

# 一种融合节点属性信息的社会网络链接预测方法

张 昱 高克宁 于 戈

(东北大学计算机科学与工程学院 沈阳 110819)

**摘 要** 随着大规模社会网络的发展,链接预测成为了一个重要的研究课题。研究了在社会网络中融合节点属性信息进行链接预测,在传统的社会-属性网络图模型的基础上,将节点属性的类别这一重要参量加入到网络构建中。基于此,提出了一系列为网络中不同类型的连边分配边权重的方法,最后通过随机游走的方法进行网络链接的预测。实验表明,所提链接预测方法相比同类方法有明显的效果提升。

**关键词** 链接预测,社会网络,社会-属性网络,社会节点,属性节点

**中图分类号** TP311 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2018.06.007

## Method of Link Prediction in Social Networks Using Node Attribute Information

ZHANG Yu GAO Ke-ning YU Ge

(School of Computer Science and Engineering, Northeastern University, Shenyang 110819, China)

**Abstract** With the development of large social networks, link prediction has become an important research subject. The link prediction problem in social networks using rich node attribute information was studied in this paper. Based on attribute-augmented social network model, which means rebuilding an augmented network by adding additional nodes with each node corresponding to an attribute, called social-attribute network, the classification of node attributes was added to the model as an important parameter. Several methods of assigning weights for different kinds of links were proposed. Then a random walk method was used for link prediction in the network. Experimental results reveal that this method has better performance compared with other similar methods.

**Keywords** Link prediction, Social network, Social-attribute network, Social node, Attribute node

## 1 引言

近年来,在线社会网络已成为人与人之间相互联系以及信息传播的重要载体,因此社会网络演化的基本机制成为社会网络研究的重点。其中,链接预测研究是一种从微观的角度研究网络演化机制的尝试,且取得了很大的进展。经典的链接预测问题的定义如下<sup>[1]</sup>:给定  $t$  时刻的网络  $G(V, E)$ , 其中  $V$  为节点集合,  $E$  为边集合; 预测  $t'$  ( $t' > t$ ) 时刻的网络  $G(N, E')$ , 即预测加入到网络中的新边。此外,链接预测还有另一种定义,即预测网络中缺失的边:假定网络是不完整的,预测应该出现在网络中的边。

早期的链接预测问题研究集中于利用网络的结构特征来计算网络中两个节点之间的相似性,并根据相似性越大则两者产生链接的可能性就越大的思想进行预测。例如,共同邻居数(Common Neighbors)就以网络中两个节点间的共有邻居数作为两者相似性的衡量指标。除共同邻居数之外,研究者还提出了几十种用于度量节点相似性的指标和方法<sup>[2-3]</sup>,如杰卡德系数(JA)、Adamic-Adar(AA)、Random Walk with Re-

start(RWR)等。这些相似性度量指标有的极其简单,有的极为复杂,计算量各异。但是,并没有一个能够适用于所有网络的指标,有些极其复杂的指标在某些网络上的预测效果甚至不及共同邻居数的预测效果。另一方面,随着在线社会信息网络的兴起,越来越多的研究集中于如何针对大规模在线社会信息网络进行链接预测,这一研究最典型的应用是通过为网络用户推荐好友来提高其对网络的忠诚度。大型社会网络数据的出现带来了新的问题,即大多早期的研究因受限于数据集只使用了网络图结构信息进行预测,而现在的社会信息网络中则包含了大量的用户属性信息,并且这些属性信息对链接的形成可能产生了很大的作用。因此,在网络的结构信息基础上加入用户的属性信息进行链接预测的研究,就显得尤为必要。

加入用户属性信息对社会网络进行链接预测时需要解决一个重要的问题,即如何度量用户属性对建立链接的贡献。对于两个用户而言,他们既可能因为有共同的朋友(结构上有联系)而产生链接,也可能因为有共同的兴趣爱好(属性上相似)而产生联系。结构信息和属性信息如何作用在一起从而

到稿日期:2017-03-11 返修日期:2017-06-18 本文受教育部基本科研业务费项目青年教师科研启动基金(N151603001),辽宁省科技攻关项目博士启动基金(201601026)资助。

