $O(N^2)$.

3 实验结果与分析

3.1 实验环境及实验场景

从实验环境来说,作者所在实验室已经搭建了一个基于手机通信录的移动社交网络服务系统^①.该新型手机通信录系统除了具有传统的电信服务功能外,还具有即时通信功能和社交服务能力,如图 5 所示.

为了说明本文所提出的社交网络中一种基于用户上下文的信任度计算方法在实际应用中的价值,我们设计了一个基于用户地理位置的兴趣点社区推荐服务(如饭店、旅游景点、旅馆、酒吧、KTV 等)的

① <http://mobilesns.int.bupt.cn/2008/>

应用场景. 目前的移动社交网络服务系统中关于兴趣点社区的评论中大都类似于大众点评一样的模式, 用户不知道哪些是好友的点评, 哪些是一般大众的点评, 更没有考虑用户对不同评论的信任程度. 因此, 现实生活中“熟人推荐”的行为模式并没有得到充分的挖掘和利用. 为了提供更好的用户体验, 项目组研发的基于手机通讯录的移动 SNS 实验系统中的基于地理位置的兴趣点社区应用中(如图 6 所

示), 对兴趣点社区的点评进行分类, 使得用户不仅可以看到一般大众的点评, 如果自己的好友有评价, 也可以单独看到; 更为重要的是, 利用本文提出的用户信任度计算方法, 对点评社区的用户进行了信任度打分, 然后根据信任度高低来依次将用户的评论展现给用户, 充分发挥了现实生活中“熟人推荐、口碑营销”的作用, 从而进一步提升了用户体验.



图 5 基于手机通讯录的移动 SNS 服务器端 Web 界面及手机客户端界面



图 6 基于用户地理位置的兴趣点社区推荐服务

3.2 实验数据采集

为了仿真更真实的用户情况和大量实验样本的需求, 项目组让实验室的 22 名在读研究生利用业余时间从国内著名的社交网站人人网^①上基于自己的社交圈采集好友数据, 同时让研究生发动各自的好友也参与数据的采集, 历时 1 个月共采集了 1286 个用户的个人信息. 然后将这些用户信息导入到实验

系统中, 通过用户及其之间的好友关系连接形成一个小规模的复杂社会化网络, 该网络包含 1286 个节点、21612 条边. 为了直观地了解这个社会化网络, 本文利用社区划分中的 Label Propagation 算法将这 1286 个用户组成的社会关系网络划分为若干个

① 人人网, <http://www.renren.com>

子社区,经过划分之后共出现了 19 个子社区,如图 7 所示,图中不同颜色的结点簇代表不同的社区.

下面的实验中将随机选取数据集中的一名用户作为根用户,随机浏览该应用中的某一社区,并采集根用户与这些用户之间的消息交互次数(打电话、留言、即时消息等)作为计算由熟悉性产生的信任度的上下文;其基本属性(实验中选取年龄和家庭所在地)作为计算外部相似性的上下文;另外采集应用中根用户及目标用户对共同评分过的地理社区的评分作为计算他们内部相似性的上下文.

