

一种基于资源传输路径拓扑有效性的链路预测方法

王 凯 李 星 兰巨龙 卫红权 刘树新*

(国家数字交换系统工程技术研究中心 郑州 450002)

摘 要: 链路预测旨在利用网络中已有的拓扑结构或其他信息, 预测未连边节点间存在连接的可能性。资源分配指标具有较低复杂度的同时取得了较好的预测效果, 但在资源传输过程的描述中缺少对路径有效性的刻画。资源传输过程是网络演化连边产生的重要内在动力, 通过分析节点间资源传输路径周围拓扑的有效性, 该文提出一种基于资源传输路径有效性的链路预测方法。该方法首先分析了节点间潜在的资源传输路径对资源传输量的影响, 提出资源传输路径有效性的量化方法。然后, 基于资源传输路径的有效性, 通过对双向资源传输量进行刻画, 提出了节点间传输路径的有效性指标。在12个实际网络数据集上的实验测试表明, 相比其他基于相似性的链路预测方法, 该方法在AUC和Precision衡量标准下能够取得更好的效果。

关键词: 复杂网络; 链路预测; 资源传输路径; 有效性

中图分类号: TN915, TP391

文献标识码: A

文章编号: 1009-5896(2020)03-0653-08

DOI: 10.11999/JEIT190333

A New Link Prediction Method for Complex Networks Based on Topological Effectiveness of Resource Transmission Paths

WANG Kai LI Xing LAN Julong WEI Hongquan LIU Shuxin

(National Digital Switching System Engineering and Technological R&D Center, Zhengzhou 450002, China)

Abstract: Link prediction considers to discover the unknown or missing links of complex networks by using the existing topology or other information. Resource Allocation index can achieve a good performance with low complexity. However, it ignores the path effectiveness of resource transmission process. The resource transmission process is an important internal driving force for the evolution of the network. By analyzing the effectiveness of the topology around the resource transmission path between nodes, a link prediction method based on topological effectiveness of resource transmission paths is proposed. Firstly, the influence of potential resource transmission paths between nodes on resource transmission is analyzed, and a quantitative method for resource transmission path effectiveness is proposed. Then, based on the effectiveness of the resource transmission path, after studying the two-way resource transmission amount between two nodes, the transmission path effectiveness index is proposed. The experimental results of 12 real networks show that compared with other link prediction methods, the proposed method can achieve higher prediction accuracy under the AUC and Precision metrics.

Key words: Complex network; Link prediction; Resource transmission path; Effectiveness

1 引言

作为复杂网络的重要研究方向之一, 链路预测近年来一直是网络科学领域^[1-4]的研究热点, 其旨

在利用网络中已有的拓扑结构或其他信息, 研究预测现有网络中任意两个未连接节点之间存在连接的可能性^[5-7]。基于网络结构相似性的链路预测方法仅仅利用网络拓扑信息, 具有计算简单且效果明显的特点, 受到各领域学者普遍关注。基于网络结构的相似性方法可分为基于局部拓扑、基于路径信息、基于随机游走等方法。

基于局部拓扑的链路预测方法主要利用共同邻居及其局部拓扑信息的方法, 包括共同邻居指标CN^[8]、资源分配指标RA^[9]、Adamic-Adar指标(AA指标)^[10]、社区联系的CN指标CAR^[11]、偏好连

收稿日期: 2019-05-13; 改回日期: 2019-09-10; 网络出版: 2019-09-19

*通信作者: 刘树新 liushuxin11@126.com, liushuxin11@gmail.com

基金项目: 国家自然科学基金(61803384), 国家自然科学基金创新研究群体项目(61521003)

Foundation Items: The National Natural Science Foundation of China (61803384), The National Natural Science Foundation Innovation Research Group Project of China (61521003)

接指标PA^[12]、余弦相似性指标Salton^[13]等。局部拓扑相似性方法对于网络结构信息的利用较少,预测效果相对一般。因此,有学者提出利用多跳路径计算节点相似性的方法。文献[14]对节点间的3阶路径进行了部分考虑,提出了局部路径指标LP。文献[15]将RA指标扩展到3阶路径上,提出了扩展的资源分配指标ERA。文献[16]通过计算所有的路径信息,提出了全路径指标Katz。基于随机游走的方法主要是通过随机游走对节点间的拓扑信息进行分析得出相应的相似度刻画方法,包括全局和局部的随机游走指标。其中,全局性的随机游走指标有平均通勤时间指标ACT^[17]、余弦随机游走相似性指标Cos+^[18]、重启的随机方法RWR^[19]等。

现有链路预测研究已经对许多演化机理进行了应用,却忽略了网络演化中的资源传输过程^[20]。资源传输过程作为复杂网络传播动力学的重要部分,对于连边的产生具有重要的影响^[21],也是网络演化的内在动力。当前,虽然有少量文献对资源传输过程进行了一定程度的研究^[9],但仍然缺少对资源传输过程的传输路径周围拓扑有效性。因此,资源传输过程中的传输路径^[22]有效性对于传输过程的刻画也具有的重要意义。

基于上述分析,本文提出一种基于资源传输路径有效性的链路预测方法。首先,通过详细分析资源传输过程中的多跳路径资源传输情况,提出节点间路径有效性的量化方法;然后,从传输资源角度,利用节点间传输路径量化方法,提出资源传输路径有效性指标,并在实际网络中验证了方法的有效性。