张 昱(1980—),男,博士生,讲师,CCF会员,主要研究方向为社会网络,E-mail:zhangyu@mail.neu.edu.cn;高克宁 教授,主要研究方向为Web信息处理、社会网络;于 戈 教授,博士生导师,主要研究方向为数据库理论与技术等,E-mail:yuge@mail.neu.edu.cn(通信作者)。

建立链接预测的模型,是一个尚未解决的问题。

本文提出了一个将社会信息的网络结构信息和用户属性信息相结合的模型,并通过一系列方法计算网络中边的权重,最终采用随机游走的方法进行链接预测,取得了很好的预测结果。

## 2 融合节点属性信息的链接预测

### 2.1 网络的构建

链接预测可被看作一个二元分类问题,即两个节点间有链接或无链接。因此,采用有监督机器学习的分类方法均可被应用到链接预测中,如神经网络、支持向量机等。这些机器学习类的方法都需要提供训练样本的特征集,采用此类方法的研究者通常都是将网络结构信息(如两个节点间的共同好友数、路径长度等)作为训练集的特征来参与学习<sup>[4]</sup>;也有研究者将网络中节点或边的属性(如两个人在网络中发表的话题匹配数量)加入到训练特征中<sup>[5]</sup>,并取得了一定的效果。通常情况下,加入用户属性信息特征都能提高预测的准确度。

有些研究者将节点属性转换成网络边的边权重<sup>[6-7]</sup>,然后利用有权图的节点相似性度量方法进行链接预测或推荐。这种方法的缺点是:如果网络中的两个节点在图结构上距离较远或是没有连通路经,那么即使它们的属性非常相似,也很难有机会产生链接。

还有的研究者将用户属性也作为网络中的节点,从而形成一个带有属性节点的增量社会网络图<sup>[8-10]</sup>,如图1所示。网络中的节点分为两类:1)表示人的社会节点(Alice等);2)表示各种属性信息的属性节点(football, computer science等)。通常情况下,属性信息表示一种分类信息,例如一个人的专业、兴趣爱好等。虽然在实际网络中除了这种分类属性信息外还有许多连续或离散的数值属性,但在实际操作中仍可以通过对数值进行分类或聚类使其成为若干个属性节点,然后应用到网络中。

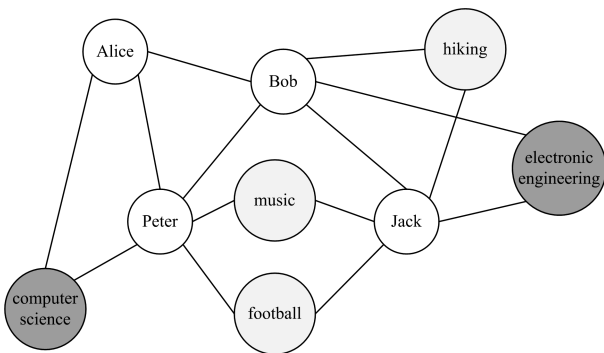


图1 社会-属性网络示意图

Fig. 1 Illustration of social-attribute network

有研究者将这种网络命名为“社会-属性网络”(Social-Attribute Network, SAN),并在此基础上进行链接预测。以这种方式构建网络的优点是:将用户的属性信息和网络结构信息组合在同一个层面共同用于预测。相比有的研究者直接将属性信息换算成边权重应用于社会网络中<sup>[7]</sup>,SAN方法将属性信息提升到与网络结构信息同等重要的高度,这也更加符合社会网络中链接生成的原因:1)社会影响,表现为两者在

网络上相近;2)同质性,表现为两者具有共同的属性。采用SAN方法,可以保证两个节点即使在网络结构上的距离很远,也有可能因为具有相同的属性而变为邻居。

文献<sup>[8-10]</sup>在处理SAN网络时将不同类别的属性节点做统一处理,而实际上不同类别的属性对链接形成的贡献不一。例如图1中,football等属性节点属于爱好类别,而computer science则属于专业类别,它们对网络中特定节点之间是否形成链接的作用不同。因此,本文在重新定义SAN网络时加入了属性类别这一重要因素。

