

一种基于数据场的复杂网络聚类算法

刘玉华 张翼 徐翠 晋建志

(华中师范大学计算机学院 武汉 430079)

摘要 针对当前复杂网络研究中聚类的热点问题,提出了一种基于数据场的复杂网络聚类算法,该算法通过一种基于互信息的方法计算出复杂网络中节点的重要性,然后通过数据场中节点的势来划分网络的簇结构。实验证明,该算法在计算时间和精度上具有一定的优势。

关键词 聚类,复杂网络,数据场,互信息,重要性,势

中图分类号 TP393.01

文献标识码 A

New Clustering Algorithm Based on Data Field in Complex Networks

LIU Yu-hua ZHENG Yi XU Cui JIN Jian-zhi

(School of Computer, Central China Normal University, Wuhan 430079, China)

Abstract Focusing on current clustering hot issues, the article presented a new clustering algorithm based on data field in complex networks. It calculates nodes' importance combining with a mutual-information method, and divides network cluster structures according to node's potential. Experiments show that the algorithm has certain advantages upon the accuracy and the computable complexity.

Keywords Clustering, Complex network, Data field, Mutual information, Importance, Potential

1 引言

簇结构(cluster structure)是复杂网络的拓扑结构属性之一(如图1所示),正所谓“物以类聚,人以群分”;同一簇中节点相似度较高,不同簇间节点相似度较低^[1];例如 WWW 可以看成是由大量网站社团构成,其中同一社团内部的各个网站讨论的是共同的话题。复杂网络聚类算法旨在寻找复杂网络中真实存在的簇结构。聚类算法的研究对分析复杂网络的拓扑特性、探索复杂网络隐藏的规律以及预测复杂网络的行为特性都具有十分重要的意义。簇结构的分析在数据挖掘、模式识别、统计等领域都有广泛的应用,比如挖掘恐怖组织、控制疾病的传播等^[2]。

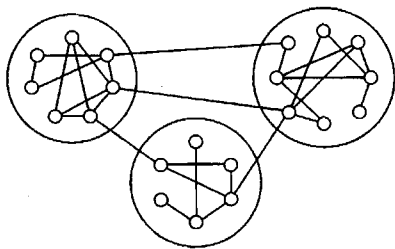


图1 网络簇结构示意图

复杂网络的聚类算法大致分为基于优化的聚类算法^[1]、启发式的聚类算法以及其他的一些聚类算法(见图2)。基于优化的聚类算法的思想是把聚类问题转化为优化问题,主要

有两类,一类是谱方法,另一类是局部搜索算法。启发式的聚类算法预先设定启发性规则来寻找簇结构,包括著名的 GN (Girvan-Newman)算法和 MFC(maximum flow community)算法等^[2]。

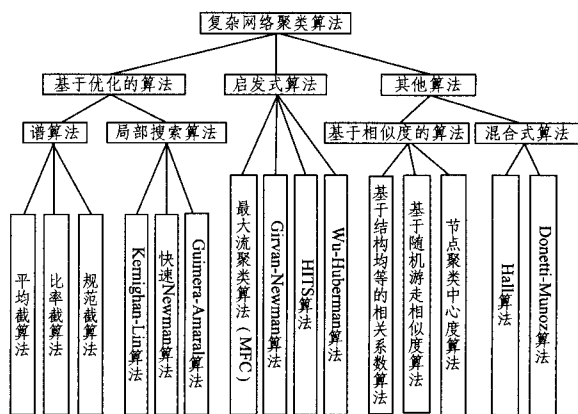


图2 复杂网络聚类算法分类图^[2]

现有的复杂网络聚类方法都具有一定的缺点,在计算精度、时间复杂度上无法面面俱到,并且需要外部监督(即给出先验条件和参数)。通过比较分析以上几种聚类算法可知,计算准确的聚类算法往往时间复杂性高于 $O(n^2)$,而计算快的聚类方法往往准确度不让人满意,因此设计一种又快又准的复杂网络聚类方法是一个亟待解决的问题^[3]。

