

基于空间映射的顶点带属性网络的链接预测

姜卯生¹ 葛剑飞¹ 陈 峻^{1,2}

(扬州大学信息工程学院 扬州 225127)¹ (南京大学软件新技术国家重点实验室 南京 210093)²

摘 要 提出了基于空间映射的顶点带属性网络的链接预测算法。顶点带属性网络包含拓扑及顶点属性两种信息,为了综合考虑这两种信息,将二者同时映射到另一空间。完成空间映射后,在新的空间计算顶点的相似度,并以此来预测链接存在的可能性。提出分步交叉迭代的方法来取得最优的映射矩阵,以在新的空间中有效融合拓扑信息与顶点属性信息。实验结果证明了空间映射方法的正确性,所提出的基于空间映射的方法能够取得较高质量的预测结果。

关键词 空间映射,链接预测,复杂网络,顶点属性

中图法分类号 TP181 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.07.045

Link Prediction in Networks with Node Attributes Based on Space Mapping

JIANG Mao-sheng¹ GE Jian-fei¹ CHEN Ling^{1,2}

(College of Information Engineering, Yangzhou University, Yangzhou 225127, China)¹

(State Key Lab of Novel Software Tech, Nanjing University, Nanjing 210093, China)²

Abstract A link prediction algorithm in networks with node attributes based on space mapping was proposed in this paper. In general, networks with node attributes have two types of information, i. e. topology and node attributes. To integrate these two types of information, we mapped them into a new space. After information being mapped, similarities between nodes were calculated in the new space, which are used to predict the possibility of links between nodes. Alternative iteration method was proposed to get the optimal mapping matrix, so as to integrate topology information and node attribute information effectively. The experimental results verify the correctness of space mapping approach, and show that the proposed space mapping based algorithm can obtain high quality prediction results.

Keywords Space mapping, Link prediction, Complex networks, Node attributes

1 引言

现实世界中,许多系统都可以用复杂网络来描述和刻画,如交通网^[1]、神经网络^[2]、社交网^[3-5]等。对复杂网络的拓扑性质进行深入分析和挖掘,是解决社会及科技发展中许多问题的前提条件。复杂网络中的链接预测是复杂网络分析中的一个重要分支^[6],链接预测的目的是发现网络节点间的潜在联系。链接预测可被应用到诸多领域,如社会网络、生物学网络、互联网、学术网络和通讯网络等。例如,在生物学领域,链接预测被应用到蛋白质相互作用网络和新陈代谢网络中:酵母菌的蛋白质之间存在着相互作用,但我们对其中的80%一无所知^[7],链接预测能够揭示这类网络中未知的相互作用关系,从而推动人类对生物的进一步认知。链接预测在社会网络分析中可以作为准确分析社会网络结构的有力的辅助工具,在社会关系分析中可用于发现人际之间的潜在联系^[8]。链接预测的思想和方法还可以用于学术网络以预测一篇学术

论文的影响力以及作者未来的研究课题^[9-10]。链接预测的方法也可直接用于信息推荐,如在电子商务中可以用于向客户推荐商品^[11-13];在社交媒体网络中,链接预测可以检测信息传播的途径^[14];在蛋白质相互作用网络中,链接预测可以用来预测蛋白质之间动态的相互作用^[7]。

早期, Sarukkai^[15]应用马尔可夫链进行了链接预测及路径分析的工作。随后,马尔可夫链的预测方法被 Zhu 等人^[16]应用到了 WWW 网络的链接预测中。近年来,越来越多的科研工作者开始关注网络结构特征的研究^[6],他们利用网络的结构信息进行链接预测研究。这些基于结构信息的链接预测算法通常具有较高的预测精度和适用性,而且这些信息通常会被转化成关于结构的相似性。对网络链接的预测包括基于节点相似性、基于最大似然估计、基于概率模型等方法。基于节点相似性方法假设两个节点之间相似性越大则它们之间存在链接的可能性就越大^[17-18]。为此,有很多有关节点相似性的定义,如基于局部信息的相似性指标^[6],主要包括共同邻居

到稿日期:2016-04-11 返修日期:2016-08-31 本文受国家自然科学基金(61379066,61379064,61472344,61402395),江苏省自然科学基金(BK20130452,BK2012672,BK2012128,BK20140492),江苏省教育厅自然科学基金(12KJB520019,13KJB520026),江苏省六大人才高峰项目(2011-DZXX-032)资助。