SAN网络的传统定义为有向图 $G'(V', E')$ ,其中 $V' = V_p \cup V_a$ , $V_p$ 表示社会节点集合, $V_a$ 表示属性节点集合。为所有社会节点 $p \in V_p$ 与其所具有的属性 $a \in V_a$ 建立双向链接,并与原始的社会网络图中社会节点间的双向链接一起组成 $E'$ 。在此基础上,为SAN网络加入节点属性类别这一参量,即为所有的节点 $v \in V'$ 标记其属性的类别 $c, c \in \{p(\text{记为 } c_0), c_1, c_2, \dots, c_n\}$ ,其中 $p$ 表示社会节点的类别(人), $c_1, c_2, \dots, c_n$ 表示 $n$ 种不同的属性类别(爱好、专业等)。

### 2.2 边权的分配

已有的SAN研究有的简单地将网络设定为无向无权图<sup>[10]</sup>;有的虽然将网络设定为有向有权图,但对网络中边权重的分配较为单一<sup>[8]</sup>。举例而言,社会节点与属性节点之间的边权在已有方法中是平均分配的,而实际上每个人对某一种属性的链接强度是不同的。本文将SAN定义为有向有权图,并重新定义各类节点之间的边权分配方式。

SAN网络中包含了社会节点和属性节点。构建后的网络通过随机游走的方法进行链接的预测,因此将SAN定义为有向图。网络中的边可以分为如下几类:

- 1) 社会节点 $\rightarrow$ 社会节点(简记为 $p \rightarrow p'$ );
- 2) 社会节点 $\rightarrow$ 属性节点( $p \rightarrow a$ );
- 3) 属性节点 $\rightarrow$ 社会节点( $a \rightarrow p$ )。

社会节点之间的连边表示网络中的好友关系,社会节点与属性节点的连边表示人具有该种属性。因为将SAN定义为有向图,所以在为社会节点与属性节点连边分配权重时,需要从两个方向分别进行处理。

- 1) 属性节点 $\rightarrow$ 社会节点

链接预测最终预测的是社会节点之间是否产生联系。在SAN中,属性节点到社会节点的连边更多地是起到一种桥梁的作用,为人与人之间建立链接提供了可能。

根据属性节点的独立性,一个属性节点只可能与社会节点产生连边,因此可以简单地定义属性节点到社会节点的边权为<sup>[9]</sup>:

$$w(a \rightarrow p) = \frac{1}{|N_p(a)|} \quad (1)$$

其中, $N_p(a)$ 为与属性节点 $a$ 链接的所有社会节点集合。这样的边权分配表明属性节点到与它有链接的社会节点的强度是等同的。如图1所示,“football”这个属性节点对于“Peter”和“Jack”这两个邻居节点的权重一样,都为 $\frac{1}{2}$ 。

实际上,同一个属性对于不同的人应该有不同的权重。如果某个社会节点的某一类属性下的属性节点过多,则应适

当削弱每一个属性节点对其的权重;如果社会节点所链接的某类属性节点过少,那么仅有的若干个属性节点对其产生的影响更大。这种思想非常类似于著名的相似性度量指标 Adamic-Adar,该指标的思想是度小的共同邻居节点的贡献应大于度大的共同邻居节点的贡献,且该指标被认为是相似性度量指标中对链接预测效果最好的指标之一。本文依据这样的思想,在对属性分类的基础上重新定义了属性节点到社会节点的边权:

$$\omega(a \rightarrow p) = \frac{1}{|N_{a'}(p)|} \quad (2)$$

$$\sum_{p' \in N_p(a)} \frac{1}{|N_{a'}(p')|}$$

其中,  $N_{a'}(p)$  表示  $p$  的所有与  $a$  同类的属性邻居节点。举例说明,如图 1 所示,对于“football”这个属性节点,“Peter”和“Jack”这两个社会节点均与其相连,“Jack”除了“football”之外,还具有“hiking”和“music”两个与“football”同属爱好类别的属性节点,而“Peter”则只有一个“music”同类属性节点。“Peter”的“爱好”类属性比“Jack”少,因此边权更应该向“Peter”倾斜,即给予其更大的权重。对于“Peter”和“Jack”的其他类别属性,则不予考虑,因为不同类别的属性是无法进行比较的。在此例中:

$$\omega(\text{football} \rightarrow \text{peter}) = \frac{1}{\frac{1}{2} + \frac{1}{3}}$$