2 基于传输路径有效性的链路预测方法

2.1 节点间传输路径有效性分析与量化

传输路径是网络中资源传输过程得以顺利进行的主要通道^[23],如图1所示。由于多跳路径上存在大量的资源损耗和不确定性,相比长路径,短路径的资源传输能力更强,其对于节点间资源传输量的影响也相对较大。因此,在许多预测方法中,通常利用短路径进行资源传输过程的刻画,其效果较好且复杂度低。

网络内部多样化拓扑结构并存,传输路径周围

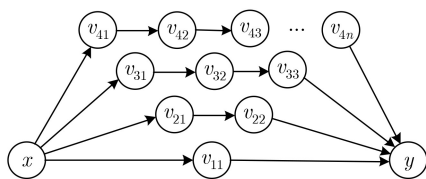


图1 网络中节点间多路径传输示意图

局部拓扑会对资源传输能力影响较大。图2给出了网络拓扑与路径有效性的关系示意图,其中对于始发节点 x 和终结节点 y , z_1 和 z_2 是其共同邻居节点。 l_{xz_1} 是关键路径,其周围存在大量点和边,可为资源传输提供潜在的传输路径。因此,通过量化传输路径周围拓扑结构对传输有效性的影响,对基于短路径资源传输的相似性指标刻画有重要价值。传统方法中,往往利用介数等^[24]统计特征刻画边的传输影响力。然而,由于具体计算中涉及全局的最短路径,复杂度非常高。此外,介数更侧重于去掉当前连边后对其传输能力的影响,而缺少局部拓扑对路径传输有效性的刻画。因此,需要从局部拓扑对传输过程影响的角度,分析刻画路径连边资源传输的有效性。

一般情形下,节点间路径数目是影响资源传输路径有效性的重要因素。图3中示意了不同网络结构下局部拓扑对路径有效性影响的对比。很明显,路径数目越多,可提供的潜在传输能力越大,对连边 l_{ij} 来说,其传输有效性则越高。一定程度上可以直接通过端点路径数目刻画传输的有效性,但需考虑所有跳的路径,其复杂度非常高。此外,端点间的路径数目固然非常重要,而端点本身邻居节点数目也决定了其可能接收到资源的量。因此,将从最短的2阶路径和节点度的角度分析路径节点间有效性的量化。

为了探讨分析局部拓扑中对路径节点间资源传输有效性影响的重要因素,图4构造了4个典型的拓扑结构进行对比。其中,图4(a)和图4(b)对应节点均具有相同的节点度,但图4(b)多一个共同邻居,故后者路径节点 i 和 j 间具有更高的资源传输能力,其路径节点间资源传输有效性更高;图4(b)和图4(c)具有相同的共同邻居数目,但节点度并不相同,图4(c)

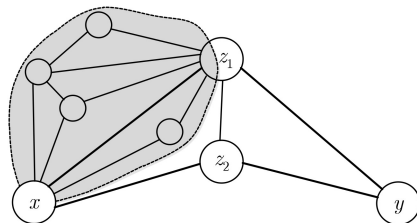


图2 网络拓扑结构与路径传输有效性的关系示意图

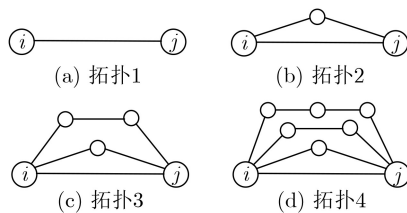


图3 不同网络结构下路径数目对比分析

的节点度较小, 因此对于路径节点*i*和*j*间的资源传输来说, 后者共同邻居具有更大的资源分配机会, 故图4(c)路径节点间的资源传输有效性更高; 图4(c)和图4(d)对应节点具有相同的节点度, 但图4(d)的共同邻居数目较多, 因此, 相比而言图4(d)路径节点间的资源传输有效性更高。因此, 路径节点间资源传输有效性与路径节点度成反比, 与共同邻居数目成正比。

经过上述分析, 对于网络中任意两个存在连接的节点*i*和*j*(图5所示), 其资源传输有效性可量化为

$$wl_{ij} = 1 + \frac{|\Gamma(i) \cap \Gamma(j)|}{\sqrt{k_i k_j}} \quad (1)$$

其中, $|\Gamma(i) \cap \Gamma(j)|$ 表示*i*和*j*间共同邻居的数目, k_i, k_j 分别为两个节点的节点度, 对于图5网络结构下, 资源传输紧密性 $wl_{ij} = 1 + 4/\sqrt{7 \times 8} = 1 + 2/\sqrt{14}$ 。对于未连边的节点对, 其局部拓扑也同样为其提供了潜在的资源传输可能。因此, 对于网络中任意两个未连接的节点*i*和*j*(图6所示), 其资源传输有效性可量化为

$$wl_{ij} = \frac{|\Gamma(i) \cap \Gamma(j)|}{\sqrt{k_i k_j}} \quad (2)$$

因为节点间没有直接连边, 故其有效性均来自周围拓扑结构。对于图6, 有效性为 $wl_{ij} = 4/\sqrt{7 \times 8} = 2/\sqrt{14}$ 。