本文提出了一种基于数据场的复杂网络聚类算法:利用

到稿日期:2013-01-25 返修日期:2013-05-10 本文受华中师范大学中央高校基本科研基金项目(120002040476)资助。

刘玉华(1951—),女,博士,教授,博士生导师,主要研究方向为计算机网络与通信技术、无线网络、复杂网络等,E-mail: yhliu@mail.ccnu.edu.cn (通信作者);张翼(1987—),男,硕士生,主要研究方向为复杂网络和无线网络;徐翠(1988—),女,硕士生,主要研究方向为复杂网络;晋建志(1982—),男,博士生,主要研究方向为复杂网络、网络通信。

物理学中数据场的理论,首先根据一种基于互信息的方法对复杂网络的节点重要性进行评估,挖掘网络中的骨干节点,然后利用这些骨干节点划分复杂网络的簇结构。实验证明,该方法能准确地对网络进行聚类分析,计算速度快并且可以根据实际需求来决定划分的粒度。

2 数据场与势

分布在一定空间区域的对象之间是相互联系的,对象与对象之间相互影响,不同对象的影响辐射范围不一样。19世纪早期,物理学家法拉蒂提出“场”的概念来描述这种对象间的相互作用。如果物理量在空间的不同位置仅有数量上的区别,则称对应的“场”为标量场;反之,如果物理量在空间的不同位置还有方向上的区别,则称对应的“场”为矢量场,如电场、磁场、重力场都是矢量场。通常将粒子间的相互作用及“场”方法引入数域空间,用来描述某个物理量或数学函数的分布规律^[4]。我们将数域空间看成一个数据场,每个对象都对其他对象产生作用,同时每个对象也受到其他对象的影响。

在复杂网络中,节点之间也相互产生影响,重要节点的影响范围较大,其他节点的影响范围相对较小。因此可以利用这种特性来对复杂网络进行聚类,在这里我们把复杂网络看成一个势场,应用数据场理论来对复杂网络进行聚类分析。

势场是一种标量场,通常使用势函数来描述势场。势函数是随着位置产生变化,并且可以叠加,等势线是用来划分势函数在空间中的分布规律,类似于物理学中的等高线。因此,数域空间中每一个数据对象对场中任何一点的势都有贡献,且贡献的大小与两者间的距离平方成反比。显然数据密集的地方势较大,数据稀疏的地方势相对较小。

定义 1 空间 Ω 中共有 n 个对象 $\{x_1, x_2, \dots, x_n\}$, 令 x_i 代表对象所处位置,则在该数据场中 x 处对象 x_i 所产生的势为:

$$\phi_i(x) = m_i e^{-\left(\frac{\|x-x_i\|}{\vartheta}\right)^2} \quad (1)$$

式中, $\|x-x_i\|$ 为对象 x_i 到位置 x 的距离,可以用拓扑距离、欧式距离来代表; m_i 为对象 x_i 的质量,满足归一化条件,通常取值是一些代表对象特性的值,这里我们使用节点在复杂网络中的重要性来代表节点的质量,具体选择第 3 节中提出的基于互信息的方法; ϑ 称为影响因子,用于控制对象间的相互作用程度(见图 3)。

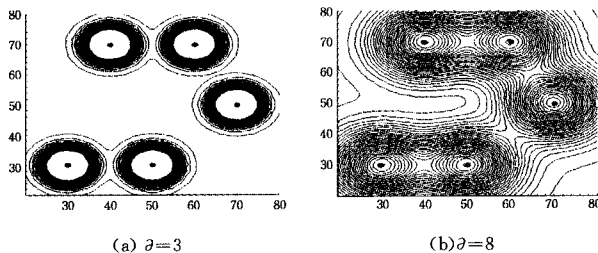


图 3 不同的 ϑ 对簇密集程度的影响^[5]