姜卯生(1974—),男,博士生,讲师,主要研究方向为机器学习、数据挖掘,E-mail:msjiang@yzu.edu.cn;葛剑飞(1989—),男,硕士,主要研究方向为机器学习、数据挖掘;陈 峻(1951—),男,教授,博士生导师,主要研究方向为并行计算、人工智能、机器学习、数据挖掘。

(CN)指标、Salton 指标、Jaccard 指标、Sorenson 指标、大度节点有利指标 HPI 和 LHN_I、大度节点不利指标(HDI)、优先链接指标(PA)等;基于路径的相似性指标,主要包括局部路径指标(LP)、Katz 指标等。

为了提高链接预测的准确率,一些更为复杂的技术被用来计算相似性,例如将局部指标进行加权的方法^[19]、基于蚁群优化的方法^[20]、基于抽样的方法^[21]、基于网络社区划分的方法^[5,22]等。基于网络社区划分的方法根据网络的社区结构来预测网络顶点之间存在链接的可能性,在具有明显层次结构的网络中具有很好的效果。

在许多实际网络中,顶点所代表的对象具有一定的属性,而具有相似属性的对象之间往往具有比较密切联系。现在越来越多的研究工作开始关注顶点带属性网络的分析问题^[19,23-24]。本文研究此类顶点带属性网络上的链接预测问题,提出了将网络顶点的属性信息和拓扑信息有机结合的空间映射方法,以提高链接预测的准确性。现在已经有一些对顶点带属性的网络进行链接预测的方法,如矩阵比对方^[25]、增广图方法^[26]、基于概率图的方法^[27]、基于图立方体的方法^[28]、基于随机游走的方法^[29-30]等。这些方法基本出于两种思路:1)将其中一种信息转化为另一种,即将顶点的各类链接的信息作为顶点的一个属性来处理,或者将顶点上属性的相似度过转化为连边上的权重来处理。这样的方法不能反映拓扑信息和属性信息的演化关系,在两类信息不相关时很难得出精确的预测结果。因此,直接将一种信息转换成另外一种会降低链接预测的质量。2)将两种信息分开考虑,然后再简单地综合起来。这样的做法忽视了各类链接信息和属性信息之间可能存在的内在一致性或互斥性,不能实现它们之间的无缝结合。基于以上分析,本文提出将链接信息和属性信息相结合的空间映射方法。这种结合由于不是简单地将两种信息合并相加,而是共同向同一个空间进行映射,然后在映射后的空间进行链接预测,因此有效提高了链接预测的准确性。

2 空间映射的基本思想

一个网络可表示为图 $G=(V,E)$,其中 V 是所有顶点的集合, E 是所有边的集合。设 $|V|=n$,则 $n \times n$ 阶矩阵 $A=[a_{ij}]_{n \times n}$ 为这个网络的邻接矩阵, $a_{ij}=1$ 表示顶点 v_i 和 v_j 之间存在链接, $a_{ij}=0$ 表示不存在链接。设顶点上有 m 个属性, $n \times m$ 阶矩阵 $B=[b_{ij}]_{n \times m}$ 表示该网络的属性信息, $b_{ij}=1$ 表示第 i 个顶点具有第 j 个属性, $b_{ij}=0$ 表示第 i 个顶点不具有第 j 个属性。

邻接矩阵 A 和属性矩阵 B 都直接反映了网络的拓扑信息和属性信息,在邻接矩阵中行元素和列元素都分别表示网络中对应的顶点。矩阵中的元素则表示顶点与顶点在结构上的关系,即是否存在链接。而属性矩阵的行元素表示网络中对应的顶点,列元素表示每一个属性,矩阵中的每一个元素表示该行的顶点是否具有该列的属性。由于这两种信息在不同的线性空间内,因此若直接对这两个矩阵进行运算则无法达到综合拓扑信息和属性信息的目的。