## 2) 社会节点→属性节点和社会节点→社会节点

社会节点到属性节点的边权分配不能简单地等同于属性节点到社会节点的边权,因为社会节点还需要与其他社会节点相连。最简单的分配社会节点到其他邻居节点边权的方式是平均分配:

$$\omega(p \rightarrow a) = \frac{1}{|N_a(p)| + |N_p(p)|} \quad (3)$$

其中,  $N_a(p)$  和  $N_p(p)$  分别表示社会节点  $p$  的属性邻居节点集合和社会邻居节点集合。

但实际上,考虑到社会节点可能链接到不同类的属性,边权值不应该简单地平均分配,而是不同的属性类别具有不同的权重。文献[9]将属性节点和社会节点作为两个大的类别进行区分,即为这两类节点分配不同的比重,但其并没有考虑到属性节点实际上还可以细分成多种不同的类别。因此,考虑到不同类别的属性节点,并把社会节点邻居也当作一类,那么社会节点到其他节点的边权分配即可统一在一个框架下,如下式所示:

$$\omega(p \rightarrow a) = \frac{\lambda_{c(a)}}{|N_{a'}(p)|} \quad (4)$$

其中,  $N_{a'}(p)$  表示节点  $p$  所链接的属性节点中与  $a$  属性类别相同的节点集合,  $\lambda_{c(a)}$  表示属性  $a$  节点所属这一属性类别在全部属性类别中的权重比例。显然,对于所有的  $c \in \{c_0, c_1, \dots, c_n\}$ , 要求  $\sum \lambda_c = 1$ 。式(4)实际上就是将总体平均分配权重的方法改成只在同一类别下平均分配,并且按这一类别的比例进行加权。

在不同的社会网络中,不同类别的属性应具有不同的权重比例。例如,在作者合作网络中,专业属性和地址属性的权

重显然应更大。当然,对  $\lambda_c$  不能简单地按照经验分配,本文采用的方法是对每一类属性进行频率统计。举例说明,如果一条边的两个社会节点都链接到了一个属性节点,即由两个社会节点和一个属性节点组成了一个三角形,则可以直观地认为该属性对于两个社会节点的链接是有帮助的。一种属性类别所拥有的此类三角形越多,则其权重比例应该越大。因此,可计算出每种属性类别的三角形数目,并以此来分配权重比例。

接着唯一需要确定的是  $\lambda_c (c = p)$ , 即社会节点到社会节点的权重比例。这个参数决定了网络结构信息对于预测链接的重要程度,其取值越大,网络结构信息在预测中所占的比重就越大;反之,属性信息在预测中的重要程度就越大。一般地,虽然网络结构信息和属性信息对于链接形成的帮助的大小并不明确,但通常网络结构信息对于链接预测更重要,因此  $\lambda_c (c = p) \geq 0.5$  是比较合理的选择。此参数的最优值会根据网络数据集的不同而不同,因此本文在实验中对该参数分别设置了多种取值进行计算,以获得最为理想的结果。

## 2.3 随机游走

在网络中所有边的边权都确定之后,可通过多种基于有向图的随机游走方法进行网络中两个节点的相似性度量。本文选择带有重启的随机游走方法<sup>[11]</sup>(Random Walk with Restart, RWR),该指标可以看作是网页排序算法 PageRank 的拓展应用。该方法假设随机游走粒子  $V_x$  每走一步都会以概率  $1-c$  返回初始位置,如式(5)所示。

$$\pi_x(t+1) = c \cdot P^T \pi_x(t) + (1-c)e_x \quad (5)$$

其中,  $\pi_x$  是一个一维列向量,表示节点  $x$  游走到其他节点的概率;  $P$  为概率转移矩阵,在本文的方法中即为 2.2 节中介绍的利用式(2)和式(4)计算的边权归一化后所形成的矩阵;  $1-c$  为粒子返回源点的概率,此处  $c$  采用 PageRank 算法的标准取值 0.85;  $e_x$  表示初始态,是一个只有第  $x$  个元素为 1 而其他元素都为 0 的一维列向量。通过多次迭代,可计算游走粒子从源节点  $x$  出发走到目标节点  $y$  的稳态概率值,并将其作为  $x$  到  $y$  是否形成链接的度量值。

RWR 方法在社会-属性网络中的应用与在传统社会网络中的应用并无本质区别。游走粒子从一个社会节点  $V_x$  出发,以 2.2 节确定的边权值为转移概率,既可能走向它的社会邻居节点,也可能走向它的属性邻居节点,而通过属性邻居节点再继续游走,就可能与其他社会节点产生联系。通过式(5)反复迭代得到的  $\pi_x$  稳态值,就是节点  $V_x$  游走到其他节点的概率,这个概率可以作为两个节点间是否产生链接的度量值。

