

基于因子图模型的用户可信度评估

白 昀, 蔡皖东

(西北工业大学 计算机学院, 陕西 西安 710072)

摘要: 为克服传统方法在缺少用户个人信息及发布内容的情况下无法有效地评估用户可信度问题, 提出基于评论反馈信息和信任关系的用户可信度因子图模型(UFGM)。该模型将信任关系和评论反馈对用户可信度评估的影响形式化为一个概率模型, 并提出半监督分类的学习方法构建模型。在Extended Epinions数据集上验证了模型的有效性, 并发现信任关系比评论反馈更易对用户可信度评估产生积极影响。与传统算法相比, UFGM能在缺少用户描述信息及评论信息的情况下将用户可信度评估的准确度提高12%~29%。

关键词: 社交网络; 因子图; 用户可信度; 信任关系

中图分类号: TP391

文献标志码: A

User Credibility Evaluation Method Based on Factor Graph Model

BAI Yun, CAI Wandong

(School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: In order to overcome the inability of traditional methods to effectively evaluate user credibility in the absence of user personal information and published content, a user credibility factor graph model (UCFGM) based on comment feedback and trust relationship is proposed, which formalized the influence of trust relationship and comment feedback on user credibility evaluation into a probability model, and proposed a semi-supervised classification learning method to build the model. The effectiveness of the model is verified on the Extended Epinions dataset, and it is found that the trust relationship is more likely to have a positive impact on user credibility evaluation than comment feedback. Compared with traditional algorithms, UCFGM can improve the accuracy of user credibility evaluation by 12% to 29% in the absence of user description and comment information.

Key words: social networks; factor graph; user credibility; trust relationship

社交网络平台(如Twitter、Facebook、E-bay、Wikipedia、E-pinions、微博、微信、QQ等)在全球拥有数量众多的用户, 用户通过这些社交平台获取信息的同时也生成各种信息(如多媒体文件和博客文章等)。许多平台开始使用用户生成内容的方式提供服务, 由于网络的开放性, 网络用户行为也存在复杂性和不确定性, 这就使得伪造评论、刻意灌水或恶意诋毁的现象越来越明显。目前, 网民面对海量信息(如零售商家、转发帖子、推荐文章和推荐用户)大多采用查看信誉可信度的方式来进行选择, 用户可信度评估的准确性就决定了网民选择的正确度。研究用户可信度有助于发现虚假评论发布者, 及时处理发布虚假内容, 维护网站信誉并增强用户使用信心。另外, 通过用户可信度的评估辅助用户进行信息获取及交易, 降低商业活动中的风险。因此, 如何判断这些用户的可信度是目前急需解决的问题。

国内外学者对用户可信度评估进行了大量研究, 目前主要有以下几种网络用户可信度评价模型: ①根据用户评论的可信度来确定用户的可信度。如: Shang等^[1]采用拉格朗日算法计算用户可信度。Mukherjee等人^[2]提出GS-Rank算法可以对虚假评论发布者群进行检查。文献[3]用Review Graph的方法对编造虚假评论信息进行检测, 构造出评论诚实度和评论者可信度的关系曲线来判断评论者可信度。②通过加入用户的社交关系^[4-6](如: Facebook中的好友关系, Twitter中的追随关系以及E-pinions中的信任关系)来计算用户的信任度, 但这些算法

收稿日期: 2020-10-20

基金项目: 陕西省教育厅专项科学研究计划(19JK0526); 榆林市科技计划项目(2016-24-4); 陕西省科技厅专项科学研究计划(2021JQ-576)