由于正态分布的对称性和集中性,存在这样一个法则:其 99.993666% 的概率分布在均值左右 4 个标准差的范围内,也就是说当距离大于 $\sqrt{8\vartheta}$ 时,势已经几乎为零了,因此可以把每个数据对象的影响范围也即半径看成 $\sqrt{8\vartheta}$, 这样就可以适当地减少一些计算量。

定义 2 空间 Ω 中 x 处的势函数为所有对象的势叠加,即

$$\phi(x) = \sum_{i=1}^n \phi_i(x) = \sum_{i=1}^n m_i e^{-\left(\frac{\|x-x_i\|}{\vartheta}\right)^2} \quad (2)$$

可见,某点的势是由距离和质量所决定的,本文使用最短距离;由于质量客观代表了吸引力,这里我们用复杂网络节点的重要性来代表节点的质量。

3 基于互信息的节点重要性评估方法

量化分析节点的重要性是复杂网络领域中一个极其重要的基本问题。对节点的重要性进行评估,发掘重要的节点,具有很高的实用价值。评估网络中节点重要性的方法很多,包括:度数法、介数法、生成树法^[6]、PageRank 算法等^[7]。不同的方法是从不同的角度来评估节点的重要性,所以每种方法都有自己的优缺点(见表 1)^[8]。节点的重要性取决于两方面:节点的网络拓扑和邻居节点的重要性。本节从网络拓扑结构出发,同时考虑节点之间连接的强弱和多少来量化节点的重要性,提出一种基于互信息的评估指标,该指标充分挖掘网络的结构特性,并且考虑邻居节点重要性的反馈,相对于其他一些指标更加准确,并且计算复杂度低。

表 1 各指标优缺点

指标	优点	缺点
度数	简单直观,方便计算	只反映了局部特性,忽略很多因素,不够准确
介数	从流量的角度分析节点的重要性,反映了网络的动态特性	计算复杂度过高
紧密度	从全网拓扑量化重要性,反映了节点的中心化程度	对于部分网络不适合,有一定的局限性
特征向量	考虑到邻居节点的重要性	简单的将各节点进行线性叠加,过于简化实际情况
Pagerank	既考虑到邻居的重要性反馈,又考虑到全网的拓扑特性	忽略了一些实际因素,较理想化

对于任意一个节点,如果将它的 n 条边看作 n 条数据流,那么复杂网络可以看成一个复杂的通信系统模型(如图 4 所示为信息通信模型),因此可以运用信息论的观点来评估节点的重要性,在这种基于互信息的评估方法中,每个节点所包含的信息代表了节点的重要程度,计算出这些节点的信息并进行比较,就可以得到节点之间的相对重要性。

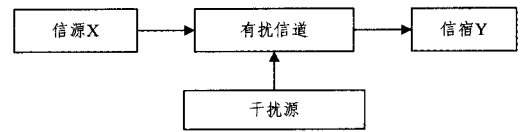


图 4 信息通信模型

3.1 互信息的定义

基于互信息的评估方法通过节点的信息量来评估节点的重要性,每个节点所包含的信息量由它的边决定。节点的信息量是这样计算的:先计算节点之间的互信息,互信息原本是信息论中的一个十分重要的概念,用来表示信息之间的关系,互信息是两个随机变量统计相关性的测度,这里用互信息来描述边的特殊性,然后计算该节点所包含的边的互信息总量。

定义 1 设节点 i 到节点 j 的互信息为 $I(i, j)$, 定义如下:

$$I(i, j) = \begin{cases} \log \frac{1}{p_{ij}} - \log \frac{1}{p_i}, & i \text{ 与 } j \text{ 直接相连} \\ 0, & \text{其他} \end{cases} \quad (3)$$

式中, p_{ij} 为节点 i 的边 (i, j) 的概率。这里我们以无权无向网络为例, 因此同一节点的每条边的概率分布视为等概率分布。所以

$$I(i, j) = \begin{cases} \log k_i - \log k_j, & i \text{ 与 } j \text{ 直接相连} \\ 0, & \text{其他} \end{cases} \quad (4)$$

