逆链路预测方法研究综述

李 晶, 蒋忠元, 马建峰

西安电子科技大学 网络与信息安全学院 西安 中国 710071

摘要 链路预测旨在挖掘或预测隐藏的或即将出现的链路或链接, 已被广泛应用到诸多实际网络系统, 为用户挖掘潜在的高价值的关联关系。然而, 链路预测也可被攻击者用来攻击隐藏的敏感链接, 泄露用户隐私, 给用户与社会带来难以估量的损失。因此, 近年来, 链路预测的安全性引起了广大科研工作者的关注, 我们称之为逆链路预测问题。逆链路预测的核心思想为通过扰动一定数量的链路来降低链路预测方法对已隐藏的敏感链路的预测概率, 从而实现一定的安全性目标。现有机制从研究视角不同可主要分为三类: 基于对抗的逆链路预测模型, 基于鲁棒性攻击的逆链路预测分析, 基于隐私保护的逆链路预测研究。本论文系统综述目前链路预测、逆链路预测、逆链路预测防御方法的研究进展, 并对未来研究热点进行了展望, 为未来链路预测安全性研究提供基础。

关键词 逆链路预测; 鲁棒性; 链路预测对抗; 隐私保护
中图法分类号 TP309.2 **DOI 号** 10.19363/J.cnki.cn10-1380/tn.2021.03.03

A Survey of Reverse Link Prediction Methods on Graphs

LI Jing, JIANG Zhongyuan, MA Jianfeng

School of Cyber Engineering, Xidian University, Xi'an 710071, China

Abstract Link prediction aims to mine or predict links that have been hidden or will emerge in the near future. It has been widely used in many real network systems to mine potential high-value associations for users. However, link prediction can also be used by attackers to attack hidden sensitive links, which lead to the disclosure of user privacy and cause incalculable losses to users and society. Therefore, in recent years, the security of link prediction has attracted the attention of a large number of researchers, and we call it the reverse link prediction problem. The core idea of reverse link prediction is to reduce the probability of hidden sensitive links being predicted by the link prediction methods, which can be achieved by perturbing a certain number of links, and finally achieve the goal of security. The existing mechanisms can be divided into three categories from different research perspectives: the reverse link prediction models based on adversarial attack, the reverse link prediction analysis based on robustness and the reverse link prediction research based on privacy protection. This paper reviews the current research progress of link prediction, reverse link prediction and the defense of reverse link prediction methods. We also discuss the possible future research directions and provides a basis for future link prediction security research.

Key words reverse link prediction; robustness; adversarial link prediction; privacy protection

1 引言

在现实世界中, 许多系统都可以表示为网络(network/graph), 例如社交网络[1-2], 生物网络[3], 通信网络[4], 交通网络[5]等, 可谓网络无处不在。通常, 节点(node/vertex)为网络中的行为对象, 而链路或链接(link/edge)为对象之间的某类关联关系, 比如好友关系、邮件交流、化学反应、物质传递等。特别地, 某些网络系统(尤其是社交网络)是动态的, 并且此类网络中的链接始终随时间而不断变化[6-7]。基于此类网络数据进行分析, 进而判断网络中未发现的或将来会形成、更改的链接, 通常被称为链路预测过程。链路预测对我们进行网络深入分析、获取网络未知信息具有重要的作用, 能够为各种实际应用带来好处, 因此已被应用于许多重要的领域中。例如, 如果已知了恐怖分子的部分通信网络或亲属关系网络, 则可以通过链路预测来发现一些隐藏的链接, 从而发现潜在的恐怖分子[8]。链路预测也可以用于推荐系统[9-10], 网络重构[11]和节点分类[12]等。

在过去的几十年中, 许多关于链路预测的研究

通讯作者: 蒋忠元, 博士, 副教授, Email:zyjiang@xidian.edu.cn。
本课题得到国家自然科学基金(No. 61502375)与陕西省自然科学基础研究计划(No. 2020JM-203)资助。
收稿日期: 2020-04-07; 修改日期: 2020-05-20; 定稿日期: 2020-12-21

被相继提出。然而, 在链路预测被广泛用于网络分析的同时, 也导致了越来越多的个人信息可以在网上被他人获取, 引起了人们对于隐私问题的担忧。例如, 在社交网络中, 可能存在一些敏感链接, 表示两个人之间的某种关系, 比如通信关系、交易或恋爱关系[56]。这些敏感关系通常涉及个人隐私, 因此人们不愿意将其公开并将其隐藏。Mislove 等人[13]证明, 通过分析 Facebook 的社交网络结构, 不仅可以推断出其他 Facebook 用户的私人信息, 而且可以推断某些用户的属性。随着恶意攻击者利用链路预测等分析方法引发的各种攻击带来的严重安全威胁, 如何保护网络中的敏感隐私链接, 避免被链路预测这一网络分析工具攻击引起了越来越多研究者的关注。

针对链路预测可能导致的隐私泄漏这一问题, 大量关于个人隐私保护, 欺骗链路预测模型的方法被提出, 我们将这一类研究称为逆链路预测问题。逆链路预测研究的出发点与链路预测相反, 它试图分析各种链路预测模型的脆弱性, 产生可以对抗链路预测的方法以及保护网络中的敏感链接。基于研究的不同目的和适用场景, 我们将目前已有的逆链路预测方法分为了三类: 基于对抗的逆链路预测方法、基于鲁棒性攻击的逆链路预测方法和基于隐私保护的逆链路预测方法。

基于对抗的逆链路预测方法是指通过对原始网络进行一些扰动, 使得链路预测方法在扰动后的网络上的结果与本应得到的在原始网络上的预测结果大相径庭, 从而达到保护隐私链接的作用。由于链路预测旨在发现一些大概率存在但不可观察的链接, 因此针对链路预测的对抗方法也以一些敏感链接为隐藏目标, 也就是说必须防止目标链接被链路预测成功预测。链接干扰是此类研究中的一种常用技术[57], 例如, 数据发布者可以随机地修改原始网络上的链接以防止敏感链接被识别。它是一种针对链路预测的防御方法, 也是对链路预测的攻击方法。

基于鲁棒性攻击的逆链路预测方法研究了在各种链路预测对抗策略下一些预测方法的鲁棒性。尽管先前已经出现了许多有效的链路预测方法, 但其鲁棒性尚未得到广泛的讨论。链路预测在遭受欺骗时能否依然保持预测结果的准确性, 是在评价一个链路预测方法时同样需要考虑到的因素。上面的链路预测对抗方法被用来分析预测模型的脆弱性。

基于隐私保护的逆链路预测方法从隐私保护的角度切入, 对需要进行隐私保护的敏感链路进行隐私函数定义、证明与提出解决方案, 并通过大量的实验进行系统验证。

当然, 隐私保护方法也可能被攻击者不正当地使用。个人、社区和数据发布者可采用链路预测对抗网络作为应对过多网络分析的隐私保护工具, 但当黑客滥用对抗性攻击掩盖其非法社区时, 将会有极大的安全威胁。例如, 他们可以通过更改很小部分的链接来掩盖穆罕默德 • 阿塔在世贸中心恐怖主义网络中的领导地位[14]。有效的防御将协助反恐部门和执法机构根据大众对社交媒体的日益依赖发现罪犯和恐怖分子。许多基于链路预测方法的推荐系统也可能被欺骗, 这可能会使欺诈者将他们的产品推荐给无辜的用户。因此, 迫切需要衡量链路预测算法的鲁棒性, 研究链路预测对抗的防御方法。目前关于防御方法的研究还比较少, 我们将在后面的章节进行介绍。

本文对现有的逆链路预测方法及其防御对策进行了综述。首先从链路预测方法入手, 介绍了目前使用较为广泛的预测方法的分类及原理。接下来将从链路预测对抗、链路预测方法的鲁棒性和隐私保护三个方面来总结现有的逆链路预测工作。然后介绍了针对逆链路预测方法的防御策略。本文的主要贡献如下:

1. 我们对这一方面的工作进行了完整的综述, 总结了现有研究工作的核心贡献, 并根据较合理的依据在攻击和防御任务方面从系统的角度对它们进行了分类阐述。

2. 对现有研究的相关评价指标进行了系统梳理与对比分析, 对可用的网络资源等进行了梳理与介绍。

3. 综合分析了现有逆链路预测方法的优劣性, 展望了未来可能的研究方向。

2 链路预测

链路预测能够基于可观察的链接和网络中的其他信息尝试发现未知的链接或用来预测节点之间未来形成链接的可能性。它源于数据挖掘领域, 随着网络科学研究的蓬勃发展而引起了众多学者的关注。

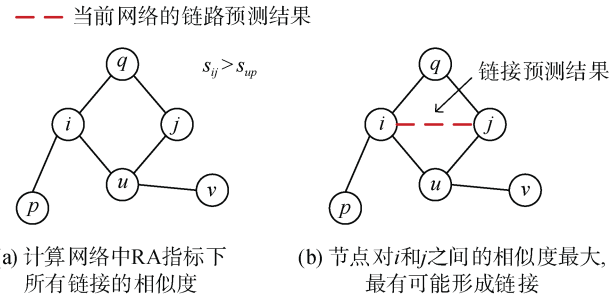
目前已经有许多成熟的链路预测方法可以被广泛地应用于各项研究领域, 这一章节将对现有的链路预测方法依据不同的分类进行阐述说明, 并对一些常见算法的原理进行说明。将预测方法分为了以下三大类: 基于相似性的链路预测方法、基于机器学习的链路预测方法以及基于最大似然的链路预测方法。