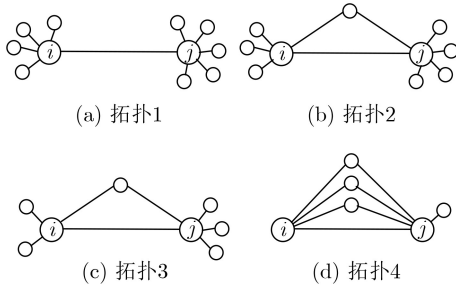


图 4 不同网络结构下路径有效性对比分析

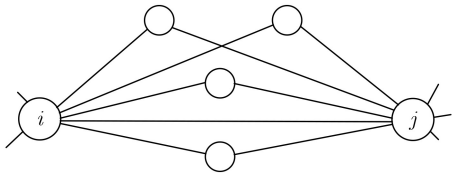


图 5 存在直接连接的两点之间路径有效性的量化

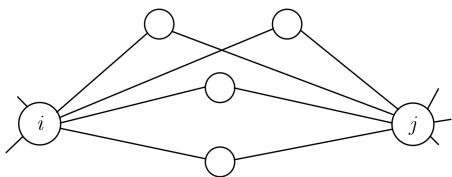


图 6 未直接连接的两点之间路径有效性的量化

结合上述分析, 对于网络中任意两个节点*i*和*j*, 从路径资源传输的角度, 其资源传输有效性可量化为

$$wl_{ij} = a_{ij} + \frac{|\Gamma(i) \cap \Gamma(j)|}{\sqrt{k_i k_j}} \quad (3)$$

其中, a_{ij} 表示邻接矩阵中*i*和*j*中的元素值。可以看出, 任意两点间的资源传输有效性量化分为两部分, 一部分是节点间本身连边的资源传输能力; 另一部分是周围局部拓扑结构对节点间传输能力的影响。

2.2 基于资源传输路径有效性的链路预测方法

为了从节点间路径有效性对资源传输影响的角度刻画任意两点间的相似性, 本文将基于路径传输有效性分析节点间资源传输能力。对于任意两个未建立连边的节点*v_x*和*v_y*, 其中*v_z*是其共同邻居, 如图7。假定节点*v_x*的连边上存在一个单元的资源经过节点间复杂拓扑结构传输到*v_y*, 则节点*v_y*接收到的资源量包含了经过网络本身共同邻居和*v_x*, *v_z*间路径有效性量化后的传输资源, 可以量化为

$$\begin{aligned} Ep(x \rightarrow y) &= \sum_{z \in \Gamma(y)} \frac{wl_{xz}^\psi}{W_z} \\ &= \sum_{z \in \Gamma(y)} \frac{\left(a_{xz} + \frac{|\Gamma(x) \cap \Gamma(z)|}{\sqrt{k_x k_z}} \right)^\psi}{W_z} \end{aligned} \quad (4)$$

ψ 为调节参数, 用于调节路径有效性的强度, W_z 表示节点*z*所有连边的路径有效性量化值之和, 即

$$W_z = \sum_{i \in \Gamma(z)} wl_{iz}^\psi \quad (5)$$

同样, 分析从*v_y*到*v_x*的资源传输量化过程。对于网络中的任意两个未建立连接的节点*v_x*和*v_y*, 其中*z*是节点*x*和节点*y*的共同邻居, 如图8。假定节点

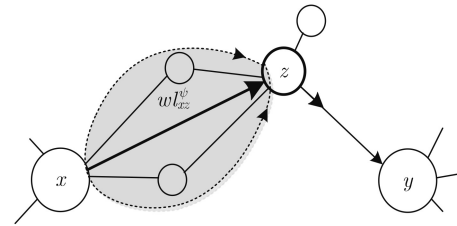


图 7 节点*x*到*y*传输路径有效性量化举例

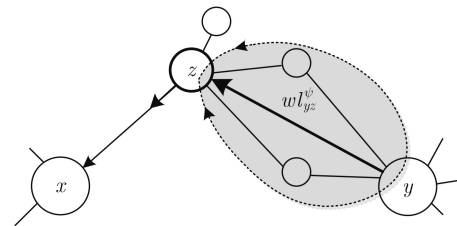


图 8 节点*y*到*x*传输路径有效性量化举例

v_y 的连边上存在一个单元的资源经过节点间复杂拓扑结构传输到 v_x , 则节点 v_x 接收到的资源量包含了经过网络本身共同邻居和 v_y, v_z 间路径有效性量化后的传输资源, 表示为

$$\begin{aligned} Ep(y \rightarrow x) &= \sum_{z \in \Gamma(x)} \frac{wl_{yz}^\psi}{W_z} \\ &= \sum_{z \in \Gamma(x)} \frac{\left(a_{yz} + \frac{|\Gamma(y) \cap \Gamma(z)|}{\sqrt{k_z k_y}} \right)^\psi}{W_z} \end{aligned} \quad (6)$$

在分析网络中 v_x 到 v_y 之间互相资源传输后的量化分析, 后续从资源传输角度重新刻画节点间的相似性。

定义1 节点间传输路径有效性指标(EP) 对于一个无权无向网络 $G(V, E)$, 其中任意两个未连接的节点 v_x 和 v_y 。两个节点之间传输路径有效性可用于量化其相似性, 其相似性可以定义为两节点 v_x 和 v_y 相互之间资源传输的本身和路径有效性影响下的潜在资源量之和, 具体表示为

$$\begin{aligned} s_{xy}^{Ep} &= Ep(y \rightarrow x) + Ep(x \rightarrow y) \\ &= \sum_{z \in \Gamma(x)} \frac{wl_{yz}^\psi}{W_z} + \sum_{z \in \Gamma(y)} \frac{wl_{xz}^\psi}{W_z} \end{aligned} \quad (7)$$

