基于用户信息的社交网络信任评估方法*

王培人，毛 剑，马寒军，崔 键
（北京航空航天大学 电子信息工程学院，北京 100191）

摘 要： 社交网络的信任评估是社交网络应用安全的核心组件，公平客观的信任评价结果对于用户获取信息的正确性非常关键。利用用户之间信息的相似度进行用户聚类，实现用户群体的划分，依据用户集群的结果对用户之间关系进行修正，调整信任评估结果；同时还考虑到了恶意用户的影响，加入了信任检测的方案来保证方案的鲁棒性。经过实际社交网络数据实验仿真证明，该算法不仅可以使得信任评估结果与同类用户预期更为一致，而且可以大幅度降低恶意用户刷分行为的影响，实现有效而可靠的信任评估。

关键词： 社交网络；信任评估；数据聚类
中图分类号： TP309.2 **文献标志码：** A **文章编号：** 1001-3695(2018)02-0521-06
doi:10.3969/j.issn.1001-3695.2018.02.043

User information based trust evaluation mechanism for social network

Wang Peiren, Mao Jian, Ma Hanjun, Cui Jian
(*School of Electronic & Information Engineering, Beihang University, Beijing* 100191, *China*)

Abstract: Trust evaluation is a key part of social network security. An impartial and objective way to evaluate trust is important for users to acquire requisite information. This paper utilized data clustering method to classify user based on users' background information, and corrected relationship between users by their background information to estimate trust value. In addition, it introduced behavior tracking mechanism to withstand malicious acts. After simulation test, this proposed method can not only get trust value closer to users' expectation, but also restrict the malicious effects of malicious users, thus achieving effective trust evaluation.

Key words: social network; trust evaluation; data clustering

0 引言

社交网络是一个由一系列成员、动态链接与社交联系组成的社会结构[1]。Facebook 是社交网络的典型代表，其用户可以互相结为好友并分享内容[2]。现在社交网络有许多种类[3]，比如 LinkedIn[4]（关注于职场社交）、Instgram[5]（图片分享型网址）、Tumblr[6]（轻型博客）。除此之外，还有许多关注与音乐、美食、电影分享的社交网络。

目前社交网络突出的安全方面问题主要有[7]：a) 信任评估问题，代表着用户是否应该相信在社交网络中其他用户提出的信息；b) 隐私保护问题，如何防止其他用户获取使用者的隐私信息。

本文重点针对社交网络的信任评估问题。方案基于用户不同社交网络因素影响程度对用户信息建模，利用自组织映射方法进行用户的聚类分析，实现用户基于相似度关系的信任评估；同时针对用户恶意刷分行为，引入信任监测机制，该机制以 Caverlee 等人的模型[8]为核心，监测用户一段时间信任值变化以修正信任值波动，实现方案鲁棒性的提高。

1 相关工作

TidalTrust[9]是较早的关于社交网络信任评估研究算法之一，用户对已有接触的用户进行评分，当用户需要信任评估时，系统会遍历整个网络搜索到该用户节点的最短路径，并综合计算出两个用户之间的信任值。然而仅仅考虑最短路径的信任传递与聚合未必能满足信任评估的需求，这样可能会遗漏其他路径的重要信息。针对这个问题，Avesani 等人[10]提出了一个考虑所有不超过阈值长度路径的信任评估算法 MoleTrust。Lesani 等人[11]在 FuzzyTrust 算法中进行信任评估时采用了语义分析的方法。Lin 等人[12]在信任评估方案中采用了“yes/no”机制，而不是评分机制来进行信任评估。Jiang 等人[13]提出了一种优化方案，基于用户关系链与话题相关性预处理建立大型社交网络的信任分布图，以提高广度优先的用户遍历效率。

文献[14]的方案在搜寻使用者与目标用户之间的路径时，会根据用户之间的相似度调整遍历的顺序进而影响信任评估的结果。Chang 等人[15]提出了一种基于数据聚类的信任评估方法，是一种基于机器学习的信任评估方法。该方案利用数据聚类方法对用户关系进行信任评估，但并没有考虑到恶意用户的存在，而恶意用户的刷分等行为可能会影响方案的准确性。

文献[16]在信任评估中利用信任监测来减少恶意用户的影响，利用历史的信任行为变化来修正现有的信任值变化。类似地，Zhang 等人[17]提出了一种监测用户在社交网络中出现的

收稿日期：2016-10-25；修回日期：2016-12-12 基金项目：国家自然科学基金资助项目(61402029)

作者简介：王培人(1992-)，男，山东烟台人，硕士研究生，主要研究方向为社交网络安全(fwfredwang@163.com)；毛剑(1978-)，女，讲师，博士，主要研究方向为网络与信息安全、大数据安全与隐私保护；马寒军(1992-)，男，硕士研究生，主要研究方向为是移动网络安全；崔键(1989-)，男，硕士研究生，主要研究方向为网络与信息安全与隐私保护.