2.1 基于相似性的链路预测方法

此类方法将节点间相似性的大小作为判断它们

之间存在链接可能性的依据。由于节点的基本属性通常是隐藏不可获得的, 难以用来评定节点间的相似性, 因此此小节中列出的方法以网络结构的相似性为基础进行判别。其中, 每对节点 x 和 y 通过方法的计算公式会获得一个得分 s_{xy}, s_{xy} 就定义为 x 和 y 之间的相似度。网络中所有未观察到的链接均根据其相似度进行排名, 最终节点间相似度越高的链接被认为具有更大的存在可能性。不同方法的主要区别之一就是它们定义的节点相似性指标不同。图 1 描述了链路预测算法的基本过程, 使用 RA 指标来计算节点间的相似性, 首先利用链路预测算法对网络中所有不可观察的链接进行相似度的计算, 节点对间的相似度越大就越有可能形成链接, 网络中节点对 i 和 j 间的相似度最大, 因此被链路预测算法成功地预测出来。

这类链路预测方法的相似性指标共有 20 个, 可以分为三种不同的评价方向: 基于局部信息、基于全局信息以及基于类局部信息[15]。



(a) 计算网络中RA指标下所有链接的相似度　(b) 节点对i和j之间的相似度最大, 最有可能形成链接

图 1　链路预测算法过程的举例

Figure 1　An example of link prediction algorithm process

2.1.1　基于局部信息的方法

共同邻居 (common neighbors, CN)指标[16]是基于局部信息的方法中最基础的相似性指标, 它关注网络的局部信息结构, 计算节点对之间的共同邻居个数, 共同邻居越多则两个节点越相似。以 CN 算法为基础, 加入对网络中节点度的考虑, 可以得到下面其他的方法: Salton 指标[17]、Jaccard 指标[18]、Sorensen 指标[19]、大度节点有利指标 (hub promoted index, HPI) [20]、大度节点不利指标(hub depressed index, HDI)和 LHN-I 指标[21]。另一个只考虑节点度的方法为优先连接指标 (preferential attachment, PA)[58]。另外, 还有考虑到度不同的共同邻居节点的重要程度不同而设计的 Adamic- Adar (AA)指标[22], 以及从网络资源分配角度提出的资源分配 (Resource Allocation, RA)指标[23]。

这类相似性指标共有 10 种, 它们在简单高效的计算过程中充分考虑了节点的共同邻居或者节点度等网络中的局部信息, 并且预测的结果具有较高的预测精确度。由于它的计算过程较为简单, 复杂度低执行效率高, 有较好的准确性, 且适用的网络范围很广, 因此是使用的最为广泛的一类指标之一。表 1 列出了这 10 种基于局部信息的相似性指标的定义, 对于网络中的节点 x, 定义它的邻居为 $\delta(x)$, $k(x)=|\delta(x)|$ 为节点 x 的度。

表 1　基于局部信息的相似性指标定义

Table 1　Definitions of similarity metrics based on local information

名称	指标定义	名称	指标定义
CN	$s_{xy}=\lvert\delta(x)\cap\delta(y)\rvert$	HDI	$s_{xy}=\frac{\lvert\delta(x)\cap\delta(y)\rvert}{\max\{k(x),k(y)\}}$
Salton	$s_{xy}=\frac{\lvert\delta(x)\cap\delta(y)\rvert}{\sqrt{k(x)\times k(y)}}$	LHN-I	$s_{xy}=\frac{\lvert\delta(x)\cap\delta(y)\rvert}{k(x)\times k(y)}$
Jaccard	$s_{xy}=\frac{\lvert\delta(x)\cap\delta(y)\rvert}{\lvert\delta(x)\cup\delta(y)\rvert}$	PA	$s_{xy}=k(x)\times k(y)$
Sorenson	$s_{xy}=\frac{2\lvert\delta(x)\cap\delta(y)\rvert}{k(x)+k(y)}$	AA	$s_{xy}=\sum_{z\in\delta(x)\cap\delta(y)}\frac{1}{\lg k(z)}$
HPI	$s_{xy}=\frac{\lvert\delta(x)\cap\delta(y)\rvert}{\min\{k(x),k(y)\}}$	RA	$s_{xy}=\sum_{z\in\delta(x)\cap\delta(y)}\frac{1}{k(z)}$

2.1.2　基于全局信息的方法

基于全局信息的方法有 7 种, 包括 Katz 指标[24]、LHN-II 指标[21]、平均通勤时间(Average Commute Time, ACT) [25]、基于随机游走的余弦相似性(Cos+)指标[26]、重启的随机游走(Random Walk with Restart, RWR) [27]、SimRank 指标[28]以及矩阵森林指标(Matrix Forest Index, MFI) [29]。

与上面的基于局部信息的方法相比, 基于全局信息的链路预测方法要求提供完整的网络拓扑信息。尽管基于全局信息可以提供比基于局部信息更准确的预测, 但基于全局信息进行计算非常耗时, 因此这类方法通常不适用于大规模网络; 并且在现实世界中经常会有无法获得网络的完整拓扑信息的情况, 这使得此类链路预测方法的应用与上一类相比受到了限制。

2.1.3　基于类局部信息的方法

基于类局部信息的方法有 3 种, 即局部路径指标(Local Path Index, LP)[30]、局部随机游走(Local Random Walk, LRW) [31]和叠加的局部随机游走

(Superposed Random Walk, SRW)[31]。

基于类局部信息的链路预测方法是基于局部信息和全局信息的折衷方案。他的计算复杂度低于基于全局信息的方法, 它不需要网络的全局拓扑信息, 但比基于局部信息的方法考虑更多的网络结构, 在计算中放弃了对链路预测精确度没有贡献或贡献很小的多余网络信息, 预测结果通常具有更高的准确性。

2.2 基于机器学习的链路预测方法

现有的链路预测方法除了基于相似性的研究之外, 随着机器学习的蓬勃发展, 许多机器学习算法被应用到链路预测领域, 诞生了许多不同类型的基于机器学习的方法。这一小节将着重介绍现有的逆链路预测研究涉及的基于机器学习的链路预测方法, 主要与深度学习模型相关。

近年来, 随着深度学习模型的飞速发展, 网络嵌入算法在许多网络任务中都取得了卓越的性能, 例如: 节点分类[76-77]、链路预测[33]等。受语言模型word2vec[32]发展的启发, 诸如 DeepWalk[33], LINE[34]和 node2vec[35]等无监督学习的网络嵌入方法取得了巨大的成功, 这些模型学习得到的嵌入结果可以直接应用于网络的链路预测任务[33]。Kipf 等人[36]提出了用于链路预测的图自编码器(graph auto-encoder, GAE)模型, 它使用图卷积网络(Graph Convolutional Network, GCN)作为编码器, 受计算机视觉领域的卷积神经网络启发, GCN 可以直接在图上实现卷积[74], 这个过程可用一行简短的公式表达:

$$Z=\mathrm{GCN}\left(\boldsymbol{X},\boldsymbol{A}\right) \tag{1}$$

将 GCN 视为一个函数, 节点的特征矩阵 $\boldsymbol{X}$ 和邻接矩阵 $\boldsymbol{A}$ 作为输入, 输出 $Z\in R^{N\times f}$, N 表示网络节点数, f 为嵌入结果的维度, Z 代表的就是所有节点的嵌入结果; 使用节点特征的点积和激活函数作为解码器来重构原始的图, 即:

$$\hat{\boldsymbol{A}}=\sigma\left(\boldsymbol{Z}Z^{T}\right) \tag{2}$$

$\hat{\boldsymbol{A}}$ 就是所有节点对的分数矩阵, 对于分数大于阈值的链接, 则认为该链接应按预期存在, 从而可以形成重构网络。好的嵌入向量矩阵 $\boldsymbol{Z}$ 应该使重构出的邻接矩阵与原始的邻接矩阵尽可能的相似, 因为邻接矩阵决定了图的结构。因此, GAE 在训练过程中, 采用交叉熵作为损失函数:

$$L=-\frac{1}{N}\sum_{ij}A_{ij}\ln\left(\hat{A}_{ij}\right)-\left(1-A_{ij}\right)\ln\left(1-\hat{A}_{ij}\right) \tag{3}$$

并且根据损失函数使用梯度下降来优化模型中的参数。

由于深度学习方法具有非线性和分层的特性, 因此它们在链路预测及其他网络分析任务中显示出强大的功能和巨大的潜力。这些模型在保留网络初始结构和节点间邻近度的前提下, 将网络中的所有节点表示为低维的向量形式, 由此得到的嵌入结果可以直接应用于各种网络任务, 包括链路预测, 且有较好的预测效果, 近年来也得到了较为广泛的应用。对于常用的网络嵌入算法, 清华大学开源了一个较为完整的网络表示学习的 python 工具包①便于搭建与调试。

2.3 基于最大似然的链路预测方法

基于最大似然的链路预测方法的基本原理是: 根据网络结构的产生和组织方式以及目前已经观察到的链路计算网络的似然值, 并认为真实的网络使得网络似然值最大, 然后通过最大化网络的似然值来计算任何未观察到的节点间产生链接的可能性。该方法适合于处理具有明显层次组织的网络类型, 且对这类网络具有较好的精确度。

基于最大似然的方法可以再具体分为层次结构模型[61]和随机分块模型[62]两种。由于目前针对此类方法的逆链路预测研究较少, 本文将不再针对这两种模型进行单独地讨论。

从实际应用的角度来看, 最大似然法的一个明显缺点是非常耗时, 计算时间复杂度高。设计较好的该类算法能够在合理的时间内处理数千个节点的网络, 但肯定无法处理包含了数百万个节点的庞大在线网络。另外, 基于最大似然的方法与前面几种链路预测方法相比, 可能通常不是最准确的方法, 但是, 基于最大似然的方法为网络组织提供了非常有价值的见解, 而这是无法从其他链路预测模型中获得的。

3 逆链路预测

随着各类型链路预测技术的不断完善及其日益广泛的应用, 近两年来, 对于逆链路预测的相关研究也引起了众多的关注。逆链路预测的目的是欺骗或攻击链路预测模型, 使得到的结果与真实情况不符或隐藏网络中的部分链接使其不被预测模型发现, 从而使链路预测的结果失效。它试图分析各种链路预测模型的脆弱性, 产生可以对抗链路预测的方法以及保护网络中的敏感链接。

