

跨社交网络的同一用户识别算法*

沈佳琪¹, 周国民²

(1.浙江工业大学 信息工程学院, 浙江 杭州 310023; 2.浙江警察学院 计算机与信息技术系, 浙江 杭州 310053)

摘要: 针对跨社交网络的同一用户识别问题, 提出了一种综合用户兴趣、写作风格和档案属性的识别方法。通过在这3种不同的特征维度下分别判定用户关系, 然后综合判定结果, 提高同一用户识别准确性。其中, 用户兴趣分为静态兴趣和动态兴趣, 静态兴趣采用 TextRank 算法从用户背景信息中提取, 动态兴趣则利用主题模型从用户发表的文本内容中挖掘出随时间变化的兴趣点。对于用户写作风格则通过 One-Class SVM 算法进行识别, 最后利用信息熵赋权法比较用户档案属性相似度。实验结果表明, 与传统机器学习算法相比, 所提算法精确率、召回率均有所提升。

关键词: 跨社交网络; 用户识别; 用户兴趣; 写作风格; 档案属性

中图分类号: TN01; TP391

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.211518

中文引用格式: 沈佳琪, 周国民. 跨社交网络的同一用户识别算法[J]. 电子技术应用, 2022, 48(1): 109-114.

英文引用格式: Shen Jiaqi, Zhou Guomin. User alignment across social networks[J]. Application of Electronic Technique, 2022, 48(1): 109-114.

User alignment across social networks

Shen Jiaqi¹, Zhou Guomin²

(1.College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China;

2.Department of Computer and Information Security, Zhejiang Police College, Hangzhou 310053, China)

Abstract: For the problem of identifying the same user across social networks, a recognition method that integrates user interests, writing style and profile attributes is proposed. By determining user relationships under these three different feature dimensions separately, and then synthesizing the results, the same user identification accuracy is improved. Among them, user interest is divided into static interest and dynamic interest, static interest is extracted from user background information by TextRank algorithm, while dynamic interest is mined from user published text content by using topic model to find out interest points that change over time. For user writing style, it is identified by One-Class SVM algorithm, and finally, the information entropy empowerment method is used to compare the similarity of user profile attributes. The experimental results show that the proposed algorithm has improved accuracy and recall rate compared with traditional machine learning algorithms.

Key words: across social networks; users identification; user interest; writing style; file attribute

0 引言

近年来, 个人信息数据随社交网络的普及变得越来丰富, 目前对社交网络的用户分析主要针对单一平台, 但由于单一平台数据存在局限性^[1], 因此可通过挖掘同一用户在不同社交网络中的多个账号, 为社交网络分析提供数据支撑^[2]。

基于用户档案属性的识别方式是研究最广的方法。Zafarani 等^[3]通过比较用户选取用户名的行为特征相似度判断是否为同一用户; Zhang 等^[4]结合用户名、头像等多个属性, 利用朴素贝叶斯进行识别。然而上述研究中

的特征容易缺失和伪造^[5]。因此, 一些研究从发表的文本内容入手, 挖掘用户兴趣, 比较兴趣相似度来判定用户关系^[6]。何力等^[7]采用 LDA 模型来挖掘文本内容中的用户兴趣; 吕志泉等^[8]在 LDA 模型的基础上引入了时间因子。但上述研究仅考虑了文本内容体现的动态兴趣, 没有结合静态兴趣, 同时, 即使是同一用户, 在不同社交平台关注和发表的内容也可能有较大差别, 这影响了用户识别效果。

针对上述问题, 本文提出一种综合用户兴趣、写作风格和档案属性的同一用户识别算法。该方法将用户静

* 基金项目: NSFC-浙江两化融合联合基金(U1509219); 浙江省自然科学基金委公益技术计划研究项目(LGF19F02006); 公安部科技强警基础工作专项(2018GABJC33, 2019GABJC36)

态兴趣和动态兴趣结合,得到综合兴趣判定用户关系;然后利用 One-Class SVM 算法识别用户写作风格相似度,写作风格为个人习惯,不随社交平台的变化而变化^[9];最后通过信息熵赋权法比较用户档案属性相似度,综合 3 种不同维度特征下的判定结果,得到用户最终归属关系。