为了使拓扑信息和属性信息能有机融合,而不以其中的某一种信息为主,可采用空间映射的方法将它们同时映射到

同一空间中。通过寻找最优的映射矩阵,将分别表示拓扑信息和属性信息的邻接矩阵和属性矩阵转换到同一空间。

属性矩阵中具体的某一行表示某一顶点是否拥有对应的每一个属性。实际上,这一行可以看作是该顶点的一组特征。考虑采用类似的形式对拓扑信息进行表示,即用一组特征来表示每一个顶点的拓扑信息,以平衡拓扑信息和属性信息。设映射的目标空间为 $r(r \leq \min\{m,n\})$ 维,将邻接矩阵 A 到新的空间的映射用一个矩阵 $U=[u_{ij}]_{n \times r}$ 来表示。通过映射得到新矩阵 AU ,使得原来的拓扑信息被转化成了新空间里的表示形式。矩阵 AU 中的某一行表示某一顶点的一组拓扑特征。同时,也将属性矩阵 B 映射到 r 维的目标空间,将属性矩阵 B 到新的空间的映射用一个矩阵 $V=[v_{ij}]_{m \times r}$ 来表示,得到的矩阵 BV 即是空间映射后的矩阵。

完成了空间映射以后,可在新的空间内计算融合两方面信息后的相似度,从而进行链接预测。上述空间映射的基本过程如图 1 所示。

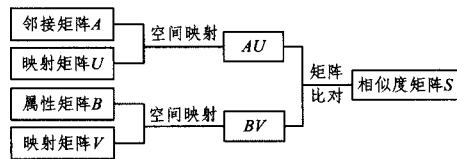


图 1 空间映射的过程

3 映射矩阵的计算

设 $U=[u_{ij}]_{n \times r}$ 是邻接矩阵 A 用以完成空间映射的映射矩阵, $V=[v_{ij}]_{m \times r}$ 是属性矩阵 B 的映射矩阵,其中 U 和 V 为各列正交的矩阵。为了将邻接矩阵 A 和属性矩阵 B 同时映射到一个 $r(r \leq \min\{m,n\})$ 维空间,需要找到最优的映射矩阵,以提高预测精度。

为了有机融合拓扑信息和属性信息,希望拓扑信息和属性信息尽可能一致,即要使空间映射之后邻接矩阵 A 和属性矩阵 B 分别被转化成的矩阵 AU 和 BV 尽可能接近。为此,采用矩阵比对的方法最小化两个矩阵的差异,从而需要最小化以下目标函数:

$$L(U,V) = \|AU - BV\|_F^2 + \lambda_1 \|U\|_F^2 + \lambda_2 \|V\|_F^2 \quad (1)$$

其中,最后两项 $\lambda_1 \|U\|_F^2 + \lambda_2 \|V\|_F^2$ 为正则项,作用是防止过度拟合; λ_1 和 λ_2 为正则项系数,是正的常数; $\|\cdot\|_F$ 表示矩阵的 Frobenius 范数。设 H 是一个 $m \times n$ 的矩阵,则 Frobenius 范数的定义如下:

$$\|H\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |h_{ij}|^2}$$

即为矩阵所有元素平方和的平方根。Frobenius 范数还有如下等价的定义:

$$\|H\|_F = \sqrt{\text{tr}(H^T H)}$$

其中, $\text{tr}(H^T H)$ 为矩阵 $H^T H$ 的迹,为矩阵 $H^T H$ 主对角线元素之和。

为了寻求合适的各列正交的矩阵 U 和 V ,需要解决如下的优化问题:

$$\min_{U,V} L(U,V) = \min_{U,V} [\|AU - BV\|_F^2 + \lambda_1 \|U\|_F^2 + \lambda_2 \|V\|_F^2] \quad (2)$$

能够使得函数 $L(U,V)$ 获得最小值的 U 和 V 便是目标最优映射矩阵。由于存在 U 和 V 两个未知矩阵,因此本文提出一种分步交叉优化的策略,通过迭代得到最优的映射矩阵 U 和 V 。分步交叉优化策略的主要思想是:轮流选择,先固定两个未知矩阵 U 和 V 中的一个,以另一个为变量来优化 $L(U,V)$ 。例如,可以先固定矩阵 V ,以矩阵 U 为变量来优化 $L(U,V)$;然后,再固定矩阵 U ,以矩阵 V 为变量来优化 $L(U,V)$ 。如此交叉迭代进行优化,直到收敛为止。

在固定矩阵 V 而以矩阵 U 为变量来优化 $L(U,V)$ 时,先求 $L(U,V)$ 关于 U 的偏导数:

$$\frac{\partial L}{\partial U}(U,V)=(A^T A+A A^T)U-2 A^T B V+2 \lambda_1 U$$

根据 KKT 条件,需要置偏导数为 0,即:

$$(A^T A+A A^T)U-2 A^T B V+2 \lambda_1 U=0$$

根据式(3)可以得出矩阵 U 中的元素 u_{ij} 的更新规则:

$$u_{ij} \leftarrow u_{ij} \cdot \frac{\left(A^T B V\right)_{ij}}{\left(A^T A U\right)_{ij}+\lambda_1 u_{ij}}$$