对于进行逆链路预测, 最常采用的手段就是添

① https://github.com/thunlp/OpenNE

加或删除网络中链接, 也可以称为链接干扰或重写, 通过这种手段就可以使网络的拓扑结构和一些节点间的相似性发生变化, 产生截然不同的预测结果, 从而隐藏网络修改者不希望被别人发现的链接或关系。不同方法的主要区别之一就是判断对哪些链接进行添加或删除的算法不同, 设计者期望通过改动尽可能少的链接数, 对预测结果的准确性产生尽可能大的影响。图 2 描述了逆链路预测算法的基本过程, 使用 RA 相似性指标作为链路预测方法进行计算, 目标链接在原始网络的预测结果中是得分最大的链接, 会被链路预测算法识别, 逆链路预测算法通过添加和删除部分链接的干扰方式形成了可用于公开的对抗网络, 最终链路预测算法对对抗网络进行预测时将得到错误的结果, 达到了目标链接被成功隐藏的目的。

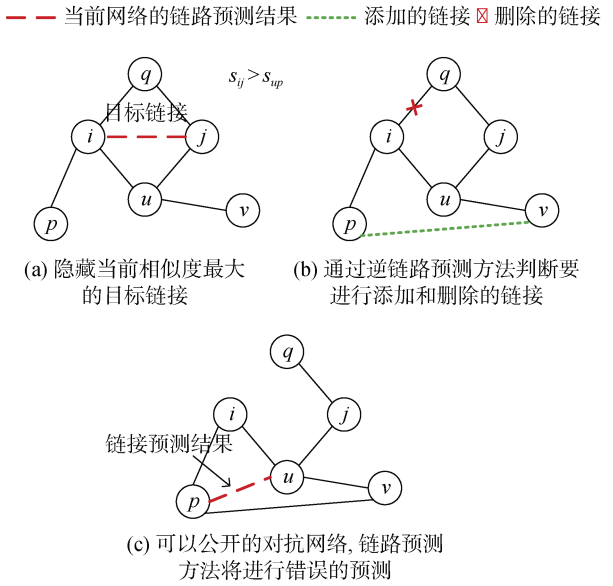


图 2 逆链路预测算法过程的举例

Figure 2 An Example of reverse link prediction algorithm process

本章节将对现有的逆链路预测研究进行全面的梳理和总结。根据研究目的和适用场景的不同, 逆链路预测研究可以分为以下三类: 基于对抗的逆链路预测方法、基于鲁棒性攻击的逆链路预测方法和基于隐私保护的逆链路预测方法。

3.1 基于对抗的逆链路预测方法

基于对抗的逆链路预测方法通常是从攻击者角度对链路预测模型进行的一种欺骗和对抗。由于链路预测的结果会依据概率由大到小对可能存在但不可观察的链接进行排序, 因此针对链路预测的对抗方法以一些敏感链接为隐藏目标, 尽可能降低其排序, 也就是说必须防止目标链接被预测方法成功预测。下文中的目标链接指的就是要进行隐藏的敏感链接。

由于基于对抗的逆链路预测方法是对链路预测模型的一种攻击, 它们通常是针对某一种链路预测模型提出的, 在近年来的研究中, 越来越多的研究者在确定所提出逆链路预测方法对该种预测模型具有较高攻击性的同时, 也倾向于判断该方法是否具有可转移性, 即证明所提出方法对其他种类的链路预测模型是否也具有攻击性。由于链路预测方法的种类较多, 因此具有可转移性的攻击方法也具有更广阔的应用场景, 引起了许多研究者的重视。下面的介绍中将对具有可转移性的逆链路预测方法进行特别地说明。

此类方法根据其算法依靠的基础方法的不同分为基于相似性指标的方法、基于深度学习模型的方法和基于其他算法的方法三类。

3.1.1 基于相似性指标的方法

此类方基于链路预测的相似性指标, 并结合其他策略来选择要进行添加或删除的链接, 通常也用于针对基于相似性的链路预测方法。

Zhou 等人[37]对通过链接删除来攻击基于相似性的链路预测问题进行了全面的算法研究, 重点研究了此类方法的两大类, 一类仅使用有关目标链接的局部信息, 另一类使用全局网络信息。基于局部信息的方法思路如下: 将 $U=\{u_i\}$ 表示为目标链接集 H 中的两端节点的集合, 称为目标节点。假设 $|U|=n$。$W=\{w_1,w_2,\cdots,w_m\}$ 是目标节点共同邻居的集合, 其中每个 $w_i \in W$ 连接到 U 中的至少两个节点。合理的攻击者只会删除 W 中的节点和 U 中的节点之间的边, 否则删除其他类型的边都不会导致两个 u_i 间共同邻居数减少, 从而使得节点间相似度减小。并且使用决策矩阵 $\boldsymbol{X} \in \{0,1\}^{m \times n}$ 表示 W 和 U 中节点之间的链接状态, 如果 w_i 和 u_j 之间存在链接, 则 $\boldsymbol{X}$ 的第 i 行和第 j 列的数值 x_{ij} 等于 1, 否则 $x_{ij}=0$。因此, 决策矩阵 $\boldsymbol{X}$ 完全捕获了目标链接集的总相似度 f_t。最终作者将基于局部相似度的攻击表述为优化问题来决定要进行删除的链接,

$$\min_{X} f_t(\boldsymbol{X}), \quad s.t.\ Sum\left(\boldsymbol{X}^0-\boldsymbol{X}\right) \leqslant k \tag{4}$$

其中 $\boldsymbol{X}^0$ 是原始决策矩阵, 而 Sum 函数表示逐元素求和, k 表示最多可以删除网络中的 k 条链接。通过求解上述优化问题得到新的决策矩阵 $\boldsymbol{X}$, 从而有节点

之间新的链接状态, 判断要进行删除的边。使用网络全局信息的方法原理与上面类似, 作者用 Katz 和 ACT 指标求解优化问题计算出了要进行删除的链接集合。该方法针对不同的相似性指标优化模型不同, 有较好的对抗效果, 当选择的相似性指标发生变化时, 要依据当前指标的特点设计不同的优化模型。

3.1.2 基于深度学习模型的方法

近年来, 深度学习模型在链路预测中显示了出强大的功能和巨大的潜力。然而, 另一方面, 深度学习模型的脆弱性也被揭露了出来。在计算机视觉领域, 精心设计的对抗样本容易欺骗深度学习模型, 这些样本对原始的图像略微施加干扰, 从而使模型无法获得正确的结果[38-39]。这些改动通常很难被人们注意到, 并且让深度学习模型以高置信度得到了错误的结果[38,40]。针对链路预测的深度学习模型是否也可以被对抗样本攻击, 从而产生错误的预测结果是这类方法的研究核心。

Chen 等人[41]提出并正式定义了链路预测的对抗攻击问题, 他们提出了一种基于图自编码器(GAE)中梯度信息的迭代梯度攻击(iterative gradient attack, IGA)方法, 并讨论了对其他预测方法的可转移性。由于在实际网络中, 不存在的链接通常比现有的链接要多得多, 换句话说, 负样本远远大于正样本, 因此作者将公式(3)中的损失函数构造为加权的交叉熵, 以防止负样本过度拟合:

$$L=\sum_{ij}-\omega A_{ij}\ln\left(\hat{A}_{ij}\right)-\left(1-A_{ij}\right)\ln\left(1-\hat{A}_{ij}\right) \tag{5}$$

其中, $\omega=\left(N^2-\sum_{ij}A_{ij}\right)-\sum_{ij}A_{ij}$ 是加权交叉熵的权重。该方法通过梯度信息判断要进行改动的链接, 与 GAE 训练过程的不同之处在于公式(5)中的损失函数考虑了邻接矩阵 $\boldsymbol{A}$ 中的所有链接, 而对抗攻击只需要考虑单个链接。对于目标链接 E_t 需要构造不同的目标损失函数:

$$\hat{L}=-\omega Y_t\ln\left(\hat{A}_t\right)-\left(1-Y_t\right)\ln\left(1-\hat{A}_t\right) \tag{6}$$

其中 $Y_t\in\{0,1\}$ 是目标链接 E_t 的真实链接状态, 而 $\hat{A}_t$ 是 GAE 计算得出的 E_t 存在的概率。在指定了损失函数的情况下, 则可以计算 $\hat{L}$ 对邻接矩阵的偏导数, 从而获得梯度矩阵:

$$\boldsymbol{g}_{ij}=\frac{\partial\hat{L}}{\partial A_{ij}} \tag{7}$$

进行链路预测时, 在梯度下降过程中会将损失函数 L 取到一个很小的值, 以获得良好的预测能力。逆链路预测方法则相反, 在这里需要最大化目标损失函数 $\hat{L}$, 以使模型错误地预测目标链接。梯度矩阵中的值可以为正或负, 正或负梯度表示最大化目标损失函数的方向是在邻接矩阵的相应位置中增大或减小该值。由于网络数据是离散的, 仅允许添加或删除链接的操作, 即只能在邻接矩阵中置 1 或置 0。要进行添加或删除的链接的选择取决于它们梯度的大小, 因为梯度表明了该链接对损失函数的影响程度, 数值越大, 链接对目标损耗的影响就越大。但要特别注意的是, 无论数值有多大都无法对其梯度为正/负的现有/不存在的链接进行修改, 这些边被认为是不可攻击的。与梯度下降过程相同, 对抗网络的生成也是迭代的。在每一次迭代中, 选择梯度最大并且同时可对其进行攻击的 n 条链接进行修改。通过将这些步骤重复 k 次, 就可以得到最终的对抗网络, 该网络可以欺骗链路预测方法。IGA 可以有效地对各种链路预测方法进行对抗, 有可转移性, 可攻击包括基于深度学习的链路预测方法(deepwalk, node2vec 等)和经典的基于相似性的链路预测方法 (CN 等)。