第一作者: 白 昀(1985—), 女, 讲师, 工学博士, 主要研究方向为社交网络的信任模型以及概率图。

E-mail: 395943619@qq.com



论文
拓展
介绍

不能缺少用户个人信息提取或者评论内容提取,特征选取和特征提取的准确性对最终评价结果有明显的影响。Ye和Akoglu^[7]基于评价者与产品构成的二分图提出GroupStrainer算法,发现虚假评论发送者。Choo^[8]通过对用户交互做情感分析检测虚假评论发送者群组,其对活跃群体检测效果更佳。Wang等^[9]通过二分图投影方法解决某些产品无群组成员评论的问题,提出检测虚假评论发送者群组的方法。Ayday^[10]提出基于信任传播的迭代信任和信誉管理模型(BP-ITRM),将信誉管理问题描述为从复杂的多变量全局函数中计算边缘概率分布的问题,利用置信传播算法对边缘概率分布进行计算,有效地评估了服务提供者的可信度。③利用稻田算法^[11]和机器学习^[12]的方法对用户信任度进行研究。例如:Gupta等^[13]首先运用回归分析法找到预测可信度的有关特征,然后通过相关反馈与机器学习相结合的方式将微博中的信息进行可信度得分排序。Josang等^[14]提出了一种基于贝叶斯网络的Beta声誉信任模型,用二元评分(诚实的或不诚实的)来更新Beta概率密度分布函数,以评估用户的信任度。Wang^[15]结合图论知识,分析每个用户的出度和入度,来评估用户的可信度值。Wang^[16]利用途中节点间的交互提出基于图的迭代模型来检测虚假评论发布者。Chu等^[17]利用随机森林算法、构造信息熵结合机器学习的方法,提出虚假用户识别算法。Zhang等^[18]提出构建基于异构产品评论网络的无监督学习模型以此鉴别不可信评论。还有一些学者借助复杂网络分析用户评价的可信度。

目前常见的研究方法存在以下不足之处:①对于影响用户评价可信性的决策因子考虑不全面;②对于用户个人信息、用户行为或评论内容无法获得的这种情况,评估结果的准确性会大大下降;③特征选取和特征提取的准确性往往对评估结果有很大的影响;④有标注的数据相对较少,而大量无标注的数据在算法中没有得到有效利用。

针对现有评估方法不足,提出一种基于评论反馈信息和信任关系的用户信任度因子图模型。将用户在社交网络上发表的评论进行分析,该模型利用大量未知标签的用户数据进行半监督分类,充分挖掘社交网络中隐藏的信任关系,同时考虑了评判者可信度,能有效避免灌水和恶意诋毁的现象,在缺少用户个人信息和评论内容的情况下有效地进行用户可信度评估。

1 基于因子图模型的用户可信模型构建

1.1 信任模型相关术语的定义

可信度指对人或事物可以信赖的程度。用户通过以往的历史经验对现在的特定事物或对象赋予一定的信赖程度。在本文中,用户可信度表示用户在社交网络中的真实性和其相应行为的可信性大小,具有高可信度的用户通常拥有真实的信息资料、高质量发布内容,并且与可信用户具有社会关系。

解决用户可信度问题有利于发现领域专家,具有影响力的人物,还可以帮助人们找到真实可靠的评论信息。另外,电子商务网站通过用户可信度的评估辅助用户进行交易,从而降低商业活动中的风险。以下为相关术语的定义。

定义1 虚假评论(opinion spam):Jindal等^[19]首次提出虚假评论的概念。在本文中,将不真实的评论(具有夸大或诋毁现象的评论)认为是虚假评论,而这样的评论通常与大部分的其他评论相反。

定义2 虚假评论发布者(review spammer):Jindal等^[19]指出虚假评论一般由一个固定群体发出,这个群体的成员都是虚假评论发布者,即指做出虚假评论的人。而在这种群体中的成员彼此之间更容易有信任关系。

定义3 评判者:社交网络用户对网中其他用户发布的评论,依据中肯性、合理性和可信性进行综合评分,本文将这种给出综合评分的用户称之为评判者。用户不能对自己发布的评论进行综合评分。

定义4 虚假评判者:指恶意诋毁或有意抬高评论真实性的评判者。

1.2 可信评估建模分析

参考一个基于动态连续的因子模型图Mood Cast方法来建模和预测用户可信度,有评论的发布者(以下简称用户)集合 $U = \{u_1, u_2, \dots, u_n\}$,共有 n 位评论发布者。用户 u_i 可信度记为 y_i ,用户可信度集合 $Y = \{y_1, y_2, \dots, y_n\}$,其中部分用户的可信度已知 $y_i = 1$ 或 0 ,一些用户的可信度未知,需求这些用户的 y_i 。用户集中的每位用户对应发布的评论集 $V = \{V_1, V_2, \dots, V_n\}$,其中用户 u_i 发布的评论集为 $V_i = \{v_1, v_2, \dots, v_m\}$, m 为用户 u_i 发布评论的篇数, V 为所有评论的数据集,用户 u_i 发布的评论集 V_i 的可信度记为 $y_{v,i}$ 。评判者对评论的综合评分矩阵 $R =$

$\{r_{v_1, p_1}, r_{v_1, p_2}, \dots, r_{v_n, p_k}\}$, 其中 r_{v_i, p_k} 表示评判者 p_k 对评论 v_i 给出的评分, 评分分为5个等级, 其分值为1分、2分、3分、4分、5分。