次数来判断信任度的方法。Xia 等人[18]提出的信任评估方法是通过考量节点的历史行为表现情况，并建立模型预计未来的表现来进行信任值评估。Nepal 等人[19]提出基于社交网络互动的信任评估模型，将信任分为流行度信任度和投入信任度，通过活跃程度来量化信任。

2 方案模型

本文主要研究一种社交网络中基于用户信息的信任评估方法。现在的社交网络大多将信息以一种“平等”的方式来显示信息，然而社交网络往往是“不平等”的[20]。用户在交流和作决策时，并不会平等对待每个人。有研究表明，人们更愿意相信有相似背景的人[21,22]。因此用户间个人信息的相似度是信任评估的重要依据，但与此同时，信任评估还要考虑恶意用户的影响。要实现基于用户信息的社交网络信任评估，需要解决两个问题：a) 如何量化社交网络中的用户信息并实现分类；b) 如何降低恶意用户行为的影响。

本文通过实现用户之间基于相似度关系进行的用户分群，利用数据聚类方法进行数据的关联性分析。经过聚类的数据集、相近的数据节点会聚集在一个集群之中，而不相近的数据会分离在不同集群。数据聚类方法有很多种，比如 K-means 聚类算法[23]、自组织映射聚类算法（self-organizing map，SOM）[24]、模糊 C-均值聚类算法[25]等。本文基于 SOM 算法进行用户数据聚类分析，主要考虑到 SOM 算法在运算效率、准确性等方面的优势。相比于文献[15]基于数据聚类的方法，本文详细量化了用户向量建立的原则与方法，使得用户向量建立更加有效。

本文通过自组织映射进行数据聚类，将用户根据其本身数据特点的相似度来形成相应的集群，在信任评估过程中，就可以使得相近用户的意见更为重要，而不相近用户的意见重要性降低。在进行自组织映射之前，需要将用户信息进行量化，提取用户特征建立用户向量，将量化后的用户信息向量经过自组织映射实现用户聚类。同时为了抵御恶意用户刷分行为，本文结合了 Caverlee 等人提出的信任监测机制[8]，当有恶意用户刷分行为出现时，对信任评估结果进行修正，降低恶意用户行为的影响，提高信任评估方案的鲁棒性。

所提出的方案架构图与工作流程如图 1 所示，当图中使用者（虚线圈中的深色用户）需要评估目标对象信任值时，方案会收集其他评价过该对象的用户评价以及用户背景信息。利用背景信息给用户建立用户向量，评价信息用于信任评估结果。

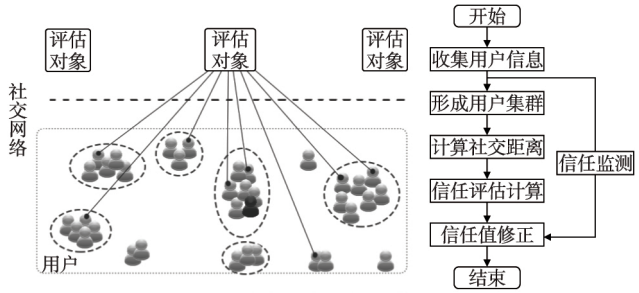


图1 方案架构图与工作流图

方案的运行包含如下步骤：

a) 形成用户集群。通过将用户的背景信息进行向量化处理，本文可以将用户向量输入自组织映射以形成多个用户群（集群是依据用户背景信息数据的相似性得到的）。在图 1 中，表现为多个虚线圈中的用户，相似的用户会靠近，而不相干的用户会不变或者疏远。

b) 收集评价者信息。当使用者需要了解目标对象信任值时，本文会收集评价者信息，并计算使用者与评价者的社交距离。

c) 计算用户间的社交距离与关系权重。用户之间的社交距离是通过经过用户集群后的向量距离计算得来的，而关系权重是综合了社交距离与社交网络特点得到的一个综合量。

d) 信任监测。方案运行时会监测目标对象的信任值变化情况，对较大的波动进行修正。当出现异常行为（波动大于阈值）时，会强行限制信任值变化情况。

e) 信任聚合。将数据聚类得到的社交关系权重与信任监测的结果综合并进行修正，得到最后的信任值评估结果。

本文方案中的数学符号说明如表 1 所示。

表 1 符号说明

符号	描述
$Tr(X)$	针对评价的目标对象 X，评价者原始的评价值
u_i	由用户 i 背景信息形成的 N 维向量，$u_i=[u_{i1},\cdots,u_{iN}]$
r_{Xj}	评估对象 X 的第 j 个分量的大小，$0<r_{Xj}<1$
k_{ij}	用户 i 的第 j 个原始分量大小
$d_X(i,j)$	针对评价的目标对象 X，用户 i、j 之间的向量距离，表示用户间关系的相似度
$I_x(i,j)$	针对评价的目标对象 X，用户 i、j 之间的社交关系权重，表示用户间影响程度
$TV_h(X)$	针对评价的目标对象 X，信任值监测得到的历史聚合信任值
$c(X)$	针对评价的目标对象 X，信任监测得到的信任修正值
$TV'(X)$	目标对象 X 经过基于用户相似度关系得到信任评估结果
$TV(X)$	目标对象 X 的信任评估结果