## 3 实验

本文实验采用的数据集是 2016 年 10 月从链接为 <http://blog.sciencenet.cn> 的网站抓取的数据,简称为 SciNet-Blog。该网站是一个为科研人员提供交流平台的博客网站,抓取的数据集不仅包含用户之间的好友关系,还包含用户的一些属性信息,如研究领域、工作单位、发表博文类别等。这几类属性信息均具有分类明确的特点,并且一个用户可能具有该种属性类别下的多种属性。举例来说,一个用户可能具有多个不同的研究领域,可能会在不同的博客版块上发表博

文。这样的属性数据对于验证本文提出的方法是十分适合且必要的。为验证算法,过滤掉完全不包含任何属性信息的节点。表 1 列出了该数据集的主要网络特征信息。

表 1 SciNetBlog 的网络特征  
Table 1 Network characteristics of SciNetBlog

SciNetBlog 网络特征	参数值
社会节点数	5041
好友连边数	39507
平均度	15.49
平均聚集系数	0.182
网络直径	9

实验中首先建立社会-属性网络,将属性分为三大类:研究领域( $c_1$ )、工作单位( $c_2$ )、发表博文的版块类别( $c_3$ )。经整理后,计算出每种属性节点的数量和它们与社会节点链接所构成的三角形数量,并得出该属性类别对应的权重在这 3 种属性类别中的比率(这个比率是不包含社会节点类别的),如表 2 所列。

表 2 各类别属性节点的特征  
Table 2 Characteristics of different categories of attribute nodes

属性节点类别	数量	三角形数量	$\lambda_c$ 比率/%
$c_1$	308	31520	90.4
$c_2$	1147	2435	7
$c_3$	92	893	2.6

其中,属性  $c_3$  即发表博文版块这一类别的权重较小,因为大多数用户并不发表博客,所以数据较少;而属性  $c_1$  即研究领域这一属性类别的权重非常大,说明该属性对链接的形成有较大的贡献。

为验证本文方法,选取了链接预测中应用广泛且效果较好的 3 种节点相似性指标进行横向比较。

#### 1) CN(Common Neighbors)

此方法为链接预测的基准方法,将网络中两个节点的共同好友数作为它们的相似度度量指标。

#### 2) AA(Adamic-Adar)

$$AA(x, y) = \sum_{z \in N(x) \cap N(y)} \frac{1}{(\log |N(z)|)} \quad (6)$$

AA 指标考虑了两节点的共同好友的度信息,相当于为每一个共同好友的贡献赋予一个权重,如果某个共同好友的好友越少,那么他的贡献就越大。本文在式(2)中提出的为属性节点到社会节点的连边分配权重的基本思想与这个指标是一致的。

#### 3) RWR(Random Walk with Restart)

这个指标即为本文所采用的带重启的随机游走方法。由于本文方法是在社会-属性网络即原有社会网络的一个增量网络上进行的,因此这 3 种可进行比较的指标还可被进一步划分为在原有社会网络中(不含节点属性信息)和社会-属性网络中分别计算。此外,为体现本文提出的边权重设置方法的有效性,在计算 RWR 指标时采用按式(1)和式(3)平均分配边权的方法(SAN-RWR1)、按式(2)和式(4)分配边权的方法(SAN-RWR2)分别进行,以进行结果比较。

将该数据集的用户节点连边分别去除 10%、20% 和 30% 用作测试集,剩余的连边用作训练集。测试时,分别从测试集随机选取有连边的节点对和无连边的节点对进行节点间相似

性度量值的比较。令  $n$  为总的比较次数, $n'$  为有连边的节点对度量值大于无连边的节点对度量值的次数, $n''$  为有连边的节点对度量值与无连边的节点对度量值相等的次数,则链接预测的 AUC 评估值如下:

$$AUC = \frac{n' + 0.5n''}{n} \quad (7)$$

实验中,还需要确定  $\lambda_c (c=p)$  即社会节点邻居这一类别的权重比例。实验分别验证了从 0.5 到 1(步长为 0.1)的 6 种取值,取值 0.5 表示属性节点和社会节点具有同等的重要性,取值 1 则表示完全不考虑属性节点的作用,即等同于原始的不包含属性节点的社会网络图。经验证, $\lambda_c (c=p)$  取值 0.8,且表 2 中的 3 种属性类别比率分别乘以 0.2 得到各自的  $\lambda_c$ ,最终得到的结果最优。