有评论网络图 $G = (V, E)$, 其中每个用户的评论集作为节点 $V = \{V_1, V_2, \dots, V_n\}$, 用户之间的信任关系作为边, 即用户 u_i 和用户 u_j 发布的评论集节点分别 V_i 和 V_j , 用户 u_i 信任用户 u_j , 则 V_i 和 V_j 之间存在边 e_{ij} , 边集 $E = \{e_{ij} | (i, j \text{ 分别为 } 1, 2, \dots, n)\}$ 。用户被信任的频次集合为 $C = \{c_1, c_2, \dots, c_n\}$ 。

给定如上条件和定义, 可以将研究的问题定义如下。

问题: 给定具有评判信息的评论网络图 $G = (V, E, R)$ 以及部分用户具有信任标签 $y_i = 1$ 或 0 , 如何判断未知信任标签用户的信任标签 y_i 。

2 用户信任度模型框架

本文目标是通过用户对用户的评论记录和用户之间

的信任关系来评判用户的可信度。在此提出一个基于评论和信任关系的用户因子图模型。

2.1 基于评论反馈信息和信任关系的用户信任度因子图模型

采用概率图方法建立一个基于评论和信任关系的用户因子图模型 (user credibility factor graph model based on comments feedback and trust relationships, UCFGM), 将用户信任度问题建模到一个统一的框架中。

提出一个基础的因子模型如图1所示, 左图包含了5个用户 $\{u_1, u_2, \dots, u_5\}$ 以及用户之间的信任关系, 图中 u_1 与 u_2 之间的箭头表示 u_1 信任 u_2 。右图是将左图作为输入建立的用户可信度因子图模型, 图中观察变量是网络中给定的用户所发表的评论集 $\{V_1, V_2, \dots, V_5\}$, 图中隐变量是用户可信度 $\{y_1, y_2, \dots, y_5\}$, 该图中定义了2组因子: 用户可信度与用户评论集的因子, 用函数 $f(y_i, V_i)$ 表示; 用户可信度与信任关系的因子, 用函数 $g(y_i, y_j)$ 表示。本文围绕这个模型图开展进一步研究。

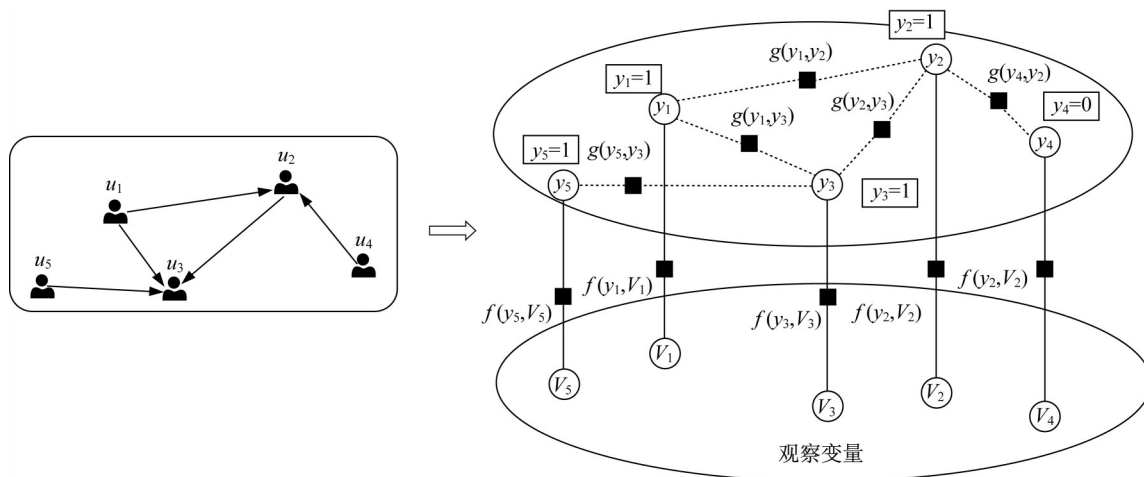


图1 用户可信度因子图模型的图结构

Fig. 1 Graphical representation of UCFGM

2.1.1 用户可信度与评论反馈的因子函数

假设用户的可信度与其发布的所有评论的可信度有关, 评价用户的可信度是一个复杂的过程, 需要借助复杂网络模型, 假设评论可信度越高那么用户可信度越高。文献[3]中提出评论可信度和用户可信度符合逻辑斯蒂曲线。以此类推, 一位用户的可信度与其评论集的可信度也符合逻辑斯蒂曲线, 即随着用户 u_i 的总评论可信度不断增高, 用户 u_i 的可信度越增越慢, 最终趋于一个稳定的值, 则有