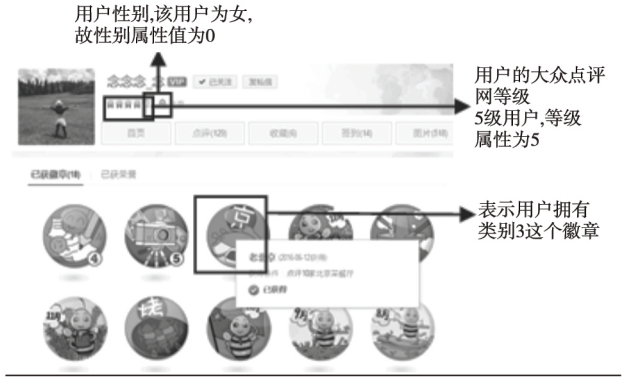
2.1 社交关系权重计算

在计算用户之间社交相似度时，第一步是建立用户向量。建立用户向量需要有以下几个特点：a) 能够代表用户的最主要特征；b) 向量本身能够具有独特的意义，即根据一个向量可以还原出一个用户大致是怎样的一个状态。在建立向量中，有一些比较常见的元素，比如性别、年龄等；还有一些专属信息，如专业信息、消费习惯、教育程度等，这些信息因不同社交网络关注信息不同而有所区别。

本文采用的测试平台为大众点评网[26]，因为它集合了“信任评估”与“社交”两个因素。主要收集的信息包括：a) 用户基本信息，这里仅收集了用户的性别信息（初始设置 1 为男性，0 为女性），没有选择年龄是因为在用户注册大众点评网的背景信息中并没有要求填写，则不作为评估量化的维度；b) 用户在网站整体等级，代表着用户的整体评论行为，高的等级代表着更多更高质量的点评；c) 用户的类别偏好。大众点评网有一个“徽章”机制，当一个用户在某个类型的美食中有过一定数量的高质量评论后，网站会给用户一个徽章。这样可以通过收集每个用户的徽章信息，判断用户是否在这个类别中有较高的偏好。而用户对于餐饮类别的偏好并不具有唯一性，即用户可以同时喜爱多个类别，因此在选择的 N 个餐饮类别设置为 N 个用户背景信息向量分量，初始设置为当用户拥有这个类别的“徽章”，则该类别设置为 1，没有则为 0。

综上，本文基于大众点评网的用户背景信息向量建立了一个 N 维向量，用户 i 的用户向量 $u_i=[r_{i1},\cdots,r_{iN}]$，其中 r_{ij} 表示

用户 i 的第 j 个分量大小。根据大众点评网的社交网络特性,建立的用户向量信息为 [用户的大众点评网等级, 性别,餐饮 1,…,餐饮 N]。如图 2 所示的用户,性别女,在大众点评网等级为 5,只偏爱类别 3,初步建立该用户的用户向量 u_i 为 [5,0,0,0,1,0]。



网络等级	性别	类别1	类别2	类别3	类别4
5	0	0	0	1	0

图2　用户原始信息收集

用户向量 u_i 的建立需要量化每个向量分量的大小,因为对于不同评估对象而言,每个向量分量的影响大小是有区别的,每个分量造成的影响大小即可量化为用户背景信息向量中每个分量的取值范围的最大值。

在社交网络中,针对不同的目标评价对象,不同因素对于用户进行决策起到的影响大小是不同的。因此本文定义每个向量维度的取值大小就是由在这个社交网络中,该因素分量能造成影响的大小程度。当一个因素在取不同值的情况下,统计得到的信任值的区别越大,则认为这个因素影响越大。如针对同一个目标 X,男性与女性用户对 X 的信任评估结果基本相同,则认为性别这个因素影响较小; 而若社交网络等级 1 与 6 的用户对 X 信任评估结果相差较大,则认为社交网络等级这个因素的影响较大。而一个因素影响的大小决定着这个因素对应着的向量分量取值大小。

衡量影响大小的方法,是通过统计目标 X 在每个因素分量不同取值下,信任评估结果的差值来衡量的。如针对目标 X,男性用户的信任评估结果平均值为 3,女性用户的信任评估结果平均值为 4,则两者差值绝对值 1 代表这个因素的影响大小。当存在多个取值时,如社交网络等级,则取差别最大的作为这个因素的影响大小。每个因素都依次可以得到一个差值的绝对值 p_{Xi}。