1 相关理论

1.1 文本向量化与特征词提取

本文采用向量空间模型来表示文本,将文本 d_j 表示为 $d_j=(t_1, w_{j1}; t_2, w_{j2}; \dots; t_s, w_{js})$,其中 w_i 表示特征词 t_i 的权重,由 TF-IDF 算法得到。在传统 TF-IDF 算法中,一个词出现在多个文本中,则其不具有较强区分度,但在对用户发表的文本内容进行分析时,包含同一个词的文本数越多,则该词越能体现用户兴趣,故需要改进传统的 TF-IDF 算法,增加激励权重^[10],如下所示:

$$w_{ij} = \frac{w_{ij} + \frac{df_i}{n+0.01}}{\sqrt{\sum_{i=1}^s \left(w_{ij} + \frac{df_i}{n+0.01} \right)^2}} \quad (1)$$

其中, w_{ij} 表示特征词 t_i 在文本 d_j 中出现的频率, df_i 表示包含 t_i 的文本数, n 是文本总数。

1.2 LDA 主题模型

LDA 模型^[11]可用于提取文本内容中所包含的主题,如图 1 所示。

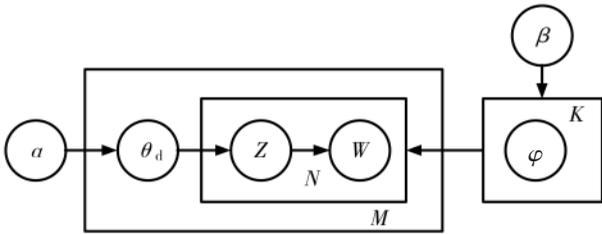


图 1 LDA 贝叶斯模型

其中, α, θ_d, β 和 φ 都是 Dirichlet 分布的参数, M 为文档数, K 为主题个数。

2 同一用户识别算法

本文算法流程如图 2 所示。

2.1 静态兴趣提取

静态兴趣指的是用户的长期兴趣,如用户个人简历、标签和职业等信息都很好地反映了用户的静态兴趣^[12]。其中,标签是用户主动选取的能代表自身兴趣的词。因此,若社交网络中直接提供了用户的标签信息,则以该信息作为用户的静态兴趣;若没有直接提供,则将用户的个人简历、职业和个性签名等信息合并到一个文本,利用 TextRank 算法提取关键词作为用户静态兴趣。

采用 Jaccard 相似度计算源社交网络用户 V_s 与目标用户 V_d 的静态兴趣相似度,如下所示:

$$\text{sim}_{\text{static}}(V_s, V_d) = \frac{|V_{s,\text{static}} \cap V_{d,\text{static}}|}{|V_{s,\text{static}} \cup V_{d,\text{static}}|} \quad (2)$$

其中, $V_{s,\text{static}}$ 和 $V_{d,\text{static}}$ 分别表示代表了用户静态兴趣

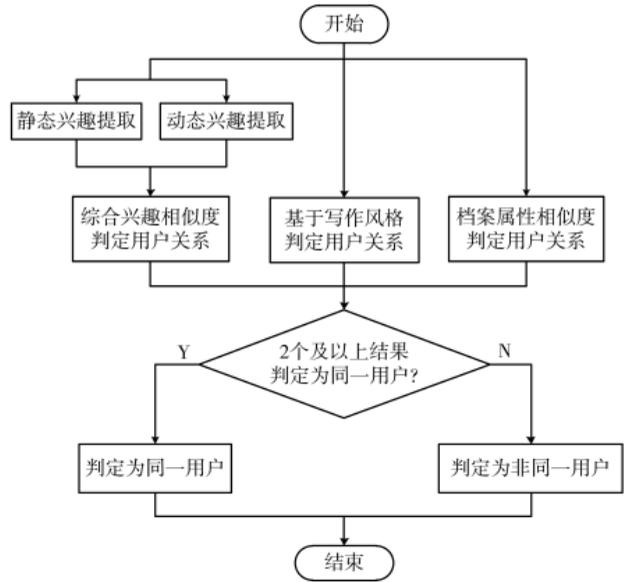


图 2 算法流程图

的词或短语。

2.2 动态兴趣提取

用户常会在社交网络中发表大量的文本内容,但同时也包含了部分与兴趣无关的噪声文本,因此,需要设计筛选器,滤除另一网络中与源用户 V_s 匹配概率较低的目标用户,再通过噪声文本过滤器滤除噪声,最后利用 Single-pass 聚类算法^[13]将文本内容按主题分为多个簇,引入时间因子计算用户动态兴趣分布的相似度。