图 2 给出了 7 种指标在 3 种测试环境下的链接预测 AUC 评估值的比较结果。可以看出,在 3 种不同的训练集比例测试环境中,SAN 下的 AUC 值均明显优于传统 SN 下的结果,而本文采用的基于属性类别的边权重分配方法 SAN-RWR2 的结果也高于平均分配边权模式的 SAN-RWR1 的结果。值得注意的是,在训练集较为完整时,本文方法的优势并不明显;而当网络的缺失边过多时,传统方法的预测准确率快速下降,而加入属性节点的方法则依然保持了较好的预测效果,这说明了当网络结构信息缺失较多时,节点的属性信息对于重建完整网络结构的作用很大。实验结果表明,本文的方法对提高社会信息网络的链接预测准确度是有效的。

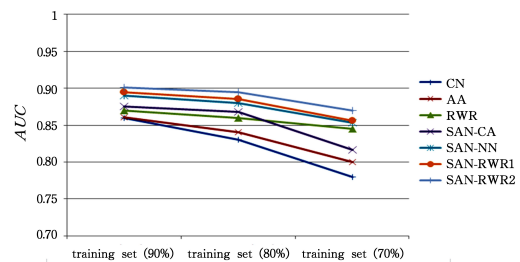


图 2 实验比较结果

Fig. 2 Performance comparison of link prediction

**结束语** 本文设计了一种根据属性分类构建社会-属性网络的方式,并在此基础上提出了一系列分配网络边权的方法。经实验验证,所提方法可有效改善社会网络链接预测的效果。此方法也可进一步推广至其他采用有向有权图构建网络的链接预测方法中。因为在不同的网络中,网络结构信息和节点属性信息对链接形成的贡献不同,所以下一步的工作是研究如何通过现有数据确定这两者的权重比例,即确定  $\lambda_c (c=p)$  的最优值。

## 参考文献

- [1] LIBEN-NOWELL D, KLEINBERG J. The link-prediction problem for social network[J]. Journal of the American Society for Information Science and Technology, 2007, 58(7): 1019-1031.
- [2] LÜ L, ZHOU T. Link prediction in complex networks: A survey [J]. Physica A Statistical Mechanics & Its Applications, 2011, 390(16): 1150-1170.
- [3] HASAN M A, ZAKI M J. A Survey of Link Prediction in Social