式(7)中包含了两部分, 一是资源从 v_x 到 v_y 的资源传输量刻画, 另一部分是从 v_y 到 v_x 的资源传输量刻画。

3 衡量指标及数据集介绍

3.1 算法衡量指标

在链路预测过程中, 一般将一个网络中的连边 E 分为训练集 E^T 和测试集 E^P , 其中 $E = E^T \cup E^P$, 且 $E^T \cap E^P = \emptyset$ 。评价链路预测算法预测效果的衡量指标主要有AUC和准确率(Precision)。

AUC衡量指标可以理解为在测试集 E^P 中随机选择一条边的分数值比未连接边的分数值大的概率^[25]。独立比较 n 次, 其中测试集中边的分数大于未连接边的分数则加1分, 测试集中边的分数等于未连接边的分数则加0.5分。分别用 n' 和 n'' 记录上述两种情况的个数, 则AUC指标表示为

$$AUC = \frac{n' + 0.5n''}{n} \quad (8)$$

准确率Precision指标(以下简称为Pre)可以理解为在前 L 个预测边中预测准确的比例^[26], 定义为

$$Pre = \frac{m}{L} \quad (9)$$

式(9)中, m 表示前 L 个预测结果排序中出现在测试集 E^P 中出现的个数。一般情况下, 设置为 $L = 100$ 。

3.2 网络数据

为了验证所提方法的效果, 本文选取了多个不同类型的实际网络进行实验, 分别介绍如下:

(1) AIDS-blog(AIDS): 一个AIDS博客相关的引用关系网络; (2)食物链网络FWEW: 一种生活在Everglades graminoids湿季的69种生物组成的网络; (3) Hamster(HS): 在hamsterster.com网页上的用户间朋友关系网络; (4) Figeys: 人类蛋白质相互作用网络; (5) UCsocial(UC): 美国加州大学UC在线社会网络; (6) Metabolic: 秀丽隐杆线虫的新陈代谢网络。上述网络的具体统计特征如表1所示, 包含节点数、集聚系数等参数。在后续实验中, 设置训练集合 E^T 中连边数目占比为0.9, 测试集 E^P 的连边占比为0.1, 每个数据测试点的结果均为20次结果的均值。

表1 网络数据特征参数

网络	AIDS	FWEW	HS	Figeys	UC	Metabolic
节点数	146	69	1858	2239	1899	453
边数	180	880	12534	6432	13838	2025
集聚系数	2.47	25.51	13.49	5.76	14.57	8.94
平均度	3.42	1.64	3.39	3.98	3.06	2.66
平均路径	-0.725	-0.298	-0.085	-0.331	-0.188	-0.226
匹配系数	0.052	0.552	0.0904	0.04	0.109	0.647

4 试验结果及分析

为了验证本文所提方法EP(后文简称为EP指标)的有效性, 本文选取了6个不同类型网络, 并利用AUC和Pre对其进行测试和结果分析。

4.1 AUC结果及分析

首先分析参数 ψ 对不同网络预测效果的影响。图9显示了不同网络中参数 ψ 对预测结果的影响。可以看出, 不同网络中 ψ 对EP指标的影响较大。FWFB和HS网络中, 在参数 ψ 较小时, AUC结果迅速提升, 其幅度在0.08左右, 随着 ψ 增大, AUC逐渐下降并持续保持一定程度的稳定; 在AIDS和Figeys网络中, $\psi = 0$ 时取得较大值均在0.96左右, 后续持续下降稳定在一定值, 说明传输路径周围拓扑对连边可能性以是否存在路径为主要影响; Metabolic中, 在 ψ 值较小时会呈现剧烈上升并达到极值, 而后轻微下降稳定在较高的AUC结果上, 这说明在这些网络中资源传输路径有效性对相似性刻画的影响较大; UC中在 ψ 值较小时会呈现剧烈上升, 到达极值后会随着参数 ψ 的增大呈现下降趋势, 说明此时资源传输路径有效性在低强度下对相似性刻画具有较大影响, 但随着有效性强度的增大, 其影响