利用式(1)对每个因素的影响大小进行量化,计算每个因素的影响大小 r_{Xi}。

$$r_{Xi}=\frac{p_{Xi}}{\sum_{k=1}^{N}p_{Xk}} \tag{1}$$

其中: p_{Xi} 表示统计得到的每个分量影响差值的绝对值, N 表示一共有 N 个因素。这样统计每个因素分量造成的影响大小,并取相对大小作为该因素所占的权重,以相对影响大小 r_{Xi} 作为第 i 个因素分量的最大值。结合收集的用户每个原始信息分量 $k_{i1},\cdots,k_{iN}$,从而用户 i 向量 u_i 建立为

$$u_i=[r_{X1}\cdot k_{i1},\cdots,r_{XN}\cdot k_{iN}] \tag{2}$$

若针对一个目标对象 X,分别统计 X 在统计评价用户的用户大众点评网等级、性别、餐饮 1、…、餐饮 N 这些维度上的信任评估差值,如分别为 3、1、1、1、0、1(假设共有四个餐饮类别),则每个因素分量的影响大小归一化后的结果为 0.43、0.14、0.14、0.14、0.14。假设使用者仍然是一个男性用户,在大众点评网等级为 1,只偏爱类别 1、2 的用户,原始用户向量为 [1,1,1,1,0,0],则其最终形成的向量为 [0.43,0.14,0.14,0.14,0,0],其他用户的背景信息向量建立依此类推。

在建立用户向量之后,利用自组织映射可以将用户向量按照其本身特点的相似度来形成集群。将用户信息向量输入自组织映射进行聚类,输出的结果仍然是向量的形式。从而两个用户之间的相似度可以用两个用户向量之间的向量距离来衡量,故针对某个信任评估对象 X,用户 i 与 j 之间的向量距离计算公式为

$$d_x(i,j)=\|u_i-u_j\| \tag{3}$$

这样就得到了两个用户 i、j 之间的向量距离,但向量距离并不能直接表示用户之间的相互影响程度,因此本文将描述相似度的变量定义为社交关系权重 $I_X(i,j)$,表示针对目标对象 X,用户 i 与 j 之间的背景信息的相似度。

社交关系权重与用户信息向量间应有如下关系: 两者应当成负相关关系,向量之间的距离越大,说明两个用户之间背景信息区别越大,两个用户越不相似,根据之前的结论,这个用户的影响权重应该尽量小。信任值评估中,本文基于式(2)得到的向量距离,提出了两种社交关系权重量化方案,如式(4)(5)所示。

a) 幂函数变换:

$$I_X(i,j)=1-\left(\frac{d_x(i,j)}{d_{\max}}\right)^{\gamma} \tag{4}$$

b) S 型曲线:

$$I_X(i,j)=\frac{1}{1+e^{\left(1-\frac{2}{d_{\max}}d_x(i,j)\right)}} \tag{5}$$

其中: $d_x(i,j)$ 是用户之间的向量距离; $d_{\max}$ 是用户之间向量距离中出现的最大值; γ 为幂函数变换中关系权重调节参数; $\gamma>0$,控制社交距离与关系权重之间的变换关系。

两种曲线有不同的特性:

a) 在幂函数变换中,γ 控制着随向量距离增加,社交关系权重衰减的速度。γ 越大,社交关系权重衰减得越慢,原本向量距离较大的用户的影响程度越大; γ 越小,社交关系权重衰减得越快,原本向量距离较大的用户的影响程度越小。

b) 当为 S 型曲线时,在中间部分变化较快,两端部分较为平缓,在使得用户社交关系权重整体趋向于两极化分布同时,一定程度保留了相异用户的部分意见,使得整体趋向于阶梯型分布。

这两种模型分别适用于不同社交网络情况,可以作为用户的方案选择,如激进型(γ 较小)、保守型(γ 较大)、均衡型(S 型曲线)等。在之后的性能评估模块将进一步测试每个方案的表现。

2.2　信任值计算

基于此前 2.1 节计算所得的关于关系权重进行信任评估计算。若使用者 i 对目标对象 X 进行信任评估,可以得到使用者 i 与每个评论者之间的关系权重 $I_X(i,j)$(j 表示每个评论者),以及收集的相应评论者 j 对于目标对象 X 的原有评分值 $Tr(j)$。利用关系权重 $I_X(i,j)$ 对评分值 $Tr(j)$ 进行加权平均,中间信任值 $TV'(X)$ 计算公式如式(6)所示。

$$TV'(X) = \frac{\sum_{j \in rel(X)} I(j) \times Tr(j)}{\sum_{j \in rel(X)} I(j)} \tag{6}$$

其中: $rel(X)$ 表示所有评论目标对象 X 的用户，即 j 代表所有的参与评论过目标对象 X 的评论者。

式(6)得到的 $TV(X)$ 为信任评估结果，通过利用关系权重对原有的信任值进行处理，让使用者 i 可以得到更契合自己的信任评估结果。

2.3 信任监测

信任监测主要是为了抵御社交网络中可能存在的恶意用户刷分行为。刷分行为是指一些用户故意有组织地提高或者降低某些目标对象的信任值。图3测试了在大众点评网，恶意刷分行为对目标对象信任值造成的影响。评估对象信任值定为3.5(统计的的平均值)，设定恶意用户全部刷高评分，即评分为5。图3给出了当存在50、100、150、200恶意用户刷分时，信任值的波动大小情况。