同理,固定矩阵 U 而以矩阵 V 为变量来优化 $L(U,V)$ 时,对 V 求 $L(U,V)$ 的偏导:

$$\frac{\partial L}{\partial V}(U,V)=2 B B^T V-2 B^T A U+2 \lambda_2 V$$

根据 KKT 条件,置偏导为 0,即得到:

$$2 B B^T V-2 B^T A U+2 \lambda_2 V=0$$

根据式(5)可以得到矩阵 V 中的元素 v_{ij} 的更新规则:

$$v_{ij} \leftarrow v_{ij} \cdot \frac{\left(B^T A U\right)_{ij}}{\left(B B^T V\right)_{ij}+\lambda_2 v_{ij}}$$

对矩阵 U 和 V 取一定的初值后,通过式(4)和式(6)两组更新规则交替对矩阵 U 和 V 中的所有元素进行更新。迭代过程收敛后,得到的矩阵 U 和 V 就是所求的最优矩阵。

4 算法的基本框架

在求得映射矩阵 U 和 V 之后,邻接矩阵 A 和属性矩阵 B 被映射到另一空间中,转化为矩阵 AU 和 BV ,它们分别表示网络的拓扑信息和属性信息,可在这两个矩阵上计算顶点之间的相似度。由于需要平衡两种信息,因此可以求这两个矩阵对应元素的算术平均值作为新空间的特征矩阵 X 中元素的值,即 $X=(AU+BV)/2$ 。由于 X 中所蕴含的拓扑特征和属性特征的综合信息以每一行的向量形式表示,因此可通过计算矩阵 X 中每两行之间的欧氏距离得到相似度评分矩阵 S 。

综上所述,基于空间映射的顶点带属性网络的链接预测

算法 Spmap(Space mapping)的框架如图 2 所示。

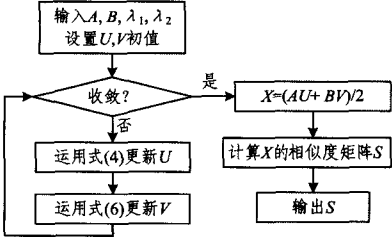


图 2 Spmap 算法的框架

上述算法首先进行交叉迭代,得到最优映射矩阵 U 和 V 。因为 U 为 $n \times r$ 阶矩阵, V 为 $m \times r$ 阶矩阵,每一步迭代,计算矩阵 $A^T AU$ 和 $A^T BV$ 分别需要 $O(n^2 r)$ 和 $O(nmr)$ 的时间。类似地,计算矩阵 $B^T AU$ 和 $BB^T V$ 分别需要 $O(nmr)$ 和 $O(m^2 r)$ 的时间。而对 U, V 矩阵元素的更新过程各需要 $O(nr)$ 和 $O(mr)$ 的时间。然后,算法通过计算欧几里得距离求得相似度评分矩阵 S ,此过程需要 $O(n^3)$ 的时间。

5 实验结果及其分析

5.1 数据集

本节利用顶点带属性的网络来测试所提出的基于空间映射的算法 Spmap,测试的 8 个数据集来自 DBLP(Digital Bibliography Library Project)数据库。该数据库是计算机领域内以作者为核心的一个计算机类英文文献集成数据库系统。测试的 8 个数据集包括 Ad Hoc Networks(AHN), International Journal of Computational Geometry and Applications(IJCGA), Computers and Artificial Intelligence(CAI), Complex, Intelligent and Software Intensive Systems(CISIS), Computing and Informatics(CI), Computer Communication Review(CCR), Applications of Natural Language to Data Bases(ANLDB), Journal of Communications and Networks(JCN)。这些数据集都是 DBLP 数据库中的学术会议或者学术期刊上的文献。

每一个数据集可以被描述为一个网络,其中每一位作者代表一个顶点,作者之间存在的合著关系表示对应的顶点间存在链接。换言之,作者和作者间的关系矩阵即为邻接矩阵 A ,矩阵中的元素为 1 表示两个作者存在合著关系,否则不存在合著关系。提取作者文章标题中的关键词作为该作者对应顶点的属性,作者和关键词矩阵即为属性矩阵。