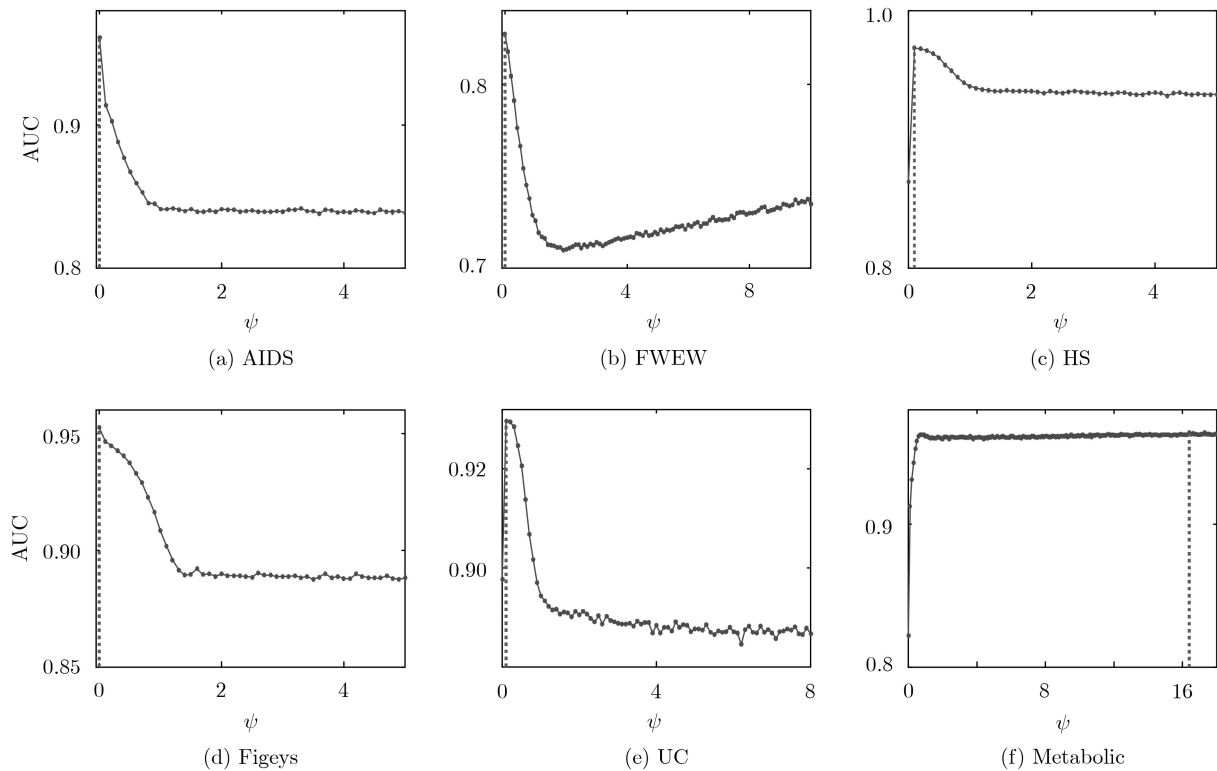


图9 调节参数对AUC结果的影响曲线图

逐渐减小。总之，多数网络中，在 ψ 较小时，AUC值均呈现较大程度的增长。因此在实际网络应用中，选取较小的 ψ 值便具备较高的链路预测准确性。

表2给出了多个不同类型网络中EP指标与其他相似性指标的AUC结果对比情况。可以看出，EP指标具有较高的预测精度，6个网络中AUC结果最高，一定程度上说明了资源传输路径有效性指标的有效性。

CN仅考虑了共同邻居，多数网络中CN的结果较差，仅在部分网络中表现略好。RA通过计算节点间传输的资源量量化相似度，其效果明显好于一

般局部相似性指标，尤其是FWEW和Metabolic网络中，部分甚至高于全局指标。AA指标利用共同邻居的节点度对共同邻居进行加权，其效果普遍好于CN。CAR指标考虑了共同邻居之间存在的部分连接对相似性的影响，但多数网络中比RA和AA效果略差。LP和Katz是基于长路径和全路径的相似性指标，在多数网络中效果较好且相对稳定。ACT指标是基于随机游走的相似性指标，多数网络中接近Katz等全局性相似性指标。Cos+同样是基于随机游走的相似性指标，相比ACT，部分网络中的预测效果有了一定的提升。在考虑了资源传输路径有效性后，EP指标在6个网络中表现最好。多个网络中如AIDS, FWEW, HS, Figcys和UC中，明显高于其他方法，其中对于局部指标提高比率为17%~67%之间，而对于全局指标Katz提高则在3%~17%左右。相比上述方法，EP指标平均提升比率为19%左右，最高可达68%。具体应用中， ψ 建议0.8左右，此时各个网络中表现普遍较高。

4.2 Pre结果及分析

图10所示为调节参数 ψ 对Pre结果的影响曲线图，与AUC相似，多数网络中在 ψ 较小时就能够取得很好的预测精度。其中，在Metabolic网络中，Pre结果随着 ψ 的增大迅速上升，上升幅度在0.03~0.50之间，之后呈现缓慢上升的趋势，说明资源传输路径有效性能显著提高链路预测中的Pre结

表2 AUC结果对比分析

方法	AIDS	FWEW	HS	Figcys	UC	Metbolic
CN	0.599	0.684	0.812	0.566	0.781	0.921
RA	0.609	0.702	0.816	0.570	0.787	0.959
AA	0.609	0.695	0.815	0.569	0.787	0.955
CAR	0.599	0.685	0.812	0.567	0.783	0.920
LP ^(a)	0.836	0.702	0.933	0.888	0.893	0.920
LP ^(b)	0.833	0.728	0.940	0.903	0.903	0.921
Katz ^(a)	0.854	0.704	0.933	0.887	0.893	0.920
Katz ^(b)	0.852	0.734	0.937	0.898	0.903	0.920
ACT	0.954	0.779	0.868	0.917	0.896	0.767
Cos+	0.591	0.510	0.960	0.844	0.869	0.904
本文方法	0.961	0.827	0.971	0.952	0.929	0.964