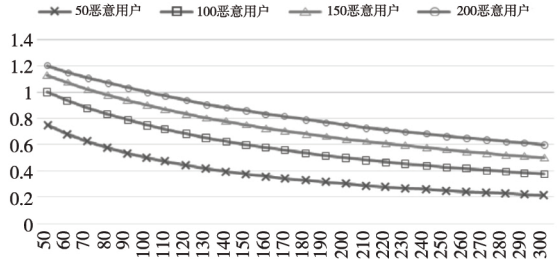


图3 信任值随恶意用户与评论数变化情况

X 轴表示商家已有评论的数量，带不同标志符的线表示恶意用户数量分别为50、100、150、200时的信任值变化情况。由图3可知，随着原有评论增多，恶意用户刷分行为造成的影响下降；恶意用户越多，影响也会越大。当用户拥有的评论数较少时，恶意刷分行为会导致信任值变化较为剧烈。

本文基于 Caverlee 等人提出的信任监测模型，监测过去的 N 个周期内的信任值，利用过去的信任值聚合来修正现在的信任值评估结果[8]。系统通过监测目标对象在一段时间内的信任值变化情况，监测并降低恶意用户行为的影响。本文方案会固定记录 N 个周期的值，随着时间的推移，最早的时间期会被抛弃并引入新的时间窗。信任监测模块的结构如图4所示。

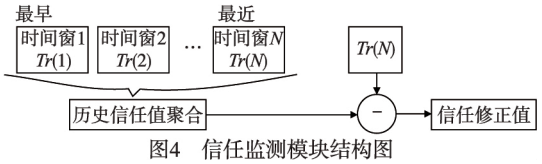


图4 信任监测模块结构图

式(7)为针对目标对象 X 的历史信任聚合值计算方法:

$$TV_h(X) = \frac{1}{\gamma} \times \sum_{k=1}^{N} TV_k(X) \times \alpha^{N-k} \tag{7}$$

目标对象 X 的信任值修正值 $c(X)$ 计算公式如式(8)所示。

$$c(X) = TV(X) - TV_h(X) \tag{8}$$

其中: $\gamma = \sum_{k=1}^{N} \alpha^{N-k}$ 是用来限制整体的值仍然处于原有的信任值范围内; α 是历史信任值控制参数，控制历史信任值对现在信任值评估的影响大小。$0 < \alpha < 1$，α 越大，历史信任值的影响程度越大; α 越小，历史信任值的影响程度越小。

利用信任修正值 $c(X)$ 对之前式(6)得到的信任评估值进行修正，得到的信任值计算公式为

$$TV(X) = \frac{\sum_{j \in rel(X)} I_X(i,j) \times Tr(X)}{\sum_{j \in rel(X)} I_X(i,j)} - c(X) \tag{9}$$

$TV(X)$ 为信任评估结果，方案基于用户之间的相似性进行用户集群，让信任评估的结果更符合使用者的特点; 同时利用信任监测机制，修正了恶意刷分行为存在(尤其是在目标仅有少量评价，易受影响情况)时，信任值的不正当波动。

3 方案性能评估

3.1 信任评估

文献[21,22]指出，用户更愿意相信与自己更为相似用户的意见，在评估信任评估值结果时，评判的标准就是“针对某一用户，给出的结果是否与所在类群的结果更为一致”。

大众点评网用户的整体评论习惯除了受餐厅本身特性影响外，餐厅受欢迎程度，参与评论的用户数对信任评估行为同样有较大的影响。消费者在网络上消费的满意度会受到商品原有的网络评价影响，形成原有的评价期望值，进而影响个人对该目标的评价。同时用户的行为也会受到从众效应的影响，倾向于作出更为一致的行为[23]。本文的测试将针对受欢迎程度不同的几个评价对象，并观察其表现。

本文选取了四个评价对象，分别对应着不同的大众点评网受欢迎程度(受欢迎程度由该评价对象拥有的评价数量决定，评价数量越多，该对象越受欢迎)，四个评价对象受欢迎程度分别为前10%、前10%～30%、前30%～50%、50%之后，评估对象拥有的评论数为423、136、89、8。

将用户分为高级用户与一般用户(有特别“徽章”，且用户等级大于3为高级用户，其他为一般用户)，统计四个评估对象，在一般用户与高级用户中每个分值人数占其该类用户总人数比例分布如图5～8所示。

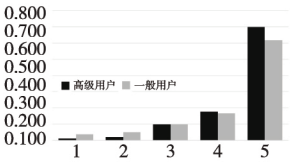


图5 评估对象1用户评分采样分布

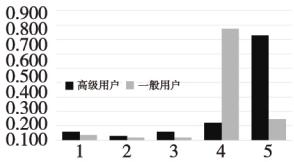


图6 评估对象2用户评分采样分布

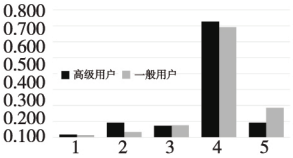


图7 评估对象3用户评分采样分布

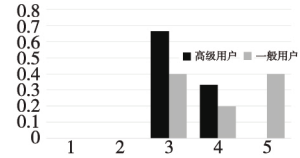


图8 评估对象4用户评分采样分布

由图5～8可得，评估对象1、3整体分布保持稳定，其中评估对象2两类用户差别较大，则运行方案后的信任评估结果与所在集群靠近的同时，会与整体平均值偏差更大。

针对评估对象1～4进行信任评估，使用者为一般用户与高级用户，使用不同的社交关系权重模型进行测试，得到表2和3，其中 Δ 表示评估结果与该用户所在类群(一般用户、高级用户)信任值平均值的差值绝对值。