2.2.1 用户筛选器设计

用户筛选器的设计应选取较为简单且填写完整度较高的用户属性,因此利用用户名来设计筛选器,采用编辑距离计算用户名之间的相似度:

$$\text{sim}_{\text{lev}}(V_s.\text{name}, V_d.\text{name}) = 1 - \frac{d_{\text{lev}}(V_s.\text{name}, V_d.\text{name})}{\max(\text{len}(V_s.\text{name}), \text{len}(V_d.\text{name}))} \quad (3)$$

其中, $V_s.\text{name}$ 表示源社交网络用户名, $V_d.\text{name}$ 表示目标社交网络用户名, $\text{len}(V_s.\text{name})$ 、 $\text{len}(V_d.\text{name})$ 表示用户名的编辑距离。利用计算得到的用户名相似度筛选另一网络中的所有目标用户 V_d 。

2.2.2 噪声文本过滤器设计

首先对用户发表的所有文本内容进行分词、去除停用词等预处理操作,统计预处理后的文本集中出现频率最高的前 T 个词,计算得到向量 F ,具体公式如下:

$$\begin{cases} F=(f_1, f_1, \dots, f_r) \\ f_i=\lambda_i \cdot \text{fq}(t_i) \end{cases} \quad (4)$$

其中, $\text{fq}(t_i)$ 是特征词 t_i 在文本集中出现次数, λ_i 表示特征词 t_i 对应的词性,定义如下:

$$\lambda_i = \begin{cases} 1.5 & \text{词 } t_i \text{ 为名词或动词} \\ 1.0 & \text{词 } t_i \text{ 为形容词} \\ 0.5 & \text{其他} \end{cases} \quad (5)$$

得到向量 F 后, 计算用户每条文本相应得分, 如下所示:

$$\text{Score}(d_j) = (w_{j1}', w_{j2}', \dots, w_{jT}') \cdot F^T \quad (6)$$

其中, w_i' 是改进的 TF-IDF 算法计算得到的特征词权重。若 $\text{Score}(d_j)$ 小于阈值, 则将其视为噪声文本舍去。

2.2.3 LDA 模型提取用户兴趣

用户在社交平台中发表的单条文本字数较少, 难以直接挖掘主题, 故先利用聚类算法, 将同一用户的文本内容按话题分为多个簇, 将同一簇内的文本合并为一个文档。

首先利用 Single-pass 聚类算法对源用户 V_s 和候选目标用户 V_d 发表的文本内容分别进行聚类, 得到 V_s 的簇文本集 $W_u^* = \{w_{u1}^*, w_{u2}^*, \dots, w_{uQ}^*\}$ 和聚类个数 N_s 与另一网络中的候选目标用户 V_d 所对应的各簇文本集 $\{W_j^*\}_{j=1}^N$; 然后将源用户 V_s 的簇文本集 W_u^* 与候选目标用户 V_d 的簇文本集 W_j^* 合并, 通过 Single-pass 聚类算法得到聚类个数 K , 作为 LDA 模型的参数提取文本-主题矩阵, 并引入基于时间的遗忘因子, 如下所示:

$$\gamma = e^{-h \times (\text{nowTime} - \text{lastTime})} \quad (7)$$

其中, h 是调节系数, 用来调节衰减的速度; nowTime 是当前系统时间, lastTime 是每一个聚类簇对应的最近一次更新时间。

可得源用户 V_s 与候选目标用户 V_d 的遗忘因子与文本-主题矩阵的对应关系如下:

$$\begin{matrix} & T_1 & T_2 & \dots & T_K \\ \gamma_1 & w_1^* & \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1K} \end{bmatrix} \\ \gamma_2 & w_2^* & \begin{bmatrix} p_{21} & p_{22} & \dots & p_{2K} \end{bmatrix} \\ \vdots & \vdots & \begin{bmatrix} \vdots & \vdots & & \vdots \end{bmatrix} \\ \gamma_N & w_N^* & \begin{bmatrix} p_{N1} & p_{N2} & \dots & p_{NK} \end{bmatrix} \end{matrix} \quad (8)$$