^(a)可调参数 $\alpha=0.001$ ^(b)可调参数 $\alpha=0.01$

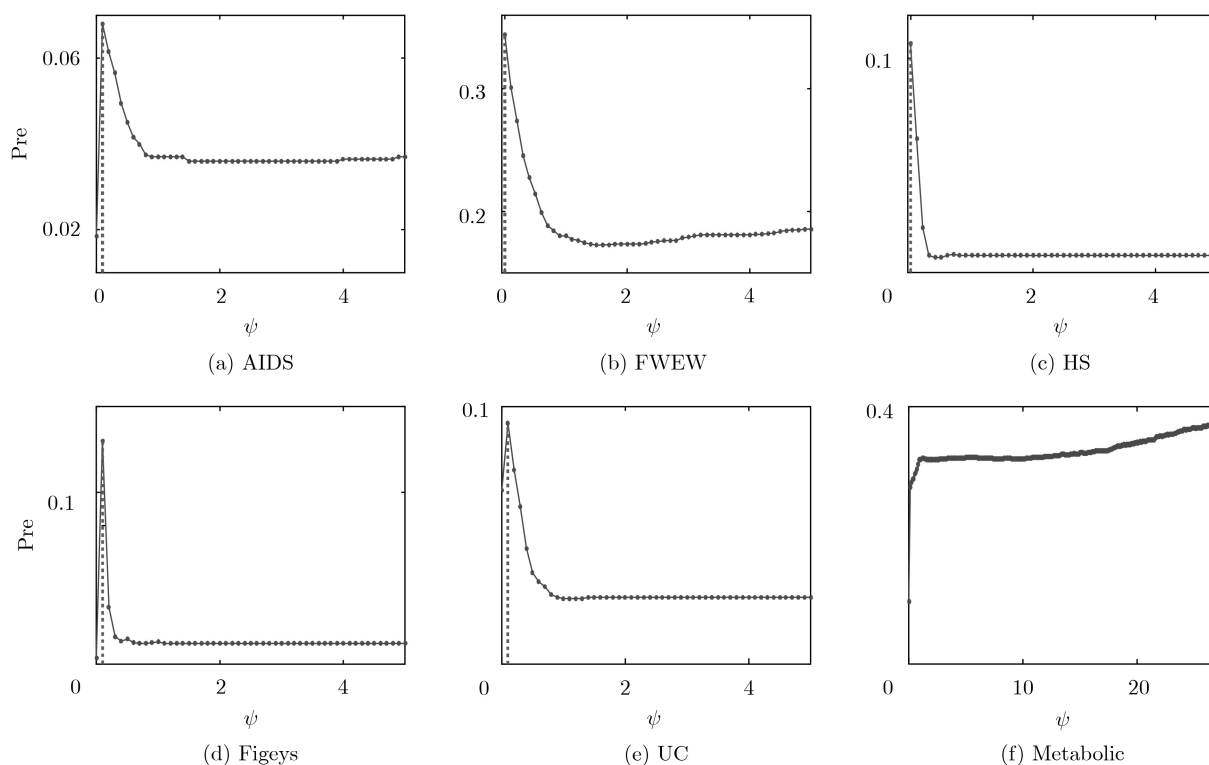


图 10 强度参数对Pre结果影响曲线图

果。部分网络中(AIDS, HS, Figeys和UC), Pre结果在达到最大值后随着参数的增大迅速下降, 并趋于平稳。在FWEW网络中, Pre结果在达到最大值后随着参数的增大迅速下降, 然后又缓慢上升, 说明资源传输路径有效性的强度对Pre的影响较为复杂。总体来看, 在调节参数 ψ 较小时, 多数网络中已经取得较高的Pre值。

表3给出了11个方法中资源传输路径有效性指标和其他相似性方法的对比结果。可以看出, 所提方法在11个方法的Pre结果明显好于基于局部相似性的CN, AA, RA和CAR指标和基于随机游走的

ACT和Cos+指标, 并且也明显好于全局相似性指标LP和Katz, 仅在Mac中排名第2。与AUC相似, 全局相似性方法一般好于局部相似性方法。不同的是ACT和Cos+指标在12个方法的Pre结果普遍较差, 可能对于随机游走相关的链路预测方法更适合AUC衡量标准。与其他方法相比, 所提方法在大多数网络中, Pre结果有比较明显的提高, 提升比率在1~10倍之间, 说明了所提方法的在多数方法中具有明显的效果。总体上, 相比上述对比的相似性方法, EP指标的Pre值平均提升3.69倍左右(部分对比方法如Cos+接近于0, 未作计算), 最高提升可达十几倍(部分方法的精度较低)。在实际应用中, ψ 建议0.2左右, 此时不同网络中的整体预测结果比较稳定且精度较高。

5 结束语

针对传输路径有效性对资源传输过程中的影响问题, 本文提出一种基于资源传输路径有效性的链路预测方法。本文方法首先从资源传输过程开始分析路径周围拓扑信息的重要影响, 以不同网络结构下连边有效性对比分析为突破点, 提出相应的节点间路径有效性量化方法。基于传输路径有效性, 对任意未连边节点间的双向资源传输量进行刻画, 提出资源传输路径有效性指标。多个实际网络数据中实验发现, 两种衡量标准下, 本文方法指标具有较好的预测效果。相比其他方法, 在AUC标准下,