针对基于深度学习模型的链路预测, Chen 等人[42]还对动态网络进行了研究, 提出了对动态网络链路预测(dynamic network link prediction, DNLP)进行对抗性攻击的第一项研究。所提出的攻击方法为时间感知梯度攻击(time-aware gradient attack, TGA), 利用跨不同快照的深度动态网络嵌入(DDNE[43])生成的梯度信息来修改一些链接, 从而使 DDNE 无法准确预测目标链接。论文通过两种方式来实现 TGA: 一种是基于遍历搜索(traversal search)的方法, 即 TGA-Tra; 另一个为提高效率简化为贪婪搜索(greedy search), 即 TGA-Gre。TGA 在攻击 DNLP 算法方面具有出色的性能。与上面的方法类似, 根据梯度信息可以找到要进行修改的链接, 从而实现攻击。修改涉及节点 i 和 j 间链接的梯度 $g_t(i,j)$ 的大小和符号, 分别决定了候选链接以及应如何修改它们。为了降低修改成本, 修改的重点将放在梯度最大的链接上, 因为此类链接的变化对 $\hat{L}$ 的影响比其他链接更大。TGA 在 DDNE 上生成的对抗示例也可以用于有效攻击其他 DNLP 算法这一事实证明了 TGA 方法的可转移性。作者通过大量实验发现, 长时间的动态预测更容易受到对抗攻击的影响, 而已知更长的历史信息可以增强 DNLP 算法的鲁棒性。稀疏网络相对比较容易受到对抗性攻击的干扰, 也就是说, 稀疏网络上的 DNLP 算法不那么健壮。由于 TGA-Tra 比较了大量的修改方案, 因此通常是更有效的, 但是它具有相对较高的时间复杂度。TGA-Tra 的性能优于 TGA-Gre, 但后者的效率要高得多, 因此在实际

应用中更加实用。TGA-Tra 和 TGA-Gre 之间性能的显著差距也说明每次迭代中 $\hat{L}$ 的最大下降有时不会导致最佳的攻击性能。二者在一些具体细节和适用场景方面也有一些区别, 首先, TGA-Tra 更有可能在较早的历史快照上修改链接, 而 TGA-Gre 倾向于更改最新链接的链接; 其次, TGA-Tra 倾向于添加而不是删除链接, 而 TGA-Gre 具有相反的趋势。这样的观察表明, TGA-Tra 对网络进行的修改应该比 TGA-Gre 更不容易被发现, 因为人们倾向于更加关注最近的事件, 例如最近网络快照中的链接更改。另一方面, 如果想获得一些短期攻击效果, 则应该首选 TGA-Gre。此外, 由于在真实的社交网络中添加链接总是比删除链接容易, 因此 TGA-Tra 似乎具有较低的社交成本。

3.1.3 基于其他算法的方法

除了以上两种较有针对性的基于对抗的逆链路预测方法外, 还有一些方法是基于其他各种类型的算法提出的。

Yu 等人[44]提出了一种对抗 RA 指标链路预测模型, 防止敏感链接泄露的方法。在论文中主要运用了三类方法, 包括了随机链接干扰、启发式链接干扰以及进化式链接干扰。每种方法的具体实施过程如下: 随机链接干扰分为了链接重写和链接交换两种, 链接重写是随机删除和插入一定数量的链接, 链接交换指的是每次随机抽取两个节点对, 交叉连接, 这样可以保持所有节点的度不变。启发式链接干扰是一种专门针对链路预测方法精确度的方法, 降低测试集节点对在链路预测结果中的排名。它通过遍历降序的链路预测结果集, 对排名前 n 的节点对根据所属集合(训练集、测试集或者不存在的链接集合)执行不同的操作, 例如, 如果链接属于训练集则直接删除该链接, 这样该链接就可以作为相似度极高的未观察到的链接, 让预测者误以为预测到了有用的结果。进化式链接干扰运用了两种不同的算法, 包括遗传算法(Genetic Algorithm, GA)[63]和分布估计算法(Estimation of Distribution Algorithm, EDA)[64], 两种方法根据适应度函数来选择要进行添加和删除的链接, EDA 与遗传算法 GA 有着明显的区别。GA 采用交叉和变异等操作产生新个体, EDA 则通过对搜索空间采样和统计学习来预测搜索的最佳区域, 进而产生优秀的新个体。相比于 GA 基于基因的微观层面的进化方式, EDA 采用基于搜索空间的宏观层面的进化方法, 具备更强的全局搜索能力和更快的收敛速度。该方法还考虑了提出的进化式算法的执行效率问题, 对计算过程进行了改进来加速适应度的计算: 只计算相似度发生变化的链接集, 进行增量更新, 从而避免了冗余计算, 降低时间复杂度。研究发现在大部分网络中进化式链接扰动的防御效果更高, 尤其是分布估计算法, 启发式链接扰动的精确度更高, 这是意料之中的, 因为它的设计原理就是针对预测结果中相似度排名较高的链接。最后作者证明了分布估计算法产生的链接干扰结果具有可转移性, 可抵御基于节点对之间的高阶相似性的其他链路预测攻击, 例如 deepWalk 等部分节点嵌入算法。

3.2 基于鲁棒性攻击的逆链路预测方法

基于鲁棒性攻击的逆链路预测方法通常是研究者从理性的角度利用一些攻击手段或其他的方法对链路预测模型进行抗攻击程度的分析, 对链路预测方法的鲁棒性进行测试。

Wang 等人[45]在论文中研究了几种主流链路预测方法 CN、AA、RA、LP 及 Katz, 在多种网络攻击策略下的鲁棒性, 包括随机攻击(random attack, RDA), 基于中心性的攻击(centrality based attacks, CA), 基于相似性的攻击(similarity based attacks, SA)和基于模拟退火算法[65]的攻击(simulated annealing based attack, SAA)。其中 CA 包括基于中介性的攻击(betweenness based attack, BA)和基于权重的攻击(weight based attack, WA)两种。上述攻击方法的具体做法是不断删除每种攻击策略中计算得到的得分最大的链接。例如, SA 中现有链接的重要性由相似度指标的大小决定, 在每一次操作中删除当前相似度最大的链接。删除的链接数越多, 预测的精确度越低。作者通过大量的实验测试得出结论, 几种攻击方法中, 通常 SAA 具有最强的攻击有效性, 攻击效率最高, 其次是 SA, 最后是 CA, 有时某些 CA 攻击策略的性能, 例如基于中介性的攻击(BA)甚至比 RDA 还差, 这是一个令人惊讶的结论, 因为在其他场景中, BA 策略已被证明非常有效[46]。实验发现攻击对于链路预测的精度有巨大影响。在 5 种基于相似性的链路预测方法中, RA 可以比其他方法获得更精确的预测效果, 但是它非常容易受到网络攻击。因此作者得出链路预测方法鲁棒性的重要结论: 具有高性能的链路预测方法可能具有较低的攻击鲁棒性, 反之亦然。

Zhang 等人[47]提出目前为止的链路预测方法验证主要在假定的无噪声网络中进行, 缺少对如果观察到的网络数据不再准确将如何影响预测结果的清楚理解。在该文献中, 作者全面研究了在存在某些链路丢失, 伪造或与其他链路交换的真实网络中, 现有基于局部信息相似性的链路预测算法的鲁棒性。

提出了一种指标来量化和比较不同链路预测方法的鲁棒性, 它计算不同比例的噪声数据下, 预测精度曲线下的面积。其中, 随机噪声和偏置噪声均被考虑, 噪声实际上指的就是对网络中的链路进行改动。在将真实网络划分为训练集 E^T 和测试集 E^P 之后, 一些链接会随机添加到 E^T 或从 E^T 中删除。通过定义一个数量比率 *ratio*, 来测量随机添加或删除的链接占原训练集总链接数的比例。当比率为正时, $|ratio|*|E^T|$ 条链接被随机添加到训练集中; 当比率为负时, $|ratio|*|E^T|$ 条链接从训练集中随机删除。为了保持网络连接, E^T 中不能删除太多链接, 作者将范围设定为 $-40\% \leqslant ratio \leqslant 100\%$。在随机噪声中可以观察到, AUC 通常随 $|ratio|$ 的增大而减小。但在某些网络中, 随机添加一些链接可以提高 Jaccard 方法的准确性。这种现象与参考文献[48]中的结果相似, 可以通过添加一些链接来提高推荐的准确性。造成这种情况的根本原因是随机链接改善了网络的连通性, 使相似度矩阵更加密集, 因此使用这些随机添加的链接可以预测到许多由于连接性低而无法预测的链接。实验的另一观察结果是, 给定一种链路预测方法, 在给定相同 $|ratio|$ 的情况下, 随机删除链接比随机添加链接更具有破坏性。偏置噪声的施加方式是随机选择两个链接, $a-b$ 和 $c-d$, 然后将链接交换为 $a-d$ 和 $c-b$, 通过这样的方式保持节点度的大小不变, 但可以使网络随机化。这种噪声不会影响节点的度, 但会更改网络中的详细连接。两种噪声处理的实验证明对于链路预测的准确性, 丢失的链接比伪造和交换的链接更具破坏性。在研究的基于相似性的链路预测方法中, 某些方法的预测精度较低, 但它们往往在“噪声”的环境中更可靠。由于链接改动造成的偏差会降低链路预测算法的准确性, 从而更加有必要考虑在网络数据不干净的网络中链路预测算法的鲁棒性。论文提出了一种称为算法鲁棒性(algorithm robustness)的度量标准, 以量化链路预测算法可以在多大程度上抵抗网络中的噪声。

$$R=\frac{1}{|L|}\sum_{q=0}^{|L|}\frac{AUC(q)}{AUC(0)} \tag{8}$$

其中, $L=ratio*|E^T|$, $AUC(q)$ 是将 q 条链接添加到观察到的网络中时, 链路预测方法的 AUC 值, $ratio=0$ 时定义 $R=1$。当从训练集中删除链接时, R 也可以用于衡量算法的鲁棒性。显然, R 的大小取决于 *ratio*。在确定了 *ratio* 的范围之后就可以比较不同算法的鲁棒性。