$$\begin{matrix} & T_1 & T_2 & \dots & T_K \\ \gamma_{N+1} & w_{N+1}^* & \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1K} \end{bmatrix} \\ \gamma_{N+2} & w_{N+2}^* & \begin{bmatrix} p_{21} & p_{22} & \dots & p_{2K} \end{bmatrix} \\ \vdots & \vdots & \begin{bmatrix} \vdots & \vdots & & \vdots \end{bmatrix} \\ \gamma_M & w_M^* & \begin{bmatrix} p_{M1} & p_{M2} & \dots & p_{MK} \end{bmatrix} \end{matrix} \quad (9)$$

根据上述对应关系, 可以求得主题 T_k 在用户 i 兴趣分布中的概率值 $P(T_k)$, 计算公式如下:

$$P(T_k) = \frac{\sum_{i=1}^N [\gamma_i \times p_{ik}]}{\sum_{k=1}^K \sum_{i=1}^N [\gamma_i \times p_{ik}]} \quad k=1, 2, \dots, K \quad (10)$$

其中, N 是用户发表的文本内容聚类后所得到的簇个数。

根据式(10)可得用户动态兴趣分布 $\theta = (P(T_1), P(T_2), \dots, P(T_K))$ 。利用 KL 散度计算用户动态兴趣分布的相似

度, 如下所示:

$$\text{sim}_{\text{dynamic}}(V_s, V_d) = 1 - D(\theta_s || \theta_d) \quad (11)$$

得到用户静态兴趣相似度和动态兴趣相似度后, 综合兴趣相似度可表示为:

$$\text{sim}_{\text{ms}}(V_s, V_d) = \alpha \cdot \text{sim}_{\text{static}}(V_s, V_d) + (1 - \alpha) \text{sim}_{\text{dynamic}}(V_s, V_d) \quad (12)$$

其中, α 是调节系数, $0 \leq \alpha \leq 1$ 。

将用户综合兴趣相似度与设定阈值进行比较, 若大于阈值, 则判定为同一用户, 反之则为非同一用户, 输出判定结果 R_{ms} 。

2.3 写作风格比较

本文从以下 6 个特征研究用户写作风格^[14], 特征提取流程如图 3 所示, 算法伪代码如下:

输入: 源社交网络用户 V_s 发表的文本内容 $(d_{s1}, d_{s2}, \dots, d_{sm})$, 目标用户 V_d 发表的文本内容 $(d_{d1}, d_{d2}, \dots, d_{dm})$, 比例阈值 δ ;

输出: 判定结果 R_{ws} 。

(1) for d_i in $\{d_{s1}, d_{s2}, \dots, d_{sm}\}$

(2) 提取 d_i 写作风格特征 θ_{wi}

(3) end for

(4) 将得到的 $\{\theta_{w1}, \theta_{w2}, \dots, \theta_{wm}\}$ 输入 One-Class SVM 算法训练得到分类器