表 1 列出了描述测试数据集拓扑特征的相关系数。

表 1 网络拓扑特征系数

	<i>nodes</i>	<i>links</i>	<i>attr</i>	<i>NUMC</i>	<i>e</i>	<i>C</i>	<i>r</i>	<i>K</i>
AHN	2262	3581	1688	252/472	0.0032	0.7525	0.212	3.166
IJCGA	940	1699	1290	533/160	0.0748	0.5955	0.147	3.615
CAI	1622	2237	1922	28/522	0.0023	0.639	0.641	2.759
CISIS	2122	4125	2385	287/433	0.0078	0.7811	0.149	3.888
CI	1604	2208	1905	28/517	0.0023	0.639	0.641	2.754
CCR	1562	3960	1727	820/226	0.0018	0.7487	0.3	5.07
ANLDB	847	1211	1041	96/225	0.0072	0.713	0.211	2.86
JCN	2040	3198	1906	61/460	0.0023	0.7468	0.498	3.135

表 1 中, *nodes* 和 *links* 分别表示网络的顶点数和链接数; *attr* 表示顶点的属性数; *NUMC* 表示网络中连通分量(Con-

nected Component)的个数以及其中最大的连通分量(largest Connected Component)所包含的顶点个数,如 JCN 数据集中

对应的 61/460 就表示该网络有 61 个连通分量,其中最大连通分量所包含的顶点个数为 460;另外, e, C, r, K 分别表示网络的网络效率(efficiency of networks)、聚类系数(clustering Coefficient)、同配系数 assortative coefficient)和平均度数(average degree)。

5.2 实验结果

分别将所提出的基于空间映射的顶点带属性网络的链接预测算法 Spmap 和基于 CN, Salton, Jaccard, Sorenson, HPI, LHN_I, HDI, PA, LP, Katz 等指标的链接预测算法在上述 8 个数据集上进行链接预测,以比较它们的性能,即比较它们结果的精确度和它们在各个数据集上所需要的时间。采用 AUC 评价指标来评估各种算法在这 8 个数据集上的预测结果的精确度。

AUC(area under the receiver operating characteristic curve)是从整体上衡量算法精确度的指标。AUC 是测试集

中已存在边的链接预测得分值比随机选取的一个不存在边的得分值高的概率。也就是说,将每次随机地从测试集中任意选取一条已存在的边和从不存在的边集合中随机选取的边进行比较,如果测试集中已存在边的得分大于不存在的边的得分,则加 1;如果两个得分值相等,则加 0.5。独立地比较 n 次,如果测试集中已存在边的得分大于不存在的边的得分的情况有 n' 次,并且二者值相等的情况有 n'' 次,则 AUC 定义为:

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

显然,如果所有分数都是随机产生的,那么 $AUC = 0.5$ 。因此,AUC 大于 0.5 的程度衡量了算法比随机选择的方法更精确的程度。

表 2 列出了 11 种算法利用 AUC 评价方法得到的预测准确率。

表 2 AUC 评价结果

	AHN	IJCGA	CAI	CISIS	CI	CCR	ANLDB	JCN
CN	0.8991	0.8878	0.9239	0.9319	0.9206	0.8976	0.9054	0.9287
Salton	0.9051	0.8933	0.9246	0.9421	0.9207	0.9054	0.9122	0.9321
Jaccard	0.4135	0.4657	0.4249	0.4557	0.3746	0.4769	0.5121	0.5109
Sorenson	0.9053	0.8912	0.9241	0.9434	0.9207	0.9051	0.9125	0.9329
HPI	0.9013	0.8915	0.9239	0.9425	0.9212	0.9021	0.9091	0.9346
HDI	0.9051	0.8916	0.9144	0.9431	0.9206	0.9025	0.9089	0.935
LHN_I	0.9054	0.8906	0.9223	0.9446	0.9217	0.9057	0.9118	0.9349
PA	0.4744	0.6191	0.5889	0.5444	0.5558	0.4669	0.4684	0.5315
LP	0.9137	0.8994	0.9246	0.9492	0.9202	0.9177	0.9267	0.939
Katz	0.9136	0.8992	0.9258	0.9493	0.9316	0.9172	0.9267	0.9391
Spmap	0.9337	0.9387	0.9647	0.9573	0.9449	0.9294	0.944	0.9512