$$f(y_i, V_i) = \frac{2}{1 + e^{-y_{v,i}}} - 1 \quad (1)$$

式中: $y_{v,i}$ 表示用户 u_i 发布的所有评论的可信度总和。

根据评分矩阵, 可得评论的反馈信息。这些反馈信息体现了其他用户对评论可信度的评价, 可直接作为评论可信度指标依据。那么对某一用户所发布的所有评论的整体可信度则是所有评论可信度的均值。评判者对评论的反馈反映出评论的可信度。

当然这些评判者当中会有虚假评判者。依据常理,一个合理的评判者对评论给出的评分应与其他评判者给出的评分相差不大,而虚假评判者所给出的评分往往与其他评判者给出的评分有较大差距。

假设有 s 位评判者 $\{p_1, p_2, \dots, p_s\}$ 对评论 v_i 的评论合理性打分,对应的评分 $(r_{v_i, p_1}, r_{v_i, p_2}, \dots, r_{v_i, p_s})$, 其平均评分为 $a_{\text{avg}} = \frac{r_{v_i, p_1} + r_{v_i, p_2} + \dots + r_{v_i, p_s}}{s}$, 该评分均值 a_{avg} 视为大多数的评判者认为该评论应具有合理分值。评判者 p_k 对评论 v_i 给出的评分,其评分的合理性 Q_{v_i, p_k} 可通过式(2)进行判别。若 $Q_{v_i, p_k} = 1$, 则评判者 p_k 对评论 v_i 给出的评分 $r_{v_i, p_k}^{D_k}$ 合理。反之,若 $Q_{v_i, p_k} = 0$, 则评分不合理。

$$Q_{v_i, p_k} = \begin{cases} 0, & |r_{v_i, p_k} - a_{\text{avg}}| > 1.5 \\ 1, & |r_{v_i, p_k} - a_{\text{avg}}| \leq 1.5 \end{cases} \quad (2)$$

评判者 p_k 对 t 个评论进行评判,即 $\{v_1, v_2, \dots, v_t\}$, 对应的评分为 $(r_{v_1, p_k}, r_{v_2, p_k}, \dots, r_{v_t, p_k})$, 其对应的评分合理性 $(Q_{v_1, p_k}, Q_{v_2, p_k}, \dots, Q_{v_t, p_k})$, 其中评分合理的个数为 $|Q_T|$, 评分不合理的个数为 $|Q_F|$, 则评判者可信度

$$y_{p_k} = \frac{2}{1 + e^{-\frac{|Q_T| - |Q_F|}{t}}} - 1 \quad (3)$$

由于评判者与评论之间存在关系:评判者的可信度越高,且对评论的评分越高,则评论的可信度越高。有 s 位评判者 $\{p_1, p_2, \dots, p_s\}$ 对评论 v_i 的评论合理性打分,对应的评分 $(r_{v_i, p_1}, r_{v_i, p_2}, \dots, r_{v_i, p_s})$, 根据评判者与评论之间的关系,这 s 位评判者认为评论 v_i 的可信度 y_{v_i} 为

$$y_{v_i} = \frac{\sum_{k=1}^{k=s} (y_{p_k} \times r_{v_i, p_k})}{s} \quad (4)$$

用户 u_i 发表了 D 篇评论,则用户 u_i 发布的所有评论的可信度总和,计算式如(5):

$$y_{v, i} = \sum_{i=1}^{i=D} \frac{\sum_{k=1}^{k=s} (y_{p_k} \times r_{v_i, p_k})}{s} \quad (5)$$

2.1.2 用户可信度与信任关系的因子函数

用户 u_i 所信任的用户集记为 $T_i = \{u_{i,1}, u_{i,2}, \dots, u_{i,l}\}$, 其中 l 是 u_i 所信任的用户数量。若用户 u_i 信任用户 u_j , (其中 $u_j \in T_i$), u_i 对 u_j 的信任与2个因素有关:① u_i 对 u_j 的信任程度 w_{ij} 。如果这2个用户共同信任的用户数比例高,则说明他们的信任相似度高(即同为可信用户,或同为不可信用户的概