(5) for d_i in $\{d_{d1}, d_{d2}, \dots, d_{dm}\}$

(6) 对 d_i 提取写作风格特征 θ_{wi}

(7) end for

(8) for θ_{wi} in $\{\theta_{w1}, \theta_{w2}, \dots, \theta_{wm}\}$

(9) 将 θ_{wi} 输入分类器

(10) if $(R_{\text{SVM}}=1)$ num=num+1

(11) end for

(12) if $\frac{\text{num}}{|m|} > \delta$ $R_{\text{ws}}=1$

(13) else $R_{\text{ws}}=-1$

其中, δ 为比例阈值, $R_{\text{SVM}}=1$, 则判定该条文本内容为同一用户所写, 反之则非同一用户所写。

2.4 档案属性相似度

用户档案属性包括用户名、头像、所在地等, 需要选

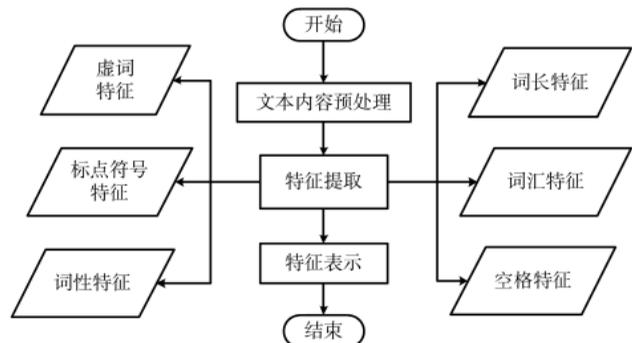


图 3 写作风格特征提取流程图

取多个不同的相似度计算方法,最大程度表示来自两个不同社交网络用户的相似程度。

(1)用户名特征。本文通过 Jaro-Winkler Distance 相似度、LCS 相似度、编辑距离和 Jaccard 相似度计算得到 4 维的用户名相似度向量 $N=(m_1, m_2, m_3, m_4)$ 。

(2)头像特征。不同社交网络对用户上传的头像格式要求并不相同,导致头像可能经过了压缩、模糊等操作。本文采用直方图相似度、SIFT 相似度和哈希相似度得到 3 维用户头像相似度向量 $A=(a_1, a_2, a_3)$ 。

(3)所在地特征。将所在地作为字符串考虑时,相似度定义为表 1 所示。

将所在地转换为对应的经纬度时,实际距离相似度计算公式定义为:

表 1 地理位置相似度

省	市	相似度
相同	相同	1
缺失	相同	3/4
相同	缺失	1/2
相同	不同	1/4
不同	/	0
缺失	缺失	0

$$\text{sim}_{\text{dist}}(V_s, V'_d) = 1 - \frac{d(V_s, V'_d)}{\pi R} \quad (13)$$

其中, $d(V_s, V'_d)$ 表示由大圆距离计算得到的用户所在地实际距离。

通过上述两种方法可得到 2 维的用户地理位置特征相似度向量 $L=(l_1, l_2)$ 。

将上述 3 种不同维度特征归一化,利用信息熵赋权法确定相应权重,得到用户档案属性综合相似度,如下所示:

$$\text{Sim}_{\text{attr}} = \sum_{i=1}^k (w_i \times S_i) \quad (14)$$

其中, S_i 表示通过第 i 种相似度计算方法得到的属性相似度值, w_i 表示第 i 种属性相似度的权重。

将用户档案属性综合相似度与设定阈值进行比较,输出判定结果 R_{attr} 。

3 实验结果与分析

3.1 数据集获取

本文选取微博和豆瓣平台用户作为实验对象。实验爬取了 24 623 名豆瓣用户数据,筛选其中主动提供了微博账号链接的用户,得到满足要求的用户共 1 562 人。获取这 1 562 人在豆瓣平台中相应的用户信息,并利用提供的微博链接定位到用户对应的微博账号,爬取相应的微博用户数据,最终得到 3 124 名来自新浪微博和豆瓣平台的用户及其数据信息。其中包含了 1 562 对已知的关联用户对,作为正样本,同时随机构建了相同数量的不属于同一用户实体的用户对作为负样本,用户发表的文本内容总计超过 32 万条。

3.2 评价指标

本次实验采用了精确率、召回率、F1 值作为算法性能的评价指标,分别记为 P 、 R 和 $F1$:

$$P = \text{tp} / (\text{tp} + \text{fp}) \quad (15)$$

$$R = \text{tp} / (\text{tp} + \text{fn}) \quad (16)$$

$$F1 = 2 \times P \times R / (P + R) \quad (17)$$

其中, tp 表示正确预测为同一用户的账号对数, fp 表示错误预测为同一用户的账号对数, fn 表示错误预测为非同一用户的账号对数。

3.3 实验结果

3.3.1 确定用户兴趣权重取值

本文在仅考虑用户综合兴趣相似度的情况下,测试不同系数 α 对用户识别效果的影响,同时与未滤除噪声文本的情况对比,如图 4 所示。

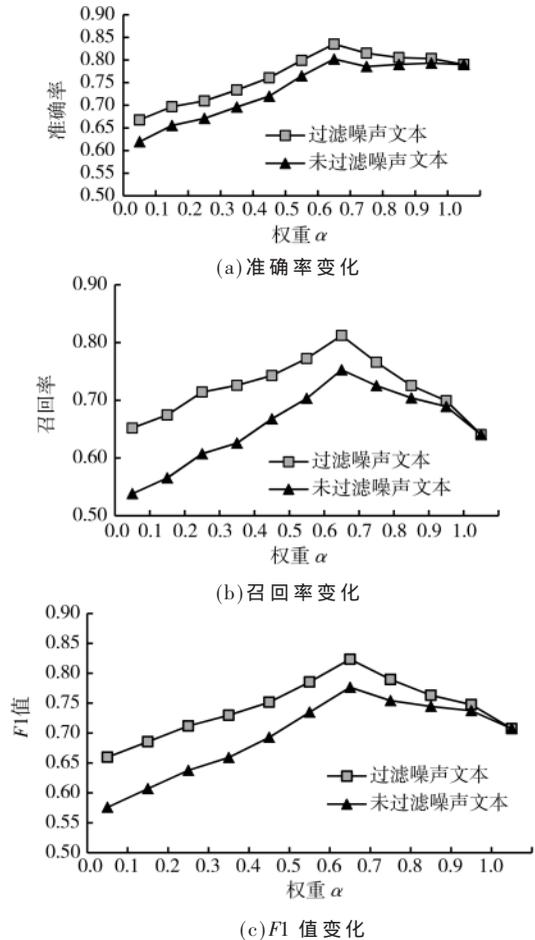


图 4 不同权重 α 下的性能参数变化

由图 4 可得,过滤噪声文本后的算法性能参数均优于对比情况,且当 $\alpha=0.6$ 时,算法效果最好。这是因为,当权重 $\alpha=0$ 时,同一用户判定完全取决于用户动态兴趣,此时滤除不能体现用户兴趣的噪声文本有利于更准确地提取用户兴趣;当 $\alpha=1$ 时,用户判定结果只由用户静态兴趣决定,此时两者计算方法相同,性能参数相等。

3.3.2 文本分类算法对比

本文把利用写作风格判定用户关系看作是一个文本分类问题。因此,为了验证 One-Class SVM 算法的分

类效果,选取了2种常用的文本分类算法KNN和Naïve Bayes进行对比,实验结果如表2所示。

从表2可得,利用One-Class SVM算法识别用户写作风格的效果优于对比算法。

表2 不同分类算法性能对比

分类算法	P	R	F1
KNN	0.583 9	0.548 2	0.565 5
Naïve Bayes	0.624 5	0.657 3	0.640 4
One-Class SVM	0.724 3	0.743 2	0.733 6

3.3.3 确定档案属性相似度阈值

为确定档案属性相似度阈值,实验统计了在不同档案属性相似度阈值下识别出的同一用户中实际的正样本和负样本所占比例,如图5所示。

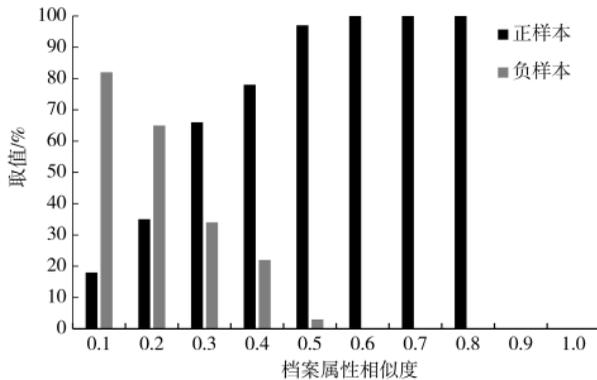


图5 档案属性相似度统计分布

由图5可得,当档案属性相似度超过0.6后识别出的所有用户对都是正样本;当档案属性相似度大于0.8后,已经没有满足条件的用户对。因此,取档案属性相似度阈值为0.6较合适。

3.3.4 不同维度特征下性能比较

为验证结合3个维度特征的判定结果后是否真正提升了算法模型的识别效果,本文设计了如表3所示的验证实验。

表3 不同维度特征下算法性能比较

不同维度特征	P	R	F1
综合用户兴趣	0.835 2	0.812 1	0.823 5
写作风格	0.724 3	0.743 2	0.733 6
档案属性	0.774 6	0.735 1	0.754 3
综合3种维度特征	0.854 3	0.842 1	0.848 2