从表 2 可以看出,本文所提的基于空间映射算法 Spmap 在所有数据集上取得的预测结果的质量都是最高的,在 8 个数据集上的 AUC 得分值都高于其他 10 种算法;在某些数据集上,如在 IJCGA 上,在其他方法都低于 0.9 的情况下,Spmap 算法依然能够达到 0.9387。可见,采用空间映射的思想去尽可能地同化拓扑信息和属性信息的方法是行之有效的,在完成映射之后,矩阵比对也是切实可行的,它能使两种信息保持一致,缩小它们之间的差异。

图 3 示出了各种算法在各个数据集上的运行时间的比较结果。

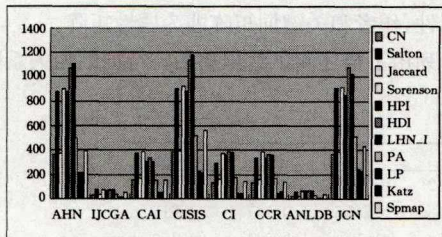


图 3 算法的运行时间

由图 3 可以看出,本文所提出的算法 Spmap 在各个数据集上的运算速度都快于 LHN-I, HDI, Sorenson, Salton, HPI 等算法的速度。同时不难发现,对于一些小的数据集,大部分算法的运行时间的差异不大。从图 3 中还可以看出,在 IJCGA, ANLDB 这两个小数据集中,所有算法的运行时间都在

一分钟以内,甚至更短。在 3 个较大的数据集 AHN, CISIS 以及 JCN 上,相较于另外一些算法而言,Spmap 算法并没有花费太多的时间,这说明算法 Spmap 具有较好的可扩展性。在大规模数据集上,算法 Spmap 可以使计算的时间复杂度合理地增长。另外,本文提出的算法需要处理属性信息,这会增加计算的复杂度。因此,综合考虑预测结果的质量,算法 Spmap 的运算速度是较快的。

结束语 针对顶点带属性网络的链接预测问题,文中提出一种基于空间映射的链接预测算法。考虑到网络拓扑信息和属性信息的异构以及异质问题,采用空间映射的方法把这两种信息映射到同一个空间中。提出分步交叉迭代的方法来取得最优的映射矩阵,以有机地融合拓扑信息和属性信息,在融合之后产生的新矩阵上完成相似度计算,从而预测链接存在的可能性。实验结果证明了空间映射思想的正确性,同时也验证了该方法确实能够有效融合拓扑和属性两种信息,提高了链接预测的精度。

参考文献

- [1] BHUYAN M H, BHATTACHARYYA D K, KALITA J K. A multi-step outlier-based anomaly detection approach to network-wide traffic[J]. Information Science, 2016, 348: 243-271.
- [2] ZHANG L, SUGANTHAN P N. A survey of randomized algorithms for training neural networks[J]. Information Science,