Pouya 等人[49]为了分析链路预测模型的鲁棒性和可解释性, 提出了一种对图进行对抗性修改的方法, 引入了一种有效的方法通过近似嵌入结果发生的变化, 来估计当前改动的效果, 及对链路预测模型的影响大小。论文通过确定对链路预测模型影响性最大的修改方式来研究其可解释性, 并通过评估链路预测模型当原网络的链接发生添加或修改时的敏感性来研究其鲁棒性。作者想设计一种方法, 当最小程度地更改图结构时, 可以使目标链接重新获得的嵌入结果有最大程度地改变, 提出了一种称为通过对抗图编辑实现的鲁棒性和可解释性(Completion Robustness and Interpretability via Adversarial Graph Edits, CRIAGE)的算法模型。首先, 论文考虑了删除目标链接的相邻链接这一干扰方式, 来确定其中对目标链接的预测影响最大的链接; 还研究了将假链接添加到网络中的干扰方式, 以评估链路预测模型对网络发生少量链接添加时的鲁棒性和敏感性。

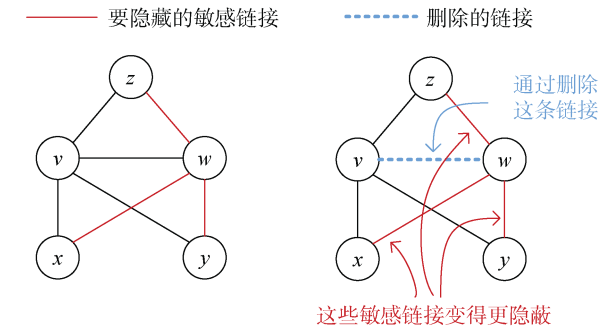
还有一些关于链路预测算法鲁棒性的结论。Marcin 等人[50]在论文中评估了 9 种不同基于相似性的链路预测算法的攻击耐受性, 发现它们的弹性随着网络节点数的增加而趋于增加, 随着网络平均度的降低而趋于降低。并且证明链路预测算法在较小的网络和密度较高的网络中操纵更容易受到攻击的影响。

3.3 基于隐私保护的逆链路预测方法

基于隐私保护的逆链路预测方法是网络中的个体或网络发布者想要保护一些敏感连接或隐私关系而产生的方法。

Marcin 等人[51]为社交平台中的用户设计了一种通过有策略地更改其链接, 使得其某些敏感链接无法被成功预测的方法, 该方法针对相似性的链路预测模型。论文提出了两种启发式方法, 可以让普通用户轻松地在现有的社交媒体上应用, 且实验结果表明这些启发式方法在各种网络上以及针对大量链路预测算法的有效性。第一种方法称为移除封闭三角形(Closed-Triad-Removal, CTR), 它从网络中删除有策略地选择的链接, 链接的删除要使网络中的封闭三角形个数减少。如果进行删除的链接会导致多个封闭三角形的消失, 且每个封闭三角形都包含要隐藏的敏感链接, 那么该算法甚至可以更有效。CTR 方法的设计旨在通过检查用户可以删除的所有可能选择并删除一个对敏感链接的隐藏最有效的链接, 它的主要思想如图 3 所示, 如果社交平台中用户 w 希望隐藏他与 x, y 和 z 的关系, 那么 CTR 建议 w 尽可

能多地取消与 x, y 和 z 的朋友的好友关系。这种方法可以轻松地应用于 Facebook 或 ins 中, 因为每个用户和他的任何朋友的共同总是可见的。在图 3 中, 通过删除 (v,w) 在网络中删除了三个封闭三角形: 一个包含节点 v, w, x, 另一个包含 v, w, y, 第三个包含 v, w, z, 因此 (w,x), (w,y) 和 (w,z) 的相似性得降低。

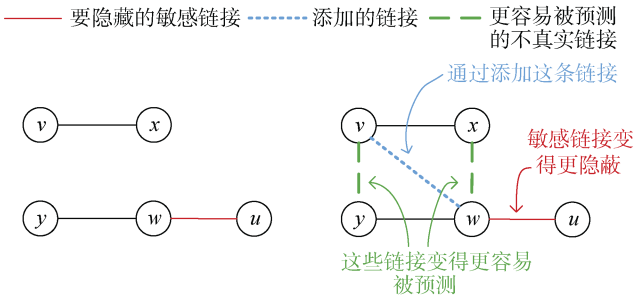


(a) 用户w对网络中三条敏感链接进行隐藏　(b) 删除可以使尽可能多的封闭三角形消失的链接

图 3　CTR 启发式方法的主要思想说明

Figure 3 An illustration of the main idea behind the CTR heuristic

另一种方法叫做创建开放式三角形(Open-Triad-Creation, OTC), 即通过添加新的链接形成不封闭的三角形。对于一条敏感连接 e, OTC 通过降低 e 的相似性得分来隐藏该链接, 同时可以增加 e 附近某些不存在链接的相似性得分, 这样的方式同样会降低 e 在所有不存在链接中, 基于相似度的排名中的位置, 从而降低链路预测算法识别出 e 的可能性。由于创建开放式三角形可以增加其中不存在链接的相似性得分, 因此通过添加新链接产生的开放式三角形越多则越好, 因为这可能会增加更多数量的不存在链接的相似性得分。根据这一观察结果, OTC 会检查用户可进行添加的所有可能选择, 并添加一条导致敏感链接在不存在链接中排名最大程度降低的链接。它的主要思想如图 4 所示, (v,w) 的添加会创建两个开放的三角形: 一个包含节点 x, v, w; 另一个包含 v, w, y。因此, (x,w) 和 (y,v) 的相似性得分增加, 而 (w,u) 的相似性得分降低。OTC 可以通过一种简单的方式应用于流行的社交媒体平台。例如, 如果用户 u 和 w 希望隐藏他们的关系, 那么他们中的任何一个(例如 w)都可以向好友列表中包含尽可能多的与自己无关的人的用户发送好友请求。即使很难找到这样的人, 仍然可以向高度联系的陌生人发送随机的友谊请求, 希望其中一些人会接受该请求, 这种做法是合理的, 因为估计有 55%的人在 Facebook 上接受来自完全陌生人的友谊请求[75]。两种启发式方法在实践中都是有效的, 前者通常比后者更有效一些。最后作者评估了 9 种不同的基于相似性的链路预测算法的攻击耐受性, 发现它们的弹性随着网络节点数的增加而趋于增加, 而随着网络平均度的降低而趋于降低。他们在另一篇论文中进行了补充性的研究[50], 表明有策略地选择要修改的链接非常关键, 因为随机地重新建立链接可能最终暴露出个人的敏感关系而不是隐藏所讨论的链接, 突出了有策略选择的重要性。



(a) 用户w对敏感链接(w,u)进行隐藏　(b) 添加可以使尽可能多的开放式三角形出现的链接

图 4　OTC 启发式方法的主要思想说明

Figure 4　An illustration of the main idea behind the OTC heuristic

本文作者蒋忠元等在文献[54]中提出目标隐私保护概念。实际网络中, 并非所有链接都是敏感的, 往往只有一小部分链接是重要且敏感的, 亟需隐私保护。因此, 我们将此类少量的重要且敏感的链接定义为目标, 因为它们往往是攻击者的攻击对象。我们提出基于网络模体(motif)的节点相似性定义, 其中模体是实际网络构建的单元, 比如三角形、四边形等。将两个节点的相似性定义为包含该两个节点间链路的某种模型的数量值, 比如参与的三角形个数, 即常用的基于共同邻居的相似性。目标集的隐私函数定义为所有目标在当前网络中的相似度之和。

实现目标隐私保护的主要步骤有: 1)从网络中删除目标集, 但仅删除目标远远不够, 因为现有链路预测算法可以有效预测出隐藏的目标, 因此需要进行步骤 2); 2)删除一定数量(预算, budget)的其他链路(被称为保护者, Protector)来进一步保护隐藏的目标, 使得现有的链路算法难以预测出隐藏的目标。

论文主体上探讨了三种场景下的目标隐私保护问题, 分别为: 1)单全局预算的保护链路选择; 2)多局部预算的跨目标链路选择; 3)多局部预算的顺序目标链路选择。论文对三种场景下多目标隐私保护分别进行了数学建模, 并理论证明了保护链路选择具有单调性和子模性, 并提出相应的贪婪算法进行求解。

为了进一步提高计算速度, 作者所有贪婪算法进行了加速, 并通过大量的实验验证。

此外, 作者对网络的可用性(utility)从端到端的距离、核数、特征值、模块度等方面进行了评估, 研究发现, 目标隐私保护一般针对少量的重要目标进行隐私保护, 隐私保护后的网络的可用性可以得到很好的保持。

更进一步, 作者还证明了在其他相似性指标(比如 RA, AA 等)下以及其他网络结构扰动机制(比如增加链路、重连链路)下隐私函数是否具有单调性与子模性。

进一步地, 作者在文献[55]中提出基于路径的相似性指标(其与 Katz 指标不同, 其选择的路径不包含回路)。定义了目标集的隐私函数, 通过理论证明了该隐私函数具有单调性与子模性, 进而提出贪婪算法可求得近似解。然而贪婪算法需要计算节点间符合一定路径长度的所有路径, 在大规模网络中变得不可用。作者提出使用无回路随机游走(self-avoid random walk)方法对需要保护的节点对之间的路径进行采样。基于采样得到的路径集, 论文提出的贪婪方法依然具有单调性与子模性, 在采样次数达到一定数量后, 得到的结果与群举法得到的结果非常相近, 但计算速度提升了成百上千倍。