由表3可以看出,在综合了3种维度特征的判定结果后,算法模型的效果优于只使用单一维度特征,证明了综合多个维度特征的判定结果对同一用户的识别是有意义的。

3.3.5 对比实验分析

本实验选择支持向量机(Support Vector Machine,SVM)、

随机森林(Random Forest,RF)、逻辑回归(Logistic Regression,LR)和IERA-UI^[15]算法作为对比的算法模型。IERA-UI算法通过划分时间片段研究用户兴趣的演变规律和相似程度。

图6表示本文算法与上述4种不同的算法模型的性能参数对比情况。

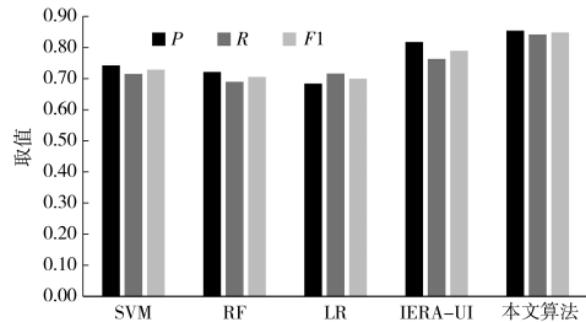


图6 不同算法性能对比

由图6可得,本文算法在不同性能指标上均优于对比算法。主要原因在于:SVM算法和LR算法对缺失数据敏感,RF算法在训练数据存在噪声时容易出现过拟合的现象,而实验中用于分析的用户数据存在部分信息缺失的情况,从而影响到最终算法效果。而IERA-UI算法仅考虑了用户文本内容中体现出的用户兴趣,没有结合其他维度特征。本文算法分别在3种不同的特征维度下分别判定用户归属关系,然后综合判定结果进行同一用户的识别,在一定程度上降低了用户信息缺失带来的影响,且相比其他机器学习算法减少了训练模型的过程,更为简单高效。

4 结论

本文针对跨社交网络的同一用户识别问题,提出了一种综合用户兴趣、写作风格和档案属性的识别算法。利用用户名设计筛选器,降低算法计算开销,并通过改进后的TF-IDF算法设计过滤器滤除噪声文本,提高用户动态兴趣提取准确性,实现了在微博和豆瓣平台之间的同一用户识别。下一步工作将重点着眼于用户发表的文本内容中所涉及的地理位置信息,通过挖掘地理位置的变化规律,进一步提高同一用户识别的准确性和适用性。

参考文献

- [1] 金诚,江婷君,闵勇,等.在线社交网络控制实验的现状与展望[J].浙江大学学报(理学版),2020,47(1):1-11.
- [2] 周小平,梁循,赵吉超,等.面向社会网络融合的关联用户挖掘方法综述[J].软件学报,2017,28(6):1565-1583.
- [3] ZAFARANI R,LIU H.Connecting users across social media sites : a behavioral-modeling approach[C]//ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.New York : ACM Press ,2013 : 41-49.
- [4] ZHANG H,KAN M Y,LIU Y,et al.Online social network

- profile linkage[C]//Asia Information Retrieval Symposium.[S. l.]: Springer International Publishing, 2014: 197-208.
- [5] 王正彬, 沈明玉, 赵皎. 基于节点重构的社交网络的隐私保护技术探讨[J]. 计算机工程与应用, 2017, 53(11): 131-136.
- [6] 仲兆满, 管燕, 胡云, 等. 基于背景和内容的微博用户兴趣挖掘[J]. 软件学报, 2017, 28(2): 278-291.
- [7] 何力, 贾焰, 韩伟红, 等. 基于用户主题模型的微博用户兴趣挖掘(英文)[J]. 中国通信, 2014, 11(8): 131-144.
- [8] 吕志泉, 李昊, 张宗福, 等. 基于主题模型的社交网络匿名用户重识别[J]. 计算机科学, 2019, 46(6): 143-147.
- [9] 金奕江, 孙晓明, 马少平. 因特网上的写作风格鉴别[J]. 广西师范大学学报(自然科学版), 2003, 21(1): 62-66.
- [10] 杨仁凤, 陈端兵, 谢文波. 微博用户兴趣主题抽取方法[J]. 电子科技大学学报, 2018, 47(4): 633-640.
- [11] ZHANG Y, EICK C F. Tracking events in Twitter by combining an LDA-based approach and a density-contour clustering approach[J]. International Journal of Semantic Computing, 2019, 13(1): 87-110.
- [12] 邢千里, 刘列, 刘奕群, 等. 微博中用户标签的研究[J]. 软件学报, 2015, 26(7): 1626-1637.
- [13] 格桑多吉, 乔少杰, 韩楠, 等. 基于 Single-Pass 的网络舆情热点发现算法[J]. 电子科技大学学报, 2015, 44(4): 599-604.
- [14] 陈得福. 基于监督学习的写作风格建模方法研究[D]. 杭州: 浙江大学, 2016.
- [15] 吴铮. 跨社交网络用户多重身份识别算法研究[D]. 郑州: 解放军信息工程大学, 2017.