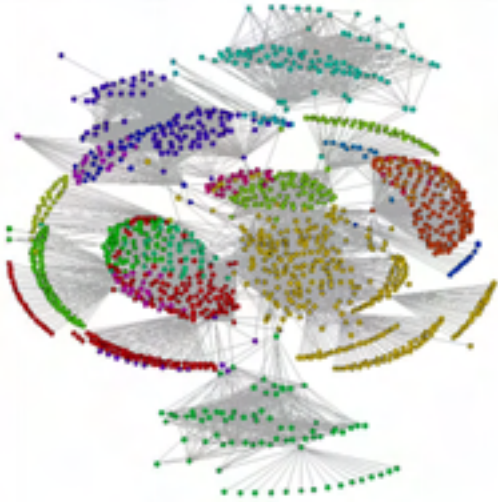


图 7 采集于人人网的关系网络社区

3.3 用户信任度的有效性验证

为了验证本文提出的社交网络中基于用户上下文的信任度计算方法的有效性,本文与文献[22]所提出的社会网络信任推理算法 TidalTrust 进行了分析和比较. TidalTrust 算法的思想是社交网络中的每个用户为与自己直接相连接的朋友进行信任度打分,然后基于这个信任度网络来计算不直接相连的用户之间的信任度. 从用户的主观判断来说,对自己的好友更容易做出相对准确的信任度评判,但对于好友的好友以及关系更疏远的用户,信任度评判的准确性会显著降低(由于用户的评分会变得越来越随意). 因此,为了降低用户对其它不直接相连的用户的信任度评分的主观性和随意性,本文采用了 TidalTrust 算法来进行不直接相连的用户之间的信任度计算,然后用 TidalTrust 算法的结果和本文所提方法的结果进行分析和比较. 实际上,绝对的信任度值对用户来说意义并不是很大,相对信任度更有意义,用户只要能够区别开不同的用户信任度顺序就可以,因此,本文将 TidalTrust 算法中目标用户的排序顺序和通过本文信任度计算方法得到的目

标用户信任度排序顺序进行了比较.

为了评估所提出的用户信任度计算方法的准确性,本文采用评价指标 $NDCG$ (Normalized Discounted Cumulative Gain) 来衡量信任度排序的准确率. $NDCG$ 是搜索系统中广泛使用的一种排序评价手段, $NDCG$ 体现了强相关搜索结果比弱相关搜索结果和不相关结果有用,而且强相关的结果出现在列表中越靠前(rank 越高)越有用,一个强相关的结果排名靠后则应该受到惩罚. $NDCG$ 的定义为

$$NDCG_p = \frac{DCG_p}{IDCG_p} \quad (7)$$

其中,本文中 DCG_p 采用的计算公式为 $\sum_{i=1}^p \frac{2^{rel_i} - 1}{\log_2(1+i)}$, rel_i 表示第 i 个目标节点用户的信任度, p 表示进行信任度排序的目标用户数, $IDCG_p$ 表示完美的排序情况下的 DCG_p 结果,完美的信任度计算算法形成的排序结果会使 DCG_p 和 $IDCG_p$ 相同,从而使 $NDCG_p$ 为 1,实际中, $NDCG$ 的取值在 0 到 1 之间. $NDCG$ 值越接近 1 就表明算法的排序结果越准确.

本文在实验中选取了朋友连接比较稠密的 50 个用户来分别对各自的好友进行信任度打分,从而形成了一个包含 476 个用户节点的社会关系子网络;然后基于这个子网络分别选择不同的根节点用户做了 3 个实验分析,实验中按照下述规则分别选取了 30 个目标节点用户来进行分析比较,通过更换不同的根节点和目标节点用户,每个实验重复 7 次来进行统计,实验 1 和实验 2 分别是针对目标用户中好友比例为 60% 和 20% 的情况,实验 3 是随机选取目标用户的情况.

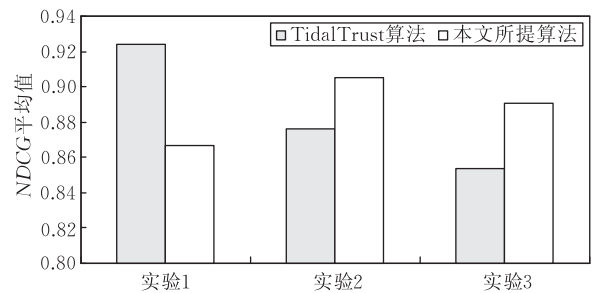


图 8 实验分析结果

从图 8 的实验结果可以看出,如果目标用户中好友数量占多数的情况下, TidalTrust 算法比本文所提算法的准确性稍高一些,主要是用户好友的信任度由用户直接评分所得,而本文所提算法是依据于用户之间的熟悉度和相似度来综合计算信任度,有的好友之间可能虽然比较熟,但消息交互量和相似度可能低一些. 而随着在目标用户中好友数量的

减少,本文所提方法在准确性方面就比 TidalTrust 算法更好一些,主要是本文所提方法综合考虑了用户之间的熟悉度和相似度,能够深层次地挖掘出用户之间的相似性,从而对用户的信任排序更符合用户的实际情况.另外,由于 TidalTrust 算法的前提条件是社交网络中的每个用户与自己的好友之间有信任度评分,然后来计算不直接相连的用户之间的信任度.这种前提条件的建立在一般的社交网络中有一定的困难,本文所提方法更符合现实生活中的商品或者社区推荐需求,可以基于用户之间的交互情况及购物或者评价信息来综合计算和分析用户之间的信任度,适用范围更广.