- Networks[M]// Social Networks Data Analytics. Springer US, 2011:243-275.
- [4] LICHTENWALTER R N, LUSSIER J T, CHAWLA N V. New perspectives and methods in link prediction[C]// ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Washington DC, USA, 2010:243-252.
- [5] HASAN, MOHAMMAD A. Link Prediction using Supervised Learning[J]. Proc of Sdm Workshop on Link Analysis Counterterrorism & Security, 2006, 30(9):798-805.
- [6] HSU C C, LAI Y A, CHEN W H, et al. Unsupervised Ranking using Graph Structures and Node Attributes[C]// Tenth ACM International Conference on Web Search and Data Mining. ACM, 2017:771-779.
- [7] BACKSTROM L, LESKOVEC J. Supervised random walks: predicting and recommending links in social networks[C]// ACM International Conference on Web Search & Data Mining. ACM, 2011:635-644.
- [8] YIN Z, GUPTA M, WENINGER T, et al. LINKREC: a unified framework for link recommendation with user attributes and graph structure[C]// International Conference on World Wide Web(WWW 2010). Raleigh, North Carolina, USA, 2010:1211-1212.
- [9] YIN Z, GUPTA M, WENINGER T, et al. A Unified Framework for Link Recommendation Using Random Walks[C]// Proceedings of the 2010 International Conference on Advances in Social Networks Analysis and Mining. IEEE Computer Society, 2010:152-159.
- [10] GONG N Z, TALWALKAR A, MACKEY L, et al. Joint Link Prediction and Attribute Inference Using a Social-Attribute Network[J]. Acm Transactions on Intelligent Systems & Technology, 2014, 5(2):1-20.
- [11] TONG H, FALOUTSOS C, PAN J Y. Fast Random Walk with Restart and Its Applications[C]// International Conference on Data Mining. IEEE, 2006:613-622.
- 
- (上接第 31 页)
- [42] MESSMER B T. Efficient graph matching algorithms for pre-processed model graphs [D]. Switzerland: University of Bern, 1996.
- [43] MESSMER B T, BUNKE H. A new algorithm for error-tolerant subgraph isomorphism detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(5):493-504.
- [44] MESSMER B T, BUNKE H. A decision tree approach to graph and subgraph isomorphism detection[J]. Pattern Recognition, 1999, 32(12):1979-1998.
- [45] MESSMER B T, BUNKE H. Efficient subgraph isomorphism detection: A decomposition approach[J]. IEEE Transactions on Knowledge and Data Engineering, 2000, 12(2):307-323.
- [46] CORDELLA L P, FOGGIA P, SANSONE C, et al. A (sub) graph isomorphism algorithm for matching large graphs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(10):1367-1372.
- [47] MCKAY B D. Practical graph isomorphism[J]. Journal of Symbolic Computation, 2014, 60(1):94-112.
- [48] MCKAY B D, PIPERNO A. Practical graph isomorphism, II[J]. Journal of Symbolic Computation, 2014, 60(1):94-112.
- [49] MESSMER B T, BUNKE H. Error-correcting graph isomorphism using decision trees[J]. International Journal of Pattern Recognition and Artificial Intelligence, 1998, 12(6):721-742.
- [50] DIAS J R, MILNE G W A. Chemical applications of graph theory[J]. Journal of Chemical Information & Modeling, 1976, 32(1):210-242.
- [51] POOLE J. Similarity in legal case-based reasoning as degree of matching between conceptual graphs: Work in progress[C]// Proceedings First European Workshop on Case-Based Reasoning. 1993.
- [52] BÖRNER K, PIPPIG E, TAMMER E C, et al. Structural similarity and adaptation[C]// European Workshop on Advances in Case-Based Reasoning. Springer Berlin Heidelberg, 1996.
- [53] FISHER D H. Knowledge acquisition via incremental conceptual clustering[J]. Machine Learning, 1987, 2(2):139-172.
- [54] LEAKE D B, PLAZA E, INTELLIGEN Z. Case-Based Reasoning Research and Development[C]// Second International Conference on Case-Based Reasoning. 1997.
- [55] EHRIG H, HABEL A, KREOWSKI H J. Introduction to graph grammars with applications to semantic networks[J]. Computers & Mathematics with Applications, 1992, 23(6-9):557-572.
- [56] MAHER P E. A similarity measure for conceptual graphs[J]. International Journal of Intelligent Systems, 1993, 8(8):819-837.
- [57] SHOWBRIDGE P, KRAETZL M, RAY D. Detection of abnormal change in dynamic networks[C]// Information, Decision and Control, 1999(IDC 99). IEEE, 1999:557-562.
- [58] WANG Y K, FAN K C, HORNG J T. Genetic-based search for error-correcting graph isomorphism[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 1997, 27(4):588-597.
- [59] YAMAZAKI K, BODLAENDER H L, FLUITER B D, et al. Isomorphism for graphs of bounded distance width[C]// Italian Conference on Algorithms and Complexity. 1997:276-287.
- [60] BALKO M, CIBULKA J, KRÁI K, et al. Ramsey numbers of ordered graphs [J]. Electronic Notes in Discrete Mathematics, 2015, 49:419-424.
- [61] JIANG X Y, BUNKE H. Optimal quadratic-time isomorphism of ordered graphs [J]. Pattern Recognition, 1999, 32(7):1273-1283.
- [62] SHEARER K R. Indexing and retrieval of video using spatial reasoning techniques [D]. Curtin University of Technology, 1998.
- [63] SHEARER K, BUNKE H, VENKATESH S. Video indexing and similarity retrieval by largest common subgraph detection using decision trees[J]. Pattern Recognition, 2001, 34(5):1075-1091.
- [64] FOGGIA P, SANSONE C, VENTO M. A performance comparison of five algorithms for graph isomorphism[C]// Proceedings of the 3rd IAPR TC-15 Workshop on Graph-based Representations in Pattern Recognition. 2001:188-199.