(收稿日期: 2021-03-16)

作者简介:

沈佳琪(1995-), 男, 硕士, 主要研究方向: 社会关系网络分析。

周国民(1971-), 男, 硕士, 教授, 主要研究方向: 信息网络安全、信息智能处理、大数据公安信息化应用。



扫码下载电子文档

(上接第 108 页)

6 结论

本文设计的可靠性优化方案通过软硬件优化相结合, 解决了 LVDS 传输初始化建链过程中出现的首帧丢失和误建链问题。通过识别帧头、发送训练帧和电气隔离等措施, 能使链路对异常状况的适应性进一步增强。本优化方案中的具体措施对于其他 LVDS 链路具备良好的可移植性和实用性。

参考文献

- [1] 乔锐, 翟成瑞, 张彦军. 一种 LVDS 高速数据传输电路[J]. 电子器件, 2021, 44(2): 312-315.
- [2] 鄢玲玲, 文丰, 李辉景. 一种基于 LVDS 高速传输的接口优化设计[J]. 电测与仪表, 2016, 53(20): 80-83.
- [3] 张晋瑛. 基于 Camera Link 的图像数据采集装置的研究与实现[D]. 太原: 中北大学, 2020.
- [4] 雷武伟, 文丰, 刘东海, 等. 基于 LVDS 的高可靠性远距离数据传输设计[J]. 电子技术应用, 2019, 45(6): 130-134.
- [5] 董静怡, 庞景月, 彭宇, 等. 集成 LSTM 的航天器遥测数据异常检测方法[J]. 仪器仪表学报, 2019, 40(7): 22-29.
- [6] 方炜, 任勇峰, 储成群, 等. 低电压差分信号远距离通信
- 电缆参数估计[J]. 科学技术与工程, 2018, 18(19): 209-212.
- [7] 李辉景, 王淑琴, 任勇峰, 等. 基于 CRC 校验的高速长线 LVDS 传输设计[J]. 电子器件, 2015, 38(6): 1346-1351.
- [8] 张波, 李杰, 张海鹏, 等. 基于 FPGA 的 LVDS 传输链路的可靠性设计[J]. 电子器件, 2018, 41(5): 1237-1241.
- [9] 朱泽琿, 任勇峰, 贾兴中. 基于 LVDS 长距离高可靠性传输的优化设计[J]. 电子测量技术, 2020, 43(20): 150-154.
- [10] 文丰, 张宇, 吴瑞华. 基于 4 bit-CRC 反馈网络的高速 LVDS 传输优化设计[J]. 兵器装备工程学报, 2019, 40(9): 108-113.
- [11] 陈晓敏, 薛志超, 张志龙, 等. 低电压差分信号长线传输的优化设计[J]. 电子技术应用, 2020, 46(11): 104-108.
- [12] 李晋涛, 任勇峰, 杨志文, 等. 基于 SATA3.0 的存储系统优化设计[J]. 电子技术应用, 2021, 47(1): 86-90.

(收稿日期: 2021-05-28)

作者简介:

杜凯旋(1997-), 男, 硕士研究生, 主要研究方向: 测试计量技术及仪器。

焦新泉(1978-), 男, 博士, 副教授, 主要研究方向: 微纳器件及测试技术。



扫码下载电子文档

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所