率高),则 u_i 对 u_j 的信任程度高,则 $w_{ij} = \frac{|T_i \cap T_j|}{|T_i \cup T_j|}$,

其中 T_j 表示用户 u_j 所信任的用户集。② u_i 的可信程度。简化起见,表示为有多少人认为 u_j 是可信的,则可用 u_j 的信任频次(c_j)表示 u_j 的可信程度,信任频次越高,则说明 u_j 越值得信任。基于以上思想,可定义 $g(y_i, y_j)$ 为

$$g(y_i, y_j) = c_j w_{ij} \quad (6)$$

为了将所有的因子函数整合在一起,根据 Hammersley-Clifford 理论^[20] 可得目标函数

$$o(\theta) = \lg P_\theta(Y|G, R) = \sum_{i=1}^{i=n} \alpha f(y_i, V_i) + \sum_{e_{ij} \in E} \beta g(y_i, y_j) - \lg Z \quad (7)$$

式中: α 和 β 分别是不同因子函数的权重; $\theta = (\{\alpha\}, \{\beta\})$ 是由训练数据得到的参数配置; Z 是归一化因子,确保概率和为1。

2.2 模型学习

因子模型学习是寻找参数 $\theta = (\{\alpha\}, \{\beta\})$ 的配置,使得目标函数 $o(\theta)$ 的值最大。即

$$\theta^* = \arg \max_\theta o(\theta) \quad (8)$$

为了求解目标函数,采用梯度下降法。以 α 为例介绍如何学习参数。先得参数 α 到关于目标函数的梯度(式(9)),其中 $E[f(y_i, V_i)]$ 是在输入网中给定数据分布下的因子函数 $f(y_i, V_i)$ 的期望,即训练集数据中因子函数 $f(y_i, V_i)$ 的平均值。 $E_{P(y_i|G, R)}[f(y_i, V_i)]$ 是在评估模型给定 $P(y_i|G, R)$ 分布下的因子函数 $f(y_i, V_i)$ 的期望。对于 β 也可以得到相似的梯度(式(10))。

$$\frac{\phi(\theta)}{\partial \alpha} = E[f(y_i, V_i)] - E_{P(y_i|G, R)}[f(y_i, V_i)] \quad (9)$$

$$\frac{\phi(\theta)}{\partial \beta} = E[g(y_i, y_j)] - E_{P(y_i|G, R)}[g(y_i, y_j)] \quad (10)$$

由于输入网络含有环路,无法通过 Junction Tree 等方法直接计算边缘分布 $P(y_i|G, R)$ 。采用 Loopy Belief Propagation (LBP)^[21] 方法近似求解。理论上 LBP 不能保证收敛并且可能导致局部最大,但实践效果良好。具体算法为,先用 LBP 算法近似求解边缘分布 $P(y_i|G, R)$, 然后使用梯度下降法来

求解目标函数 $o(\theta)$ 。该算法是一个半监督学习算法。

参数学习算法为:

(1) 输入。基于评论的用户网络 $G=(V, E)$ 和评判者评分矩阵 R , 学习速率 η 。

(2) 输出。模型参数 $\theta = (\{\alpha\}, \{\beta\})$ 。

(3) 算法步骤。先初始化 θ , 然后重复如下步骤直至参数 θ 取值收敛。①根据LBP公式计算各个期望值;

②根据式(9)和式(10), 计算梯度 $\frac{o(\theta)}{\partial \theta}$ (如式: $\frac{\phi(\theta)}{\partial \alpha}$); ③使用学习效率 η 更新参数 θ 。以 α 为例,

$$\alpha_{\text{new}} = \alpha_{\text{old}} + \eta \frac{\phi(\theta)}{\partial \alpha}, \theta_{\text{new}} = \theta_{\text{old}} + \eta \frac{\phi(\theta)}{\partial \theta}。$$

2.3 模型推理

通过已经学习的参数 $\theta = (\{\alpha\}, \{\beta\})$, 可对未知信任度的用户, 通过寻找使目标函数最大化的用户信任度的配置, 即:

$$Y^* = \arg \max P_{\theta}(Y|G, R) \quad (11)$$

在该项工作中, 再次采用LBP算法来估算未知信任度标签用户的信任标签。通过计算用户的边缘分布函数 $P_{\theta}(y_i|G, R)$, 最后给每个用户分配一个最大概率的标签。该边缘分布函数最大值时的变量值 y_i^* 即是未知信任标签用户的信任标签。