3.4 用户信任度计算在社交网络服务中的应用实证

实验所采集的用于计算熟悉性产生的信任度上下文数据(包括用户间的人际关系以及用户间的消息交互数)如图 9,其中 ID 为 193 的用户作为根用户,其它目标用户为某一项目中所有发表过点评的

用户,共 64 名;另外采集了用于计算外部相似性产生的信任度上下文数据:包括用户的年龄、家庭所在地(实验中由于大部分用户家庭所在地选填到市级,所以本实验另外采集了用户所在高中作为家庭所在地的判别依据);以及内部相似性产生的信任度上下文数据(根据二者共同评分过的地理社区评分计算得到他们之间的相似性),部分源数据如表 2 所示.为了保护用户隐私,实验中每个 ID 为系统中的真实用户的 ID 经过 Hash 之后所得,而且本实验中并不采集用户的姓名及联系方式等属性.由于数据集中大部分用户所填写的家庭所在地基本上都只到市一级,所以本实验中调整外部相似性中 S_h 的计算方法如下:若用户的家庭所在地位于同县则为 1,位于相同市则为 0.9,位于相同省则为 0.8,位于相同地区(东北、华北、华南等)则为 0.7,家庭所在地同为南北方则为 0.6,其余均为 0.

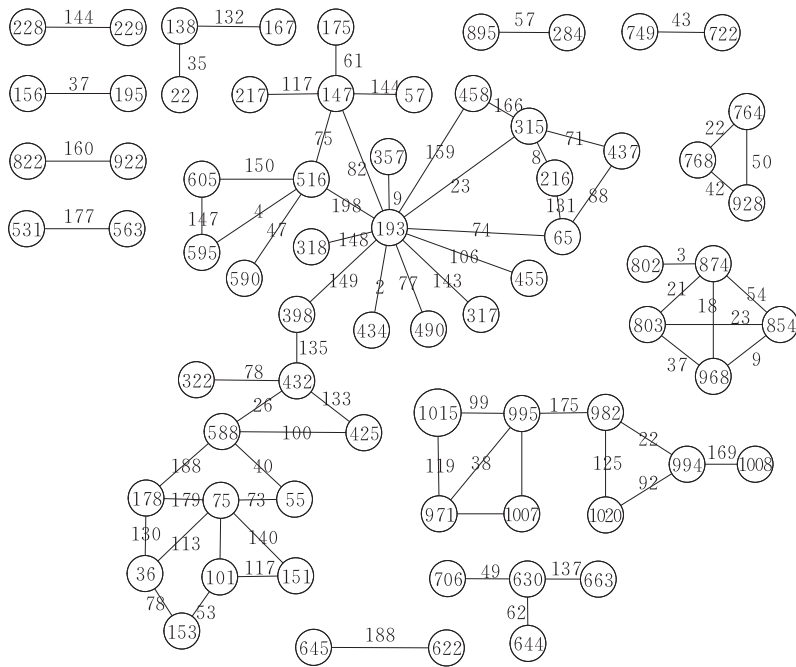


图 9 用户人际关系及消息交互数据

表 2 部分用户属性数据

用户 ID	年龄	家庭所在地	高中
193	25	山东威海市	荣成市第二中学
65	25	山东威海市	荣成市第二中学
318	24	北京海淀区	北京市十一学校
357	24	山东烟台	龙口市第一中学
437	25	山东菏泽	东明县实验中学
455	25	安徽阜阳	阜阳市城郊中学
531	35	安徽宣城	安徽省宣城中学
588	24	广东省江门市	江门一中
605	24	辽宁省辽阳市	辽阳市第一高级中学
928	24	福建省南平市	光泽一中

计算所得的部分结果如表 3 所示,本实验中选取按照熟悉性最大的前 30% 的用户作为熟悉用户,进而计算相似性产生的信任度,最后将熟悉性产生的信任度和相似度产生的信任度结合起来得到 ID 为 193 的用户对其它所有目标用户的信任度.