本章节详细介绍了现有的逆链路预测算法的原理及核心贡献, 并在表 2 中对上述的所有逆链路预测方法进行了简单的对比与总结, 便于参考。

表 2　现有逆链路预测方法的对比与总结

Table 2　Comparison and summary of existing reverse link prediction methods

文献	方法模型	类型	基础算法	干扰方式	针对的链路预测模型	可转移性
[37]		对抗	相似性指标, 求解优化问题	删除链接	基于相似性的链路预测方法	
[41]	IGA	对抗	GAE	添加和删除链接	深度学习模型和基于相似性的预测方法	√
[42]	TGA	对抗	DDNE	添加和删除连接	各种动态网络链路预测方法	√
[44]		对抗	GA, EDA	添加和删除链接	基于相似性的链路预测方法, 节点嵌入算法	√
[45]	CA、SA、SAA	鲁棒性	中介中心度、相似性指标、模拟退火算法	删除链接	基于相似性的链路预测方法	
[47]	算法鲁棒性指标 R	鲁棒性		添加和删除链接	基于相似性的链路预测方法	
[49]	CRIAGE	鲁棒性		添加和删除链接	节点嵌入算法	
[51]	CTR, OTC	隐私保护		添加和删除链接	基于相似性的链路预测方法	
[54]	贪婪算法	隐私保护		删除链接	基于模体(motif)的相似性方法	
[55]	Walk2Privacy	隐私保护		删除链接	基于路径的相似性方法	

4　逆链路预测的防御

随着近年来逆链路预测方法的出现, 一些方法确实在隐私保护方面发挥了重要的作用, 然而, 存在恶意攻击者利用现有的逆链路预测方法对链路预测模型进行攻击, 以达到他们危害性的目的。因此, 对衡量并提高链路预测算法的鲁棒性, 研究链路预测对抗的防御方法的需求日益迫切。目前, 在这一方面的有关研究非常少, 本章节将对这些防御方法进行介绍。

Chen 等人[52]在这篇论文中第一次讨论针对网络对抗攻击的防御方法, 为 GNN 提出针对网络对抗攻击的防御策略。论文针对全局性和目标性的对抗性攻击提出了不同的防御策略, 并通过大量实验测试了其防御效率。所提出的防御策略的核心是平滑 GNN 训练过程中的梯度信息, 因此减小了可用于形成对抗网络的梯度的幅度。这篇论文与文献[41]中应用的网络对抗攻击方法对应。作者提出了四种防御策略, 为网络提供多样化的防御。针对全局网络攻击, 提出了全局对抗训练(Global Adversarial Training, Global-AT)以提高 DNN 的鲁棒性, 并针对目标性网络攻击提出了目标对抗训练(Target Label Adversarial Training, Target-AT)的方法。 此外, 还有平滑蒸馏和平滑交叉熵损失函数两种防御策略来实现梯度掩盖, 这使攻击者很难利用模型的梯度信息来构造攻击。

Zhou 等人[5]从博弈论的研究角度提出了一种用

来提高基于相似性的链路预测的鲁棒性的方法, 通过为对网络进行链路预测的分析人员提供一组有限的可靠查询, 这些可靠查询可以精确地获取所查询链接是否存在。分析人员旨在通过最佳地利用可靠的查询来稳健地预测可能的链接的集合, 要进行可靠查询的链接集需要有策略地选取。论文借助贝叶斯斯坦博格博弈模型对分析人员(防御方)与隐藏链接者(攻击方)的策略和收益进行分析, 在该博弈中, 分析人员首先可以进行可靠的查询, 接下来对手删除网络中剩余的一部分链接。模型中分析人员不能确定对手试图隐藏的特定目标链接, 而对手则具有有关分析人员和网络的完整信息。首先对网络数据收集进行建模, 分析人员提交一组节点对查询并会获得每个查询是存在的连接或不存在的链接的响应, 基于查询结果, 假设分析人员将构建一个子图并使用相似性度量来评估不在查询集中的链接存在的可能性。在论文方法的设置中, 攻击者可以通过将有限查询子集的链接进行删除的方式来修改查询结果, 以隐藏目标链接。例如, 犯罪分子会恐吓一些调查者, 以使他们不透露已知的隐秘关系。为了应对此类攻击, 作者假设分析人员可以使他们的查询子集可靠, 例如, 他们可以通过多次调查以及其他方式(比如监视通信)来确定特定的关系, 从而显着降低攻击者成功隐藏现有链接的可能性。实验证明了攻击者进行的攻击不会总是损害防御者, 由于这是一个非零和博弈, 因此在某些情况下攻击实际上可能会增加防御者的收益。在攻击会降低防御者收益的情况下, 论文提出了一种启发式算法仅通过一小部分可靠的链接查询就可以大大减少攻击的损害, 论文同时也表明, 如果不仔细选择要进行可靠查询的链接, 该方法可能不会有较好的效果, 特别地, 增加随机选择的可靠查询的数量有时可能会降低防御者的收益。

5 评价指标

在本节中, 我们将介绍逆链路预测任务中常见的度量标准、数据集以及一些对照方法, 这些指标和方法有些用于攻击任务中, 另一些则用于防御方案中。对照方法是多篇论文实验中都采用了的用于突显新提出算法优越性的基础方法, 特别地, 它们是针对攻击任务的对照方法。

5.1 常用指标

精确度(Precision)[67]。它是链路预测中最基础常用的指标, 计算的是在前 k 个预测结果中被成功预测的链接比例。如果有 m 个预测是准确的, 即排在前 k 的预测结果中有 m 个在测试集。

$$precision = \frac{m}{k} \tag{9}$$

AUC 值(Area Under Curve)[66]。它的大小是 ROC 曲线与坐标轴围成的图形的面积。AUC 值计算的意义是随机选择出一对正负样本时, 正样本的预测得分高于负样本的机率。随机地选取一条测试集中的边作为正样本, 再从网络所有不存在的链接中随机选取一条作为负样本, 若正样本的预测得分大于负样本, 那么就加 1, 若二者相等则加 0.5, 若负样本的预测指标值大于了正样本则不加分。假如正负样本比较了 n 次, 其中 n' 次正样本预测指标值大于负样本, n'' 次两者值相等, 那么 AUC 指标的计算公式定义如下:

$$AUC = \frac{n' + 0.5n''}{n} \tag{10}$$

排序得分(Ranking Score)[68]。它主要考虑测试集中的边在预测结果最终排序中的位置。令 H 为测试集中的边和网络中不存在的边的并集, r_i 表示测试集中链接 $i \in E^P$ 在排序中的排名。则该条未知边的 Ranking Score 值为 $RS_i = r_i / |H|$, 遍历所有测试集中的链接, 得到系统的 Ranking Score 值为:

$$RS = \frac{1}{|E^P|}\sum_{i \in E^P} RS_i = \frac{1}{|E^P|}\sum_{i \in E^P} \frac{r_i}{|H|} \tag{11}$$

这些指标常用于链路预测任务, 但有时也会用来测试逆链路预测任务的有效性。

5.2 攻击和防御的指标

攻击成功率(Attack Success Rate, ASR), 用于衡量攻击的有效性。ASR 是在一定的链接干扰预算内成功攻击目标的比率。也就是修改链接使得成功隐藏的目标链接与所有目标链接的比率。ASR 越大, 攻击效果越好。ASR 的公式如下:

$$ASR = \frac{\text{成功攻击的链接数}}{\text{所有攻击的链接数}} \tag{12}$$

平均防御率(average defense rate, ADR), 用于衡量防御的有效性。即有防御和无防御时攻击的 ASR 之差与无防御时攻击的 ASR 之比。ADR 越高, 防御效果越好。

$$ADR = \frac{\left| ASR^{without-defense} - ASR^{with-defense} \right|}{ASR^{without-defense}} \tag{13}$$

损失预防率(Damage Prevention Ratio, DPR)是用博弈论分析提高链路预测算法鲁棒性的论文中提出的指标。预防损失的定义是衡量可以通过防御预防的损失程度。假设 L_0 是防御者在没有攻击时的损失, L_A 是攻击策略 A 下的防御者的损失, L_D 是防御策

略 D 下的防御者的损失。更好的防御策略会得到更大的 DPR。

$$\mathrm{DPR}_{\mathrm{A}}^{\mathrm{D}}=\frac{L_{\mathrm{A}}-L_{\mathrm{D}}}{L_{\mathrm{A}}-L_{0}} \tag{14}$$

平均修改链接(Average Modified Links, AML), 用于衡量攻击算法的效率。AML 是为网络拓扑攻击而设计的, 它表示成功攻击时的平均干扰的链接数量的大小。假设攻击者用于攻击目标网络的可以干扰的链接数量有限, 那么经过修改的链接(添加或删除)数量会一直累积, 直到攻击者达到目标或预算用完为止。

$$AML=\frac{\text{所有修改的链接数量}}{\text{所有攻击的链接数量}} \tag{15}$$

算法鲁棒性, 用于量化链路预测算法可以在多大程度上抵抗网络中的虚假连接。即公式(8)。

5.3 常用数据集

表 3 总结了逆链路预测研究中常用到的一些数据集, 可以看出它们大部分都是规模较小的网络。

表 3 逆链路预测实验中常用的数据集

Table 3 Data sets commonly used in reverse link prediction experiments

	$\lvert V\rvert$	$\lvert E\rvert$
Jazz①[69]	198	2742
USAir②[70]	332	2126
Email③[71]	1133	5451
PolBlogs④[72]	1490	19090
Cora⑤[73]	2708	5429
DBLP⑥	317080	1049866