表 3 Pre结果对比

方法	AIDS	FWEW	HS	Figeys	UC	Metbolic
CN	0.019	0.143	0.017	0.011	0.034	0.202
RA	0.028	0.165	0.008	0.012	0.026	0.319
AA	0.028	0.152	0.012	0.012	0.033	0.252
CAR	0.019	0.137	0.033	0.025	0.064	0.193
LP ^(a)	0.055	0.153	0.021	0.011	0.034	0.202
LP ^(b)	0.055	0.180	0.055	0.012	0.053	0.200
Katz ^(a)	0.055	0.153	0.021	0.010	0.034	0.202
Katz ^(b)	0.055	0.183	0.071	0.011	0.054	0.198
ACT	0.000	0.128	0.000	0.000	0.000	0.000
Cos+	0.000	0.000	0.015	0.005	0.010	0.097
本文方法	0.068	0.344	0.107	0.130	0.093	0.374

^(a)可调参数 $\alpha = 0.001$ ^(b)可调参数 $\alpha = 0.01$

本文方法平均提升比率为12%, 而Pre标准下, 可平均提升2.82倍。此外, 所提方法复杂度较低(与LP相当)可以应用于大型网络链路预测。

参 考 文 献

- [1] WANG Minggang, ZHAO Longfeng, DU Ruijin, *et al.* A novel hybrid method of forecasting crude oil prices using complex network science and artificial intelligence algorithms[J]. *Applied Energy*, 2018, 220: 480–495. doi: [10.1016/j.apenergy.2018.03.148](https://doi.org/10.1016/j.apenergy.2018.03.148).
- [2] GOSAK M, MARKOVIĆ R, DOLENŠEK J, *et al.* Network science of biological systems at different scales: A review[J]. *Physics of Life Reviews*, 2018, 24: 118–135. doi: [10.1016/j.plev.2017.11.003](https://doi.org/10.1016/j.plev.2017.11.003).
- [3] DU Wenbo, ZHANG Mingyuan, YING Wen, *et al.* The networked evolutionary algorithm: A network science perspective[J]. *Applied Mathematics and Computation*, 2018, 338: 33–43. doi: [10.1016/j.amc.2018.06.002](https://doi.org/10.1016/j.amc.2018.06.002).
- [4] CHEN Zhenhao, WU Jiajing, XIA Yongxiang, *et al.* Robustness of interdependent power grids and communication networks: A complex network perspective[J]. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2018, 65(1): 115–119. doi: [10.1109/TCSII.2017.2705758](https://doi.org/10.1109/TCSII.2017.2705758).
- [5] 王凯, 刘树新, 陈鸿昶, 等. 一种基于节点间资源承载度的链路预测方法[J]. *电子与信息学报*, 2019, 41(5): 1225–1234. doi: [10.11999/JEIT180553](https://doi.org/10.11999/JEIT180553).
WANG Kai, LIU Shuxin, CHEN Hongchang, *et al.* A new link prediction method for complex networks based on resources carrying capacity between nodes[J]. *Journal of Electronics & Information Technology*, 2019, 41(5): 1225–1234. doi: [10.11999/JEIT180553](https://doi.org/10.11999/JEIT180553).
- [6] BENSON A R, ABEBE R, SCHAUB M T, *et al.* Simplicial closure and higher-order link prediction[J]. *The National Academy of Sciences of the United States of America*, 2018, 115(48): E11221–E11230. doi: [10.1073/pnas.1800683115](https://doi.org/10.1073/pnas.1800683115).
- [7] LIU Shuxin, JI Xinsheng, LIU Caixia, *et al.* Similarity indices based on link weight assignment for link prediction of unweighted complex networks[J]. *International Journal of Modern Physics B*, 2017, 31(2): 1650254. doi: [10.1142/S0217979216502544](https://doi.org/10.1142/S0217979216502544).
- [8] LORRAIN F and WHITE H C. Structural Equivalence of Individuals in Social Networks[M]. LEINHARDT S. Social Networks: A Developing Paradigm. Lausanne: Academic Press, 1977: 67–98. doi: [10.1080/0022250X.1971.9989788](https://doi.org/10.1080/0022250X.1971.9989788).
- [9] ZHOU Tao, LÜ Linyuan, and ZHANG Yicheng. Predicting missing links via local information[J]. *The European Physical Journal B*, 2009, 71(4): 623–630. doi: [10.1140/epjb/e2009-00335-8](https://doi.org/10.1140/epjb/e2009-00335-8).
- [10] ADAMIC L A and ADAR E. Friends and neighbors on the web[J]. *Social Networks*, 2003, 25(3): 211–230. doi: [10.1016/S0378-8733\(03\)00009-1](https://doi.org/10.1016/S0378-8733(03)00009-1).
- [11] CANNISTRACI C V, ALANIS-LOBATO G, and RAVASI T. From link-prediction in brain connectomes and protein interactomes to the local-community-paradigm in complex networks[J]. *Scientific Reports*, 2013, 3: 1613. doi: [10.1038/srep01613](https://doi.org/10.1038/srep01613).
- [12] XIE Yanbo, ZHOU Tao, and WANG Binhong. Scale-free networks without growth[J]. *Physica A: Statistical Mechanics and its Applications*, 2008, 387(7): 1683–1688. doi: [10.1016/j.physa.2007.11.005](https://doi.org/10.1016/j.physa.2007.11.005).
- [13] SALTON G and MCGILL M J. Introduction to Modern Information Retrieval[M]. New York: McGraw-Hill, 1986.
- [14] LÜ Linyuan, JIN Cihang, and ZHOU Tao. Similarity index based on local paths for link prediction of complex networks[J]. *Physical Review E*, 2009, 80(4): 046122. doi: [10.1103/PhysRevE.80.046122](https://doi.org/10.1103/PhysRevE.80.046122).
- [15] LIU Shuxin, JI Xinsheng, LIU Caixia, *et al.* Extended resource allocation index for link prediction of complex network[J]. *Physica A: Statistical Mechanics and Its Applications*, 2017, 479: 174–183. doi: [10.1016/j.physa.2017.02.078](https://doi.org/10.1016/j.physa.2017.02.078).
- [16] KATZ L. A new status index derived from sociometric analysis[J]. *Psychometrika*, 1953, 18(1): 39–43. doi: [10.1007/BF02289026](https://doi.org/10.1007/BF02289026).
- [17] KLEIN D J and RANDIĆ M. Resistance distance[J]. *Journal of Mathematical Chemistry*, 1993, 12(1): 81–95. doi: [10.1007/BF01164627](https://doi.org/10.1007/BF01164627).
- [18] FOUSS F, PIROTTE A, RENDERS J M, *et al.* Random-walk computation of similarities between nodes of a graph with application to collaborative recommendation[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2007, 19(3): 355–369. doi: [10.1109/tkde.2007.46](https://doi.org/10.1109/tkde.2007.46).
- [19] BRIN S and PAGE L. The anatomy of a large-scale hypertextual Web search engine[J]. *Computer Networks and ISDN Systems*, 1998, 30(1/7): 107–117. doi: [10.1016/s0169-7552\(98\)00110-x](https://doi.org/10.1016/s0169-7552(98)00110-x).
- [20] YANG Dingda, LIAO Xiangwen, SHEN Huawei, *et al.* Dynamic node immunization for restraint of harmful information diffusion in social networks[J]. *Physica A: Statistical Mechanics and Its Applications*, 2018, 503: 640–649. doi: [10.1016/j.physa.2018.02.128](https://doi.org/10.1016/j.physa.2018.02.128).
- [21] 刘树新, 季新生, 刘彩霞, 等. 一种信息传播促进网络增长的网络演化模型[J]. *物理学报*, 2014, 63(15): 158902. doi: [10.7498/aps.63.158902](https://doi.org/10.7498/aps.63.158902).
LIU Shuxin, JI Xinsheng, LIU Caixia, *et al.* A complex network evolution model for network growth promoted by information transmission[J]. *Acta Physica Sinica*, 2014,