由表2和3结果可见，除了对评估对象4之外，运行方案之后一般用户与高级用户都与所在集群的“观点”更为一致，整体的信任评估结果更为可信。评估对象4因为拥有的评论数较少，在形成数据集群样本集较少，分布规律性降低; 同时原

有评论数较少，个别用户的用户可能会对整体的评论产生较大的影响。但对于评估对象 1 ~ 3，在评估对象已经拥有了一定的评论数量之后，可见方案可以使得整体的信任评估更为有效。对评估对象 2 的信任评估结果，一般用户与高级用户有了较大差别，但优于本身数据差别较大，所以最后结果差值 Δ 也相对更大。结果还可以看出，幂函数 $\alpha = 2$ 与 S 型曲线这两种方案结果基本保持一致，因为两者社交关系权重模型较为相近。而当 $\alpha = 0.5$ 时，这种更为“激进”的策略在对评估对象 1 中产生了更好的结果，使用者与所在集群的更为相近，但也存在偏差较大的情况。

表 2　不同权重模型对高级用户信任评估结果的影响

参数	高级用户		
	幂函数 ($a = 0.5$)	幂函数 ($a = 2$)	S 型曲线
评估对象 1			
平均值 = 4.36	4.50	4.50	4.49
高级用户平均值 = 4.53	$\Delta = 0.03$	$\Delta = 0.03$	$\Delta = 0.04$
一般用户平均值 = 4.31			
评估对象 2			
平均值 = 4.07	4.28	4.33	4.35
高级用户平均值 = 4.41	$\Delta = 0.13$	$\Delta = 0.07$	$\Delta = 0.08$
一般用户平均值 = 3.97			
评估对象 3			
平均值 = 3.99	3.89	3.83	3.83
高级用户平均值 = 3.78	$\Delta = 0.1$	$\Delta = 0.05$	$\Delta = 0.05$
一般用户平均值 = 4.06			
评估对象 4			
平均值 = 3.75	3.68	3.81	3.78
高级用户平均值 = 3.33	$\Delta = 0.33$	$\Delta = 0.48$	$\Delta = 0.43$
一般用户平均值 = 4.00			

表 3　不同权重模型对一般用户信任评估结果的影响

参数	高级用户		
	幂函数 ($a = 0.5$)	幂函数 ($a = 2$)	S 型曲线
评估对象 1			
平均值 = 4.36	4.36	4.35	4.34
高级用户平均值 = 4.53	$\Delta = 0.05$	$\Delta = 0.04$	$\Delta = 0.03$
一般用户平均值 = 4.31			
评估对象 2			
平均值 = 4.07	4.1	4.04	4.02
高级用户平均值 = 4.41	$\Delta = 0.13$	$\Delta = 0.07$	$\Delta = 0.06$
一般用户平均值 = 3.97			
评估对象 3			
平均值 = 3.99	3.95	3.99	4
高级用户平均值 = 3.78	$\Delta = 0.11$	$\Delta = 0.06$	$\Delta = 0.06$
一般用户平均值 = 4.06			
评估对象 4			
平均值 = 3.75	3.57	3.70	3.74
高级用户平均值 = 3.33	$\Delta = 0.43$	$\Delta = 0.30$	$\Delta = 0.26$
一般用户平均值 = 4.00			

应用到某个具体的社交网络中时，需要考虑网络的具体特性以及用户倾向、选择参数以及社交关系权重模型的选择。本方案的应用可以让用户得到的信任评估结果与相似背景的用户更为一致，使得信任评估的结果更为可靠。

3.2　信任监测

首先进行信任检测模块的仿真，测试信任监测会对信任评估结果的影响。根据大众点评网特点，首先本文有以下两个假设: a) 恶意用户的数量比较有限; b) 恶意用户会采用最极端的评分值来影响结果。根据大众点评网的情况，给出了五种可能的信任值变化情况，具体如下:

a) **下降型:** 3.8，3.5，3.5，3.2，3.1，3，3，3.1，3。

b) **上升型:** 3.6，3.8，3.8，4，3.9，4.1，4.2，4.2，4.4，4.5。

c) **稳定型:** 4.1，4.2，4.3，4.2，3.9，3.6，4，4.1，4，3.9。

d) **恶意攻击后：** 4，4，4，4，4.6，4.8，4.1，4.1，4，3.9。

e) **恶意攻击中:** 4.3，4.2，3.9，3.6，4，4.1，4，3.9，4.5，4.7。

表 4 为经过信任监测模块后，校正的信任值随着历史信任值控制参数 α 变化的情况。

表 4　信任修正值随 α 的变化情况

变化情况	0.95	$\Delta_{0.95}$	0.8	$\Delta_{0.8}$	0.5	$\Delta_{0.5}$	0.2	$\Delta_{0.2}$	Δ'
下降型	3.27	0.27	3.15	0.15	3.04	0.04	3.02	0.02	1
上升型	4.05	0.45	4.21	0.29	4.39	0.11	4.47	0.03	0.9
稳定型	4.02	0.12	3.98	0.08	3.95	0.05	3.92	0.02	0.2
恶意攻击后	4.15	0.25	4.12	0.22	4	0.1	3.92	0.02	0.9
恶意攻击中	4.14	0.54	4.23	0.47	4.46	0.24	4.64	0.06	1.1