实验的最终结果是根据浏览用户(ID 为 193)与所有点评用户(目标用户)之间的信任度的高低将目标用户发表的点评排列出来,以减少用户浏览筛选的时间和精力,效果如图 10 所示.



图 10 按照本文信任度计算方法所得信任度排序的社区点评图

表 3 实验部分结果

用户 ID	熟悉性	内部相似度	外部相似度
65	0.0632478632478632	0.840	2.0
318	0.126495726495726	1.750	1.6
357	0.00769230769230769	1.482	1.7
437	0.00728194399742516	1.959	1.8
455	0.0905982905982905	1.945	1.0
531	0.0000000000000000	1.088	0
588	0.00197288279222498	1.781	0.9
605	0.02655294496298680	0.985	1.5
928	0.0000000000000000	0.978	0.9

4 结束语

本文提出了社交网络服务中一种基于用户上下文的信任度计算方法,将社会学与心理学中人们之间信任的产生过程引入到社交网络服务中用于计算用户间信任度,把用户之间的信任度分为两部分:熟悉性产生的信任度和相似性产生的信任度,又根据所起作用的不同,把相似性分为内部相似性和外部相似性分别进行计算,使得社交网络中用户之间的信任度计算有了理论依据,充分提高了信任度计算的合理性和有效性,并且更贴近现实人际关系网络中用户之间的信任关系。本文详细描述了该计算方法的基本流程和详细过程,最后通过真实社交网络中的数据进行仿真实验,验证了本文所提出的信任度计算方法的合理性和有效性。

参 考 文 献

- [1] Wong X D, Yen D C, Fang X. E-commerce development in China and its implication for business. *Asia Pacific Journal of Marketing and Logistics*, 2004, 16(3): 68-83
- [2] Jones K, Leonard L. Trust in consumer-to-consumer electronic commerce. *Information Management*, 2008, 45(2): 88-95
- [3] Lu Y B, Zhao L, Wang B. From virtual community members to C2C e-commerce buyers: Trust in virtual communities and its effect on consumers' purchase intention. *Electronic Commerce Research and Applications*, 2010, 9(4): 346-360
- [4] Faloutsos M, Karagiannis T, Moon S. Online social networks. *IEEE Network*, 2010, 24(5): 4-5
- [5] Sinha R, Swearingen K. Comparing recommendations made by online systems and friends. Available from http://onemvweb.com/sources/sources/comparing_recommendaions.pdf, 2011.04.19
- [6] Liu F K, Lee H J. Use of social network information to enhance collaborative filtering performance. *Expert Systems with Applications*, 2010, 37(7): 4772-4778
- [7] Chou A Y. The analysis of online social networking: How technology is changing e-commerce purchasing decision. *International Journal of Information Systems and Change Management*, 2010, 4(4): 353-365
- [8] Kamvar S D, Schlosser M T. EigenRep: Reputation management in P2P networks// *Proceedings of the 12th International World Wide Web Conference (WWW2003)*. New York: ACM Press, 2003: 123-134
- [9] Song S, Hwang K, Zhou R, Kwok Y K. Trusted P2P transactions with fuzzy reputation aggregation. *IEEE Internet Computing*, 2005, 9(6): 24-34