- 63(15): 158902. doi: [10.7498/aps.63.158902](https://doi.org/10.7498/aps.63.158902).
- [22] YAO Yabing, ZHANG Ruisheng, YANG Fan, *et al.* Link prediction in complex networks based on the interactions among paths[J]. *Physica A: Statistical Mechanics and Its Applications*, 2018, 510: 52–67. doi: [10.1016/j.physa.2018.06.051](https://doi.org/10.1016/j.physa.2018.06.051).
- [23] LIU Liang, QU Bo, CHEN Bin, *et al.* Modelling of information diffusion on social networks with applications to WeChat[J]. *Physica A: Statistical Mechanics and Its Applications*, 2018, 496: 318–329. doi: [10.1016/j.physa.2017.12.026](https://doi.org/10.1016/j.physa.2017.12.026).
- [24] SATHIYAKUMARI K and VIJAYA M S. Identification of Subgroups in a Directed Social Network Using Edge Betweenness and Random Walks[M]. SATAPATHY S C, BHATEJA V, and DAS S. *Smart Computing and Informatics*. Singapore: Springer, 2018: 115–125. doi: [10.1007/978-981-10-5544-7_12](https://doi.org/10.1007/978-981-10-5544-7_12).
- [25] WU Yiteng, YU Hongtao, ZHANG Jianpeng, *et al.* USI-AUC: An evaluation criterion of community detection based on a novel link-prediction method[J]. *Intelligent Data Analysis*, 2018, 22(2): 439–462. doi: [10.3233/IDA-173400](https://doi.org/10.3233/IDA-173400).
- [26] CHUAN P M, SON L H, ALI M, *et al.* Link prediction in co-authorship networks based on hybrid content similarity metric[J]. *Applied Intelligence*, 2018, 48(8): 2470–2486. doi: [10.1007/s10489-017-1086-x](https://doi.org/10.1007/s10489-017-1086-x).
- 王 凯: 男, 1980年生, 副研究员, 博士生, 研究方向为链路预测、社会网络分析.
- 李 星: 男, 1987年生, 助理研究员, 博士生, 研究方向为链路预测.
- 兰巨龙: 男, 1962年生, 教授, 博士生导师, 研究方向为新型网络体系, 网络动力学.
- 卫红权: 男, 1970年生, 副研究员, 硕士生导师, 研究方向为社团发现.
- 刘树新: 男, 1987年生, 助理研究员, 博士, 研究方向为复杂网络演化、链路预测.