表 4 表示随着参数选择，对信任评估值波动所能作出修正的程度。其中，第一行数字表示 α 的取值，该列代表当 α 取该值时，方案运行后的信任值; Δ_α 为 α 取某个值时信任值的修正量，该列表示运行方案后的信任值结果与运行前的信任值差值绝对值; Δ' 的一列表示运行方案前，信任值变化的最大值。

接下来测试了信任监测每一步的表现，表 5 ~ 7 分别针对“稳定型”“恶意攻击后”与“恶意攻击中”三种情况，每经过一个时间窗后，信任值相较于之前时间窗的修正情况。第一列表示不同 α 取值下，十个时间窗分别的信任值波动大小。第一列由于没有之前对比，所以没有对比数据。

表 5　“稳定型”每步时间窗的信任波动值

α	1	2	3	4	5	6	7	8	9	10
0.9	/	0.04	0.04	0.01	0.02	0.05	0.03	0.01	0.01	0.01
0.8	/	0.04	0.04	0.01	0.03	0.06	0.03	0.01	0.01	0.01
0.5	/	0.05	0.05	0.01	0.04	0.08	0.03	0.01	0.01	0.01
0.2	/	0.05	0.05	0.00	0.05	0.09	0.02	0.00	0.00	0.01

表 6　“恶意攻击后”每步时间窗的信任波动值

α	1	2	3	4	5	6	7	8	9	10
0.95	/	0.00	0.00	0.00	0.06	0.07	0.03	0.01	0.00	0.01
0.8	/	0.00	0.00	0.00	0.07	0.09	0.03	0.01	0.00	0.01
0.5	/	0.00	0.00	0.00	0.09	0.11	0.03	0.00	0.01	0.02
0.2	/	0.00	0.00	0.00	0.11	0.11	0.00	0.01	0.02	0.03

表 7　“恶意攻击中”每步时间窗的信任波动值

α	1	2	3	4	5	6	7	8	9	10
0.95	/	0.04	0.08	0.10	0.04	0.01	0.01	0.01	0.01	0.03
0.8	/	0.04	0.09	0.11	0.04	0.01	0.00	0.01	0.02	0.04
0.5	/	0.05	0.10	0.12	0.03	0.00	0.00	0.01	0.03	0.05
0.2	/	0.05	0.11	0.13	0.02	0.01	0.00	0.01	0.05	0.06

由表 5 ~ 7 可得，即便对于没有较多评论周期的目标对象，

信任监测模型已经可以较好地修正信任值的变化。整体实现了修正恶意行为信任值变化的目的，增强了方案的鲁棒性。

4 结束语

本文提出了一种社交网络信任评估方法。依据每个因素影响大小建立用户向量，利用数据聚类方法来表现用户信息基于相似度聚类，让相似的用户的意见变得更为重要，以进行社交网络的信任评估。信任监测中，通过历史信任聚合与阈值监测，保障目标的信任值不会受到恶意用户行为的剧烈影响。信任评估测试中，测试了信任监测的参数选择，并对方案在实际社交网络中的表现进行性能评估，分析几种关系圈中量化曲线的关系。

后续还可以在以下方面开展进一步研究：a) 在信任评估中引入时间衰减因素，并量化其影响大小；b) 考虑加入更多的因素，并可以让用户主动选择信任评估的参数与方案等。

参考文献：

[1] WikiPedia. Social network [EB/OL]. [2016]. https://en.wikipedia.org/wiki/Social_network.

[2] Kathy. Reasons why v-divide can succeed [EB/OL]. [2011]. http://tech.qq.com/a/20110302/000153.htm.

[3] Walther O J. Social network analysis and informal trade [R]. [S.l.]: University of Southern Denmark, 2015.

[4] LinkedIn [EB/OL]. [2016-07-18]. https://www.linkedin.com/.

[5] Instgram [EB/OL]. [2016-07-18]. www.instgram.com/.

[6] Tumblr [EB/OL]. [2016-07-18]. https://tumblr.com/.

[7] Skoudis E. What are the risks of social networking sites? [EB/OL]. http://http://searchsecurity.techtarget.com/answer/What-are-the-risks-of-social-networking-sites.

[8] Caverlee J, Liu Ling, Webb S. The SocialTrust framework for trusted social information management: architecture and algorithms [J]. Information Sciences, 2010, 180(1): 95-112.

[9] Golbeck J A, Hendler J. Computing and applying trust in Web-based social networks [D]. Maryland: University of Maryland, 2005.