5.4 对照方法

随机攻击(Random Attack, RAN), 分为链接重写和链接交换两种。链接重写是随机删除原始网络中的链接, 同时随机连接最初未连接的节点对。链接交换指的是每次随机抽取两个节点对, 交叉连接, 这样可以保持所有节点的度不变, 但网络的连接方式却发生了无变化。RAN 是最简单的攻击方法, 经常作为对照方法用于突显新算法优越性。

相似性攻击(similarity based attacks, SA), 可以利用基于相似性的具体链路预测方法作为攻击方法进行链接删除与新算法对比, 比如, 基于共同邻居的攻击(Common-Neighbor-based Attack, CNA), 也经常作为对照方法。

6 未来的研究方向

目前, 针对逆链路预测问题所进行的研究并不多, 在这一领域仍有广阔的研究前景, 存在许多值得通过观察现有研究方法进行研究的问题。在本节中, 我们尝试介绍几个主要的研究方向。我们将从逆链路预测研究中的攻击、防御以及评价指标的角度分别进行阐述。

高效有效的算法。从目前提出的攻击方法来看, 最常用的数据集通常是较小的网络。当前, 大多数提议的方法都无法攻击大型图, 原因是它们需要处理大量的网络信息, 从而导致了较高的时间复杂度。为了解决这个问题, 文献[44]做出了一些努力, 即只计算相似度发生变化的链接集, 进行了增量更新, 从而避免了冗余计算, 同时保持了有效的攻击性能。然而, 该方法目前也尚不适用于较大规模的网络, 在大网络上执行效率较低。鉴于无法在大规模网络上进行攻击的缺点, 研究一种更有效且高效的攻击方法来解决这一实际问题是非常必要的。

可转移性。由于现有的链路预测方法有较多的种类, 其中, 每一种又有许多不同的模型, 因此, 探讨所提出逆链路预测方法的可转移性也越来越受到研究者的重视, 目前已经有一些具有可转移性的攻击方法被提出。具有可转移性的逆链路预测方法同时也具有更广阔的应用场景, 因此研究一种攻击性强且具有可转移性的方法是一个值得思考的研究方向。

逆链路预测的防御任务。目前只有少数文献[52-53]尝试研究和改善网络中的链路预测任务模型的鲁棒性, 并提出有效的防御方法。因此, 对于改进各种预测模型的鲁棒性, 提出现有逆链路预测方法的防御方法或将当前防御方法转移给其他模型具有宝贵的思考价值。

攻击成本计量。当前没有太多指标可用来研究攻击模型的效率。已知的方法通过所修改链接的数量粗略地衡量了成本, 引发了今后更多的对于设计

① http://deim.urv.cat/~alexandre.arenas/data/welcome.htm
② http://networkrepository.com/inf-USAir97.php
③ http://konect.uni-koblenz.de/networks/arenas-email
④ http://www-personal.umich.edu/~mejn/netdata/
⑤ https://linqs-data.soe.ucsc.edu/public/lbc/
⑥ http://snap.stanford.edu/

一个计量攻击成本的指标的思考, 即是否存在另一个角度来更精确地量化攻击和防御的成本。在真实的网络中, 通常添加链接的成本与删除链接的成本是不同的, 因此这两种链接修改方式的成本不同, 这需要一些更合理的评估指标。

链路预测与逆链路预测的博弈。链路预测有助于数据挖掘与丰富应用, 但存在隐私泄露, 造成网络安全风险。逆链路预测可加强敏感信息的隐私保护或提高网络安全系数, 但降低了网络链路的可预测性。二者之间需要平衡或折中, 这属于博弈范畴, 有待广大学者深入研究。

隐私保护与网络可用性。当前, 逆链路预测方法以扰动网络结构为主, 会造成网络的可用性(utility)下降。现有少量逆链路预测机制考虑了网络的可用性, 并进行了评估。但大多机制缺乏系统的分析与评估, 在未来研究工作中有待进一步加强。

7 总结

逆链路预测方法的研究具有巨大的潜力, 在近年来引发了越来越多的关注。通过总结现有的经验和研究来加深对于逆链路预测的理解是非常重要的, 这些理解不仅可以提供有关攻击和防御策略的知识, 而且还可以为后续的研究提供设计的思路和见解。

本文对现有的逆链路预测方法及其防御对策进行了综合性的讨论。我们对这一方面的研究进行了完整的综述, 总结了现有作品的核心贡献, 并根据合理的标准在攻击和防御任务方面从系统的角度对它们进行了分类阐述, 并且对现有研究的相关评价指标进行了系统梳理与对比分析, 对可用的网络资源及对照方法等进行了梳理与介绍, 最后综合分析了现有逆链路预测方法的优劣性, 展望了未来可能的研究方向。

参考文献

[1] Borgatti S P, Mehra A, Brass D J, et al. Network Analysis in the Social Sciences[J]. *Science*, 2009, 323(5916): 892-895.

[2] Wellman B. The Development of Social Network Analysis: A Study in the Sociology of Science[J]. *Contemporary Sociology: A Journal of Reviews*, 2008, 37(3): 221-222.

[3] Montoya J M, Sol R V. Small World Patterns in Food Webs[J]. *Journal of Theoretical Biology*, 2002, 214(3): 405-412.

[4] Ebel H, Mielsch L I, Bornholdt S. Scale-free Topology of E-mail Networks[J]. *Physical Review E*, 2002, 66(3): 035103.

[5] Latora V, Marchiori M. Is the Boston Subway a Small-world Network[J]. *Physica A: Statistical Mechanics and Its Applications*, 2002, 314(1/2/3/4): 109-113.

[6] Neal J W. “Kracking” the Missing Data Problem: Applying Krackhardt’s Cognitive Social Structures to School-Based Social Networks[J]. *Sociology of Education*, 2008, 81(2): 140-162.

[7] Liben-Nowell D, Kleinberg J. The Link-prediction Problem for Social Networks[J]. *Journal of the American Society for Information Science and Technology*, 2007, 58(7): 1019-1031.

[8] Anil A, Kumar D, Sharma S, et al. Link Prediction Using Social Network Analysis over Heterogeneous Terrorist Network[C]. *2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity)*. IEEE, 2016: 267– 272.

[9] Chen J, Wu Y, Fan L, et al. Improved spectral clustering collaborative filtering with Node2vec technology[C]. *2017 International Workshop on Complex Systems and Networks*. 2017: 330–334.

[10] Chen J, Lin X, Wu Y, et al. Double layered recommendation algorithm based on fast density clustering: Case study on Yelp social networks dataset[C]. *2017 International Workshop on Complex Systems and Networks*. IEEE, 2017: 242–252.

[11] Guimerà R, Sales-Pardo M. Missing and Spurious Interactions and the Reconstruction of Complex Networks[J]. *The National Academy of Sciences of the United States of America*, 2009, 106(52): 22073-22078.

[12] Bilgic M, Namata G M, Getoor L. Combining Collective Classification and Link Prediction[C]. *Data Mining Workshops, 2007. ICDM Workshops 2007. Seventh IEEE International Conference on. IEEE*, 2007: 381– 386.

[13] Mislove A, Viswanath B, Gummadi K P, et al. You are who You Know: Inferring User Profiles in Online Social Networks[C]. *The third ACM international conference on Web search and data mining - WSDM '10*, 2010: 251–260.

[14] Krebs V. Mapping Networks of Terrorist Cells[J]. *Connections*, 2002, 24(3): 43-52.

[15] Lü L, Zhou T. Link Prediction in Complex Networks: A Survey[J]. *Physica A: Statistical Mechanics and its Applications*, 2011, 390(6): 1150-1170.

[16] Lorrain F, White H C. Structural Equivalence of Individuals in Social Networks[J]. *The Journal of Mathematical Sociology*, 1971, 1(1): 49-80.

[17] Salton G and McGill M. J. Introduction to Modern Information Retrieval [M]. Auckland: McGraw-Hill, 1983.

[18] Jaccard P. The Distribution of the Flora in the Alpine ZONE.1[J]. *New Phytologist*, 1912, 11(2): 37-50.

[19] Sørensen T. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. Biologiske Skrifter/Kongelige Danske Videnskabernes Selskab 5: 1-34[J]. *biol:skr*, 1948, 5:1-34.

[20] Ravasz E. Hierarchical Organization of Modularity in Metabolic Networks[J]. *Science*, 2002, 297(5586): 1551-1555.

[21] Leicht E A, Holme P, Newman M E J. Vertex Similarity in Networks[J]. *Physical Review E*, 2006, 73(2): 026120.

[22] Adamic L A, Adar E. Friends and Neighbors on the Web[J]. *Social Networks*, 2003, 25(3): 211-230.

[23] Zhou T, Lü L, Zhang Y C. Predicting Missing Links via Local Information[J]. *The European Physical Journal B*, 2009, 71(4): 623-630.

[24] Katz L. A New Status Index Derived from Sociometric Analysis[J]. *Psychometrika*, 1953, 18(1): 39-43.

[25] Klein D J, Randić M. Resistance Distance[J]. *Journal of Mathematical Chemistry*, 1993, 12(1): 81-95.

[26] [26] 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.

[27] Brin S, Page L. The Anatomy of a Large-scale Hypertextual Web Search Engine[J]. *Computer Networks and ISDN Systems*, 1998, 30(1/2/3/4/5/6/7): 107-117.

[28] Jeh G, Widom J. SimRank: A Measure of Structural-context Similarity[C]. *The eighth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '02*, 2002: 538-543.

[29] Chebotarev P, Shamis E. The Matrix-Forest Theorem and Measuring Relations in Small Social Groups[J]. *Automation & Remote Control*, 2006, 58(9): 1505-1514.

[30] Lü L, Jin C H, Zhou T. Similarity Index Based on Local Paths for Link Prediction of Complex Networks[J]. *Physical Review E*, 2009, 80(4): 046122.