- [10] Chen Hai-Bao, Zhao Sheng-Hui, Ji Cheng-Chao. Interaction trust model based on strength of relationships in P2P system. *Computer Engineering and Design*, 2009, 30(11): 2678-2630 (in Chinese)
(陈海宝, 赵生慧, 计成超. P2P 系统中基于关系强度的交互信任计算模型. *计算机工程与设计*, 2009, 30(11): 2678-2630)
- [11] Li Yong-Jun, Dai Ya-Fei. Research on trust mechanism for peer-to-peer network. *Chinese Journal of Computers*, 2010, 33(3): 390-405 (in Chinese)
(李勇军, 代亚非. 对等网络信任机制研究. *计算机学报*, 2010, 33(3): 390-405)
- [12] Jennifer G, James H. Inferring binary trust relationships in Web-based social networks. *ACM Transactions on Internet Technology*, 2006, 6(4): 497-529
- [13] Wang J C, Chiu C C. Recommending trusted online auction sellers using social network analysis. *Expert Systems with Applications*, 2008, 34(3): 1666-1679
- [14] Bhuiyan T, Xu Y, Josang A. Integrating trust with public reputation in location-based social networks for recommendation making//*Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. Sydney, Australia, 2008; 107-110.
- [15] Lei Huan, Peng Jian. Integrating reputation and subjective logic to analyze trust network in SNS. *Application Research of Computers*, 2010, 27(6): 2321-2323 (in Chinese)
(雷环, 彭舰. SNS 中结合声誉与主观逻辑的信任网络分析. *计算机应用研究*, 2010, 27(6): 2321-2323)
- [16] Peng Si-Qing. Trust establishment mechanism: Relationship operation and law means. *Sociological Studies*, 1999, (2): 53-66 (in Chinese)
(彭泗清. 信任的建立机制: 关系运作与法制手段. *社会学研究*, 1999, (2): 53-66)
- [17] Zucker L G. Production of trust: Institutional sources of economic structure, 1840-1920//Staw B M, Cummings L L Eds. *Research in Organizational Behavior*. Greenwich, CT: JAI Press, 1986: 53-111
- [18] Zuo Bin, Gao Qian. The effects of familiarity and similarity on the interpersonal attraction. *Chinese Journal of Clinical Psychology*, 2008, 16(6): 634-636 (in Chinese)
(佐斌, 高倩. 熟悉性和相似性对人际吸引的影响. *中国临床心理学杂志*, 2008, 16(6): 634-636)
- [19] Simpson J A, Rholes W S. *Attachment Theory and Close Relationships*. New York: the Guilford Press, 1998
- [20] Lo S C H, Lin C H C H. WMR — A graph-based algorithm for friend recommendation//*Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI'06)*. Washington, DC, USA: IEEE Computer Society, 2006; 121-128
- [21] Ahlgren P, Jarneving B, Rousseau R. Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. *Journal of the American Society for Information Science and Technology*, 2003, 54(6): 550-560
- [22] Jennifer G. Computing and applying trust in web-based social networks [Ph. D. dissertation]. University of Maryland, College Park, MD, 2005



QIAO Xiu-Quan, born in 1978, Ph. D., associate professor. His main research interests include the intelligent theory and technology of network services.

YANG Chun, born in 1985, M. S. candidate. His main research interests include mobile social networking services.

Background

The consumers' comments on the commodities or services have played an important role in E-commerce activities and always influence the users' consuming behaviors. How to evaluate the reliability and trusty degree of these comments is an interesting topic for consumers. With the boom of social networking services, the e-commerce activities based on the real social network are becoming more and more popular. The interpersonal relationships represented by the social network provide a good research basis of trust between different consumers. Therefore, the research on the users' trust for e-commerce based on SNS is very significant.

This group's research interests are currently focused on the mobile social networking services. The long-term goal is

LI Xiao-Feng, born in 1950, professor. Her main research interests include intelligent network and communication software.

CHEN Jun-Liang, born in 1933, professor, Ph. D. supervisor, member of Chinese Academy of Sciences, member of Chinese Academy of Engineering. His research interests include services computing, intelligent theory and technology of network services.

to build a user-centric intelligent convergent service environment based on the novel mobile phone book, which can seamlessly integrate the traditional telecommunication services, instant messaging service, social networking services and personal information management service. This work is supported by National Natural Science Foundation of China under Grant No. 60802034, National Basic Research Program (973 Program) of China under Grant No. 2012CB315802, Beijing Nova Program under Grant No. 2008B50 and New Generation Broadband Wireless Mobile Communication Network Key Projects for Science and Technology Development under Grant No. 2011ZX03002-002-01.