- 2016,364:146-155.
- [3] AHMED N M, CHEN L. An efficient algorithm for link prediction in temporal uncertain social networks [J]. *Information Science*, 2016, 331: 120-136.
 - [4] BUCCAFURRI F, FOTIA L, LAX G, et al. Analysis-preserving protection of user privacy against information leak age of social-network Likes [J]. *Information Science*, 2016, 328: 340-358.
 - [5] MICHAŁ L J, JAROSZEWICZ S, OSTROWSKI Istrowski Ł A. Wierzbicki, Verifying social network models of Wikipedia knowledge community [J]. *Information Science*, 2016, 339: 158-174.
 - [6] WANG P, XU B W, WU Y R, et al. Link prediction in social networks; the state-of-the-art [J]. *Science China Information Sciences*, 2014, 58(1): 1-38.
 - [7] GE M Q, LI A, WANG M H. A Bipartite Network-based Method for Prediction of Long Non-coding RNA-protein Interactions [J]. *Genomics, Proteomics & Bioinformatics*, 2016, 14(1): 62-71.
 - [8] BUCCAFURRI F, LAX G, NOCERA A, et al. Discovering Missing Me Edges across Social Networks [J]. *Information Sciences*, 2015, 319: 18-37.
 - [9] KLIMEK P, JOVANOVIĆ A S, EGLOFF R, et al. Successful fish go with the flow; citation impact prediction based on centrality measures for term-document networks [J]. *Scientometrics*, 2016, 107(3): 1265-1282.
 - [10] SUN Z, PENG Q K, LV J, et al. A prediction model of post subjects based on information lifecycle in forum [J]. *Information Science*, 2016, 337: 59-71.
 - [11] VIDMER A, ZENG A, MEDO M, et al. Prediction in complex systems; The case of the international trade network [J]. *Physica A: Statistical Mechanics and its Applications*, 2015, 436: 188-199.
 - [12] NIGAM A, CHAWLA N V. Link Prediction in a Semi-bipartite network for recommendation [J]. *Lecture Notes in Computer Science*, 2016, 9622: 127-135.
 - [13] XIE F, CHEN Z, SHANG J X, et al. A link prediction approach for item recommendation with complex number [J]. *Knowledge-Based Systems*, 2015, 81: 148-158.
 - [14] SAITO K, KIMURA M, OHARA K, et al. Super mediator-A new centrality measure of node importance for information diffusion over social network [J]. *Information Science*, 2016, 329: 985-1000.
 - [15] SARUKKAI R R. Link Prediction and Path Analysis Using Markov Chains [J]. *Computer Networks*, 2000, 33(1): 377-386.
 - [16] ZHU J H, HONG J, HUGHES J G. Using Markov Chains for Link Prediction in Adaptive Web Sites [C] // *Proceedings of the First International Conference, SoftWare 2002*. Belfast, Northern Ireland, April 2002: 60-73.
 - [17] AHN W, JUNG W S. Accuracy test for link prediction in terms of similarity index; The case of WS and BA models [J]. *Physica A: Statistical Mechanics and its Applications*, 2015, 429(1): 3992-3997.
 - [18] HOFFMAN M, STEINLEY D, BRUSCO M J. A note on using the adjusted Rand index for link prediction in networks [J]. *Social Networks*, 2015, 42: 72-79.
 - [19] HE Y L, LIU J N K, HU Y X, et al. OWA operator based link prediction ensemble for social network [J]. *Expert Systems with Applications*, 2015, 42(1): 21-50.
 - [20] SHERKAT E, RAHGOZAR M, AADPOUR M. Structural link prediction based on ant colony approach in social networks [J]. *Physica A: Statistical Mechanics and its Applications*, 2015, 419(1): 80-94.
 - [21] CHEN B L, CHEN L, LI B. A fast algorithm for predicting links to nodes of interest [J]. *Information Science*, 2016, 329: 552-567.
 - [22] DING J Y, JIAO L C, WU J S, et al. Prediction of missing links based on multi-resolution community division [J]. *Physica A: Statistical Mechanics and its Applications*, 2015, 417(1): 76-85.
 - [23] BARBIERI N, BONCHI F, MANCO G. Who to Follow and Why; Link Prediction with Explanations [C] // *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2014: 1266-1275.
 - [24] BLISS C A, FRANK M R, DANFORTH C M, et al. An Evolutionary Algorithm Approach to Link Prediction in Dynamic Social Networks [J]. *Journal of Computational Science*, 2014, 5(5): 750-764.
 - [25] SCRIPPS J, TAN P N, CHEN F L, et al. A Matrix Alignment Approach for Collective Classification [C] // *Proceedings of the International Conference on Advances in Social Networks Analysis & Mining*. 2009: 155-159.
 - [26] GONG N Z, TALWALKAR A, MACKEY L, et al. Jointly Predicting Links and Inferring Attributes using a Social-Attribute Network (SAN) [C] // *Proceedings of the 6th International Workshop on Social Network Analysis*. Beijing, China, 2012.
 - [27] LEROY V, CAMBAZOGLU B B, BONCHI F. Cold start link prediction [C] // *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Washington DC, USA, 2010: 393-402.
 - [28] ZHAO P X, LI X L, XIN D, et al. Graph Cube: On Warehousing and OLAP Multidimensional Networks [C] // *Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data*. 2011: 853-864.
 - [29] YIN Z J, GUPTA M, WENINGER T, et al. Linkrec: a united framework for link recommendation with user attributes and graph structure [C] // *Proceedings of the 19th International Conference on World Wide Web*. 2010: 1211-1212.
 - [30] YIN Z J, GUPTA M, WENINGER T, et al. A united framework for link recommendation using random walks [C] // *Proceedings of the 2010 International Conference on Advances in Social Network Analysis and Mining*. 2010: 152-159.