3 实验分析

3.1 实验数据

采用Extended Epinions数据集对所提方法进行验证。数据是Massa从著名的产品评论网站Epinions.com上收集。该数据集中有用户信息132 000条, 有841 372条用户信任关系信息(1代表可信, -1代表不可信), 1 560 144条评论, 其中755 722条评论进行了评分, 评论的评分数据信息有13 668 320条。评判者对其他用户发表的评论进行评分, 评分为1~5, 表示评论的可靠性从低到高。数据集中还包括进行评分的时间信息和评分是否公开的数据。

本文的问题是: 基于信任关系以及评论的可靠性评分信息, 对评论发布用户的可信度进行评估。在数据集中, 14 701个用户被评价为不可信用户, 即标记为不可信用户, 69 900个用户被标记为可信用户, 剩余17 090个用户既有可信标记又有不可信标

记, 是需要评估的用户。从69 900个用户中将被信任频次 $c \geq 6$ 的6 234个用户标记信任标签为 $y_i = 1$, 从14 701个用户中将被认为不信任频次大于5的1 742个用户标记信任标签为 $y_i = 0$ 。其余用户为未标记信任标签用户。

3.2 度量指标

通过综合评价指标 F_1 、准确率 A 、精确率 P 和召回率 R 作为评估所提出模型的标准。各指标定义如下:

$$A = \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \quad (12)$$

$$P = \frac{t_p}{t_p + f_p} \quad (13)$$

$$R = \frac{t_p}{t_p + f_n} \quad (14)$$

$$F_1 = \frac{2 \times P \times R}{P + R} \quad (15)$$

式中: t_p 表示正类判定为正类; f_p 表示负类判定为正类; f_n 表示正类判定为负类; t_n 表示负类判定为负类。

3.3 有效性评价

根据历史评分数据, 在Epinions数据集上应用所提出的因子图模型进行用户可信度预测。本文进行10次交叉验证, 在每个交叉验证中, 分别在具有可信标记和不可信标记的用户集中随机选取10%的样本作为测试集, 其余90%的数据作为训练集, 进行用户可信度预测。因此, 就可以通过精确度来验证模型的有效性。图2给出了每次的准确率。平均预测准确率达到0.91以上, 最高时达到了0.93, 因此, 本文提出的因子图方法在预测用户可信度时是有效的。

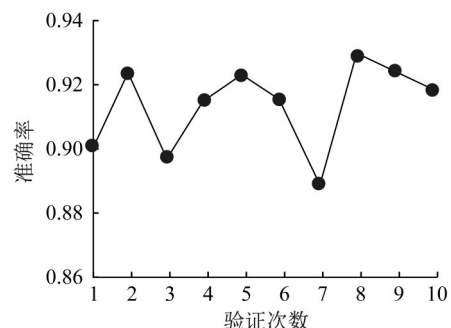


图2 因子图预测用户可信度的准确率

Fig. 2 Accuracy of factor graph in predicting user credibility

3.4 各因子对用户可信度预测的影响

在分析信任关系和评论的反馈评分在因子图模

型中所起的作用时,分别从模型中移除这2个因子,将移除后的模型与原模型进行对比。图3显示了原模型与分别移除部分因子后的 F_1 值。由图可知,移除任何一个因子都会造成 F_1 评估指标的显著降低,但不同因子对评估指标的影响力不同。这表示这些因素对用户可信度的预测具有积极作用,与不考虑这些因素的方法相比获得更好的预测效果。其中,移除信任关系因子后性能下降最大,说明相比于反馈信息,信任关系在用户可信度预测中起到更重要的作用。这是因为评论反馈中的虚假评判者对预测效果产生一定负面影响,而用户之间的信任关系在预测用户可信度时更直接。

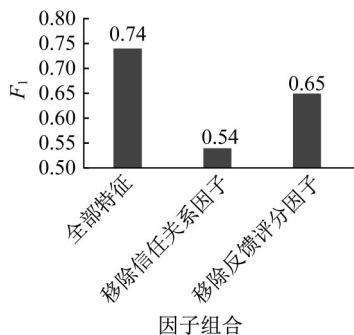


图3 模型移除不同因子对性能的影响

Fig. 3 Removal of impact of different factors on performance

3.5 对比方法和讨论

给定于评论的用户网络 $G = (V, E)$ 和评判者评分矩阵 R ,可构造训练数据集 $\{(x_i, y_i)\}_{i=1, \dots, n}$,其中 n 表示用户数, x_i 是关于用户 u_i 发布的评论集的特征向量, y_i 表示用户 u_i 是否可信,支持向量机(SVM)、逻辑斯特回归(LR)和朴素贝叶斯方法(NB)来训练分类模型并将其用于预测用户是否可信。SVM使用SVM-light。LR算法和NB算法使用weka工具包实现。与本文提出的因子图方法不同的是,分类模型没有考虑用户之间的信任关系。同时本文与基于PageRank可信度评估算法^[22]进行比较。PageRank算法将每个用户发布的信息可信度的平均值作为用户初始可信度,然后基于用户信任网计算PR值作为用户可信度。在本文中设置 $PR > 0.1$ 为可信用户,否则为不可信用户。