[31] Liu W P, Lü L. Link Prediction Based on Local Random Walk[J]. *EPL (Europhysics Letters)*, 2010, 89(5): 58007.

[32] Mikolov T, Chen K, Corrado G, et al. Efficient estimation of word representations in vector space[J]. *Computer Science*, 2013.

[33] Perozzi B, Al-Rfou R, Skiena S. DeepWalk: Online Learning of Social Representations[C]. *The 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2014: 701–710.

[34] Tang J, Qu M, Wang M Z, et al. LINE: Large-scale Information Network Embedding[C]. *The 24th International Conference on World Wide Web*, 2015: 1067-1077.

[35] Grover A, Leskovec J. Node2vec: Scalable Feature Learning for Networks[C]. *The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016: 855–864.

[36] Kipf T N, Welling M. Variational Graph Auto-Encoders[EB/OL]. 2016: arXiv:1611.07308[stat.ML]. https://arxiv.org/abs/1611.07308

[37] Zhou K, Michalak T. P, Waniek M, et al. Attacking similarity-based link prediction in social networks[C]. *The 18th International Conference on Autonomous Agents and MultiAgent Systems. International Foundation for Autonomous Agents and Multiagent Systems,* 2019.

[38] Goodfellow I. J, Shlens J, and Szegedy C. Explaining and harnessing adversarial examples[J]. *Computer Science*, 2014.

[39] Szegedy C, Zaremba W, Sutskever I, et al. Intriguing properties of neural networks[J]. *Computer Science*, 2013.

[40] Moosavi-Dezfooli S M, Fawzi A, Frossard P. DeepFool: a simple and accurate method to fool deep neural networks[C]. *Computer Vision & Pattern Recognition*. IEEE, 2016.

[41] Chen J Y, Shi Z Q, Wu Y Y, et al. Link Prediction Adversarial Attack[EB/OL]. 2018: arXiv:1810.01110[physics.soc-ph]. https://arxiv.org/abs/1810.01110

[42] Chen J Y, Zhang J, Chen Z, et al. Time-aware Gradient Attack on Dynamic Network Link Prediction[EB/OL]. 2019: arXiv:1911.10561[cs.SI]. https://arxiv.org/abs/1911.10561

[43] Li T S, Zhang J W, Yu P S, et al. Deep Dynamic Network Embedding for Link Prediction[J]. *IEEE Access*, 2018, 6: 29219-29230.

[44] Yu S Q, Zhao M H, Fu C B, et al. Target Defense Against Link-Prediction-Based Attacks via Evolutionary Perturbations[EB/OL]. 2018: arXiv:1809.05912[cs.SI]. https://arxiv.org/abs/1809.05912

[45] Wang K, Li L B, Pu C L. Robustness of Link Prediction under Network Attacks[EB/OL]. 2018: arXiv:1811.04528[physics.soc-ph]. https://arxiv.org/abs/1811.04528

[46] Holme P, Kim B J, Yoon C N, et al. Attack Vulnerability of Complex Networks[J]. *Physical Review E*, 2002, 65(5): 056109.

[47] Zhang P, Wang X, Wang F T, et al. Measuring the Robustness of Link Prediction Algorithms under Noisy Environment[J]. *Scientific Reports*, 2016, 6: 18881.

[48] Zhang F G, Zeng A. Improving Information Filtering via Network Manipulation[J]. *EPL (Europhysics Letters)*, 2012, 100(5): 58005.

[49] Pezeshkpour P, Tian Y F, Singh S. Investigating Robustness and Interpretability of Link Prediction via Adversarial Modifications[C]. *The 2019 Conference of the North*, 2019.

[50] Waniek M, Zhou K, Vorobeychik Y, et al. How to Hide One's Relationships from Link Prediction Algorithms[J]. *Scientific Reports*, 2019, 9(1): 12208.

[51] Waniek M, Zhou K, Vorobeychik Y, et al. Attack Tolerance of Link Prediction Algorithms: How to Hide your Relations in a Social Network[EB/OL]. 2018: arXiv:1809.00152[cs.SI]. https://arxiv.org/abs/1809.00152

[52] Chen J Y, Wu Y Y, Lin X, et al. Can Adversarial Network Attack be Defended?[EB/OL]. 2019: arXiv:1903.05994[cs.SI]. https://arxiv.org/abs/1903.05994

[53] Zhou K, Michalak T. P, and Vorobeychik Y. Adversarial robustness

of similarity-based link prediction[J], 2019, *arXiv preprint arXiv*:1909.01432.

[54] Jiang Z, Sun L, Yu P. S, et al. Target Privacy Preserving for Social Networks[C]. *36th IEEE International Conference on Data Engineering (ICDE), Dallas, Texas, USA, 2020*, accepted, 2020, ArXiv Preprint ArXiv:2002.03284.

[55] Jiang Z, Ma J, and Yu P. S. Walk2Privacy: Limiting target link privacy disclosure against the adversarial link prediction[C]. *2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA*, 2019: 1381-1388.

[56] Backstrom L, Kleinberg J. Romantic Partnerships and the Dispersion of Social Ties: A Network Analysis of Relationship Status on Facebook[C]. The 17th ACM conference on Computer supported cooperative work & social computing, 2014: 831–841.

[57] Hay M, Miklau G, Jensen D, et al. Anonymizing social networks[J]. *Computer Science Department Faculty Publication Series*, 2007.

[58] Barabási A L, Albert R. Emergence of Scaling in Random Networks[J]. *Science*, 1999, 286(5439): 509-512.

[59] Huang Z. Link prediction approach to collaborative filtering[J]. 2005, 141–142.

[60] Yang X, Zhang Z X, Wang K. Scalable Collaborative Filtering Using Incremental Update and Local Link Prediction[C]. The 21st ACM international conference on Information and knowledge management - CIKM '12, 2012: 2371–2374.

[61] Clauset A, Moore C, Newman M E J. Hierarchical Structure and the Prediction of Missing Links in Networks[J]. *Nature*, 2008, 453(7191): 98-101.

[62] Guimerà R, Sales-Pardo M. Missing and Spurious Interactions and the Reconstruction of Complex Networks[J]. *The National Academy of Sciences of the United States of America*, 2009, 106(52): 22073-22078.

[63] Nijhout F. An Introduction to Genetic Algorithms[J]. *Complexity*, 1997, 2(5): 39-40.

[64] Zlochin M, Birattari M, Meuleau N, et al. Model-Based Search for Combinatorial Optimization: A Critical Survey[J]. *Annals of Operations Research*, 2004, 131(1/2/3/4): 373-395.

[65] Kirkpatrick S, Gelatt C D, Vecchi M P. Optimization by Simulated Annealing[J]. *Science*, 1983, 220(4598): 671-680.

[66] Hanley J A, McNeil B J. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve[J]. *Radiology*, 1982, 143(1): 29-36.

[67] Herlocker J L, Konstan J A, Terveen L G, et al. Evaluating Collaborative Filtering Recommender Systems[J]. *ACM Transactions on Information Systems*, 2004, 22(1): 5-53.

[68] Zhou T, Ren J, Medo M, et al. Bipartite Network Projection and Personal Recommendation[J]. *Physical Review E*, 2007, 76(4): 046115.

[69] Gleiser P M, Danon L. Community Structure in Jazz[J]. *Advances in Complex Systems*, 2003, 6(4): 565-573.

[70] Colizza V, Pastor-Satorras R, Vespignani A. Reaction–Diffusion Processes and Metapopulation Models in Heterogeneous Networks[J]. *Nature Physics*, 2007, 3(4): 276-282.

[71] Guimerà R, Danon L, Díaz-Guilera A, et al. Self-similar Community Structure in a Network of Human Interactions[J]. *Physical Review E*, 2003, 68(6): 065103.

[72] Adamic L. A. and Glance N. The political blogosphere and the 2004 US Election[C]. *The WWW-2005 Workshop on the Weblogging Ecosystem*, 2005.

[73] McCallum A K, Nigam K, Rennie J, et al. Automating the Construction of Internet Portals with Machine Learning[J]. *Information Retrieval*, 2000, 3(2): 127-163.

[74] Kipf T N, Welling M. Semi-Supervised Classification with Graph Convolutional Networks[EB/OL]. 2016: arXiv:1609.02907[cs.LG]. https://arxiv.org/abs/1609.02907

[75] Nagle F, Singh L. Can Friends Be Trusted Exploring Privacy in Online Social Networks[C]. *International Conference on Social Network Analysis & Mining. IEEE*, 2009: 312–315.

[76] Tang J, Qu M, Mei Q Z. PTE: Predictive Text Embedding through Large-scale Heterogeneous Text Networks[EB/OL]. 2015: arXiv:1508.00200[cs.CL]. https://arxiv.org/abs/1508.00200

[77] Wang S, Tang J, Aggarwal C, et al. Linked Document Embedding for Classification[C]. *CIKM*. 2016.

李晶 于 2019 年在西安理工大学大学网络工程专业获得学士学位。现在西安电子科技大学网络空间安全专业攻读硕士学位。研究领域为网络空间安全。研究兴趣包括: 复杂网络与安全、链路预测与安全等。Email: li_jing@stu.xidian.edu.cn

蒋忠元 于 2013 年在北京交通大学大学信号与信息处理专业获得博士学位。现任西安电子科技大学副教授。研究领域为网络空间安全。研究兴趣包括: 隐私保护、大数据安全、社交网络计算、复杂网络安全、人工智能安全等。Email: zyjiang@xidian.edu.cn

马建峰 于 1995 年在西安电子科技大学通信与电子系统专业获得工学博士学位。现任西安电子科技大学教授。研究领域为网络空间安全。研究兴趣包括: 应用密码学、无线网络安全、数据安全、移动智能系统安全。
Email: jfma@mail.xidian.edu.cn