[10] Avesani P, Massa P, Tiella R. A trust-enhanced recommender system application: Moleskiing [C]//Proc of ACM Symposium on Applied Computing. 2010: 1589-1593.

[11] Lesani M, Bagheri S. Fuzzy trust inference in trust graphs and its application in semantic Web social networks [C]//Proc of Automation Congress. 2006: 1-6.

[12] Lin Chuncheng, Lin T S, Liu Wanyu. A trust and distrust mechanism for a social network-based recommendation system [C]//Proc of the 15th International Symposium on Wireless Personal Multimedia Communications. 2012: 172-176.

[13] Jiang Wenjun, Wang Guojun, Wu Jie. Generating trusted graphs for trust evaluation in online social networks [J]. Future Generation Computer Systems, 2014, 31(1): 48-58.

[14] Deng Shuiguang, Huang Longtao, Xu Guandong. Social network-based service recommendation with trust enhancement [J]. Expert Systems with Applications, 2014, 41(18): 8075-8084.

[15] Chang Weilun, Diaz A N, Hung P C. Estimating trust value: a social network perspective [J]. Information Systems Frontiers, 2015, 17(6): 1381-1400.

[16] Caverlee J, Liu Ling, Webb S. Socialtrust: tamper-resilient trust establishment in online communities [C]//Proc of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries. 2008: 104-114.

[17] Zhang Zhiyong, Wang Kanliang. A trust model for multimedia social networks [J]. Social Network Analysis & Mining, 2013, 3(4): 969-979.

[18] Xia Hui, Jia Zhiping, Li Xin, *et al*. Trust prediction and trust-based source routing in mobile ad hoc networks [J]. Ad Hoc Networks, 2013, 11(7): 2096-2114.

[19] Nepal S, Sherchan W, Paris C. STrust: a trust model for social networks [C]//Proc of the 10th International Conference on Trust, Security and Privacy in Computing and Communications. 2011: 841-846.

[20] Moore Madison. Online friends vs. real life friends: a comparison [EB/OL]. (2012). http://thoughtcatalog.com/madison-moore/2012/12/online-friends-vs-real-life-friends-a-comparison/.

[21] Kim Y A, Ahmad M A. Trust, distrust and lack of confidence of users in online social media-sharing communities [J]. Knowledge-Based Systems, 2013, 37(2): 438-450.

[22] Ortega F J, Troyano J A, Cruz F L, *et al*. Propagation of trust and distrust for the detection of trolls in a social network [J]. Computer Networks the International Journal of Computer & Telecommunications Networking, 2012, 56(12): 2884-2895.

[23] 孔英会，苑津莎，张铁峰，等. 基于数据流管理技术的配变负荷分类方法研究 [C]//2006 中国国际供电会议论文集. 2006.

[24] 李戈，邵峰晶，朱本浩. 基于神经网络聚类的研究 [J]. 青岛大学学报：工程技术版，2001，16(4)：21-24.

[25] FCM 聚类算法介绍 [EB/OL]. (2012). http://www.cnblogs.com/kemaswill/archive/2012/11/01/2749422.html.

[26] 大众点评网 [EB/OL]. [2016-07-18]. www.dianping.com/.

[27] 查金祥，王立生. 网络购物顾客满意度影响因素的实证研究 [J]. 管理科学，2006，19(1)：50-58.

(上接第 506 页)

[2] Eggeling T, Werner R F. Hiding classical data in multipartite quantum states [J]. Physical Review Letters, 2002, 89(9): 097905-097913.

[3] Hayden P, Leung D, Smith G. Multiparty data hiding of quantum information [J]. Physical Review A, 2005, 71(6): 062339.

[4] Gea-Banacloche J. Hiding messages in quantum data [J]. Journal of Mathematic Physics, 2002, 43(9): 4531-4536.

[5] Shaw B A, Brun T A. Quantum steganography with noisy quantum channels [J]. Physical Review A, 2011, 83(2): 498-503.

[6] Liao Xin, Wen Qiaoyan, Song Tingting, *et al*. Quantum steganography with high efficiency with noisy depolarizing channels [J]. IEICE Trans on Fundamentals of Electronics Communications & Computer Sciences, 2013, E96-A(10): 2039-2044.

[7] Martin K. Steganographic communication with quantum information [C]//Proc of the 9th International Conference on Information Hiding. Berlin: Springer, 2007: 32-49.

[8] Qu Zhiguo, Chen Xiubo, Luo Mingxing, *et al*. Quantum steganography with large payload based on entanglement swapping of χ-type entangled states [J]. Optics Communications, 2011, 284(7): 2075-2082.

[9] Zhang Yi, Lu Kai, Gao Yinghui, *et al*. NEQR: a novel enhanced quantum representation of digital images [J]. Quantum Information Processing, 2013, 12(8): 2833-2860.

[10] Wang Shen, Sang Jianzhi, Song Xianhua, *et al*. Least significant qubit (LSQb) information hiding algorithm for quantum image [J]. Measurement, 2015, 73(9): 352-359.

[11] 王宁，林崧. 基于最低有效位的量子图像水印 [J]. 量子电子学报，2015，32(3)：263-269.