图4显示不同方法在Epinions数据集上的用户可信度评估性能。由图可见,本文提出的因子图模型优于其他4种方法。对于 F_1 ,LR算法在 F_1 评估指标上的值明显低于其他方法,这是因为LR算法一般

适用线性分类,而依据用户历史评分数据特征进行分类属于非线性分类,其效果较差。 F_1 评估指标上,因子图模型和PageRank方法比SVM、LR和NB 3种方法要提高8.5%~24%,这是因为SVM、LR和NB 3种方法均未考虑用户之间的信任关系,由此可以说明信任关系是预测用户可信度的决定因素之一;PageRank算法有效利用社交网络中的信任关系,从而与SVM、LR和NB相比提高了预测性能,但PageRank算法不能从概率论角度挖掘隐含的关系,因此其 F_1 评估指标比因子图方法低6.5%。又因为PageRank算法得到的是用户的可信度排序,该方法应用于用户可信度预测时准确度与其设置的可信度的阈值有关,本文的阈值设置方式有效防止将不可信用户判定为可信用户,因此在精确率和召回率两项指标中优于基准方法。另外,真实世界的用户之间的关联在社交网络中以用户关系网的形式体现,这种结构信息用概率图的形式表示更符合数据本身的结构,而其他4种算法虽能高效利用其他特征信息,但并不具备挖掘用户潜在关系的能力。还有一个重要的原因是本文提出的模型可以利用未知标签的用户数据进一步考虑数据集中的一些潜在关系。

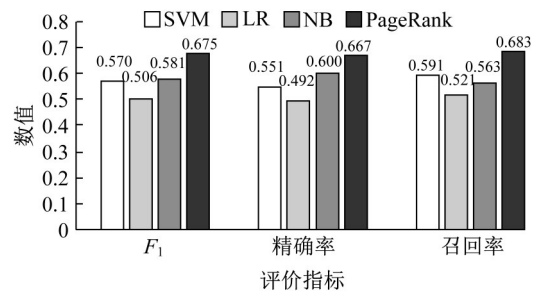


图4 不同方法的用户可信度预测性能

Fig. 4 Prediction performance of user credibility of different methods

4 结语

研究了在社交网络分析中长期关注的问题——用户的可信度评估。当前大部分的研究都聚焦在对用户个人信息和评论内容的特征提取上来评估用户可信度,本文将社交网站中评论反馈信息和用户信任关系对用户可信度评估的影响形式化为一个概率模型——基于评论反馈信息和信任关系的用户信任度因子图模型,该模型构建了评论反馈因子函数和信任关系因子函数,可在没有用户个人信息和评论内容情况下对用户可信度进行评估。提出半监督分类的学习方法构建模型,充分利用大量未知标签

的用户数据,并使用因子图方法充分挖掘社交网络中隐含的信任关系,同时对评判者与评论发布者的可信度进行评估,避免灌水和恶意诋毁的现象,有效提高了用户可信度评估精度。在 Extended Epinions 数据集上对所提出的模型进行验证,平均预测准确率达到0.91以上。通过实验发现移除任何一个因子都会降低用户可信度预测的准确性,且信任关系因子对用户可信度预测的积极影响明显高于评论反馈评分因子。

本文所提方法有效利用未知标签的用户数据,采用概率图表示用户之间的信息结构更符合真实世界的用户之间的关联,并且无需提供用户个人信息和评论内容即可有效预测用户可信度,为用户可信度研究提供了一个新思路。基于以上优势,将本文方法与其他4种传统方法进行性能对比,本文方法提出的用户因子图模型将用户可信度评估性能提高了12%~29%。

作者贡献说明:

白 昫:负责模型的设计与实现,并对模型的有效性进行验证。

蔡皖东:负责模型设计的总体指导。

参考文献:

- [1] SHANG F, LIU Y, CHENG J, *et al.* Fuzzy double trace norm minimization for recommendation systems [J]. *IEEE Transactions on Fuzzy Systems*, 2018, 26(4): 2039.
- [2] MUKHERJEE A, LIU B, GLANCE N. Spotting fake reviewer groups in consumer reviews [C]//*Proceedings of the 21st International Conference on World Wide Web*. Lyon: ACM, 2012: 191-200.
- [3] WANG G, XIE S, LIU B, *et al.* Review graph based online store review spammer detection [C]//*2011 IEEE 11th International Conference on Data Mining*. Vancouver: IEEE, 2011: 1242-1247.
- [4] CUI L, PI D, ZHANG J. Dmfasr: Deeper membership and friendship awareness for social recommendation [J]. *IEEE Access*, 2017, 5: 8904.
- [5] YAN S, LIN K J, ZHENG X, *et al.* An approach for building efficient and accurate social recommender systems using individual relationship networks [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2017, 29(10): 2086.
- [6] MAO M, LU J, ZHANG G, *et al.* Multi-relational social recommendations via multigraph ranking [J]. *IEEE Transactions on Cybernetics*, 2017, 47(12): 4049.
- [7] YE J, AKOGLU L. Discovering opinion spammer groups by net-work footprints [C]//*Joint European Conference on Machine Learning & Knowledge Discovery in Databases*. [S. l.]: Springer, 2015: 267-282.
- [8] CHOO E, YU T, CHI M. Detecting opinion spammer groups through community discovery and sentiment analysis [C]//*IFIP Annual Conference on Data and Applications Security and Privacy*. [S. l.]: Springer, 2015: 170-187.
- [9] WANG Z, HOU T, SONG D, *et al.* Detecting review spammer groups via bipartite graph projection [J]. *The Computer Journal*, 2016, 59(6): 861.
- [10] AYDAY E, FEKRI F. Iterative trust and reputation management using belief propagation [J]. *IEEE Transactions on Dependable and Secure Computing*, 2012, 9(3): 375.
- [11] WANG X, WANG X, CHE H, *et al.* An intelligent economic approach for dynamic resource allocation in cloud services [J]. *IEEE Transactions on Cloud Computing*, 2015, 3(3): 275.
- [12] JAYASINGHE U, LEE G M, UM T W, *et al.* Machine learning based trust computational model for IoT services [J]. *IEEE Transactions on Sustainable Computing*, 2019, 4(1): 39.
- [13] GUPTA A, KUMARAGURU P. Credibility ranking of tweets during high impact events [C]//*Proceedings of the 1st workshop on privacy and security in online social media*. Lyon: ACM, 2012: 2-8.
- [14] JOSANG A, ISMAIL R. The beta reputation system [C]//*Proceedings of the 15th Bled Electronic Commerce Conference*. Bled: AIS e-library, 2002, 5: 2502-2511.
- [15] WANG A H. Detecting spam bots in online social networking sites: A machine learning approach [J]. *Data & Applications Security & Privacy XXIV*, 2010, 6166: 335.
- [16] WANG G, XIE S, LIU B, *et al.* Identify online store review spammers via social review graph [J]. *ACM Transactions on Intelligent Systems & Technology*, 2012, 3(4): 1.
- [17] CHU Z, GIANVECCHIO S, WANG H, *et al.* Detecting automation of twitter accounts: Are you a human, bot, or cyborg? [J]. *IEEE Transactions on Dependable & Secure Computing*, 2012, 9(6): 811.
- [18] ZHANG Q, WU J, ZHANG P, *et al.* Collective hyping detection system for identifying online spam activities [J]. *IEEE Intelligent Systems*, 2017, 32(5): 53.
- [19] JINDAL N, LIU B. Opinion spam and analysis [C]//*Proceedings of the 2008 International Conference on Web Search and Data Mining*. Lyon: ACM, 2008: 219-230.
- [20] HAMMERSLEY J M, CLIFFORD P. Markov fields on finite graphs and lattices [J]. *Unpublished Manuscript*, 1971: 46.
- [21] MURPHY K P, WEISS Y, JORDAN M I. Loopy belief propagation for approximate inference: An empirical study [C]//*Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence*. San Francisco: Morgan Kaufmann Publishers Inc., 1999: 467-475.
- [22] GUPTA M, ZHAO P, HAN J. Evaluating event credibility on twitter [C]//*Proceedings of the 2012 SIAM International Conference on Data Mining*. Anaheim: Society for Industrial and Applied Mathematics, 2012: 153-164.