式中, k_i 为节点 i 的度。

定义 2 节点 i 的信息量是节点 i 与其他节点的互信息之和, 设为 $I(i)$:

$$I(i) = \sum_{j=0}^n I(i, j) \quad (5)$$

由以上定义可知, $I(i, j) + I(j, i) = 0$, 因此对于整个网络来说, 信息量之和 $\sum_{i=0}^n I(i)$ 为零。在计算出所有节点的信息量之后按照降序进行排序, 信息量越大的节点重要性越强。

3.2 指标的计算方法

根据以上定义, 本节提出的计算方法中节点的重要性同时取决于节点之间连接的强弱和多少, 因此这种计算方法相对于其他一些方法更加简单与准确。

设网络为图 $G=(V, E)$, 其中 V 是网络中所有节点的集合, E 是网络中所有边的集合, n 为网络中的节点数, (i, j) 代表节点 i 到节点 j 的边。设矩阵 $A=[a_{ij}]$ 为网络 G 的邻接矩阵, 其中

$$a_{ij} = \begin{cases} 1, & \text{节点 } i \text{ 与节点 } j \text{ 有边相连} \\ 0, & \text{节点 } i \text{ 与节点 } j \text{ 无边相连} \end{cases}$$

指标的计算方法如下:

Step1 输入网络 G 的邻接矩阵 A , 根据邻接矩阵 A 计算各节点的度数;

Step2 针对节点 i , 利用式(4)计算节点 i 与节点 $j(1 \leq j \leq n)$ 的互信息 $I(i, j)$;

Step3 利用式(5)计算节点 i 的信息量;

Step4 如果 $i < n$, 返回 Step2 计算节点 $i+1$ 的信息量;

Step5 根据计算出的各节点的信息量从大到小排序, 得到节点重要性顺序。

从上述算法步骤看, Step2 和 Step3 是算法的主循环, 因此整个节点重要性计算的时间复杂度取决于 Step2 节点间互信息的计算和 Step3 信息量的计算, 得出算法的时间复杂度为 $O(n^2)$ 。下面以星形网络(如图 5 所示)为例计算并说明以上计算过程。

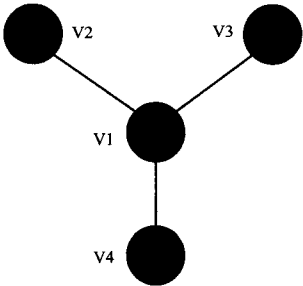


图 5 星形网络拓扑结构图

表 2 是按照以上互信息方法对星形网络进行计算的过程, 表中最右侧所示的信息量是由所在行的 4 个互信息相加得来的。计算结果显示星形网络的中心节点 v_1 的信息量最大, 其它节点的信息量相等且都小于 v_1 , 因此 v_1 最重要, 其它节点次之并且同等重要。可见, 本节提出的方法揭示了网络的拓扑结构特性, 考虑了邻居重要性的反馈, 因此比较准确地反映了节点的重要程度。

表 2 星形网络的计算过程

i \ j	节点间互信息 $I(i, j)$				节点的信息量 $I(i)$
	1	2	3	4	
1	0	1.099	1.099	1.099	3.297
2	-1.099	0	0	0	-1.099
3	-1.099	0	0	0	-1.099
4	-1.099	0	0	0	-1.099

3.3 计算实例与分析

利用 APRA 网络拓扑(如图 6 所示)来分析说明基于互信息的复杂网络节点重要性评估方法。APRA 网络拓扑是北美常用的干线拓扑, 由 21 个节点和 26 条边组成, 网络的平均度数为 2.48。去掉 APRA 网络中的任意一个节点, 网络仍然连通。表 3 是应用本节提出的互信息法与已有的度数法、介数法、生成树法以及 pagerank 算法计算出的 APRA 网络的节点重要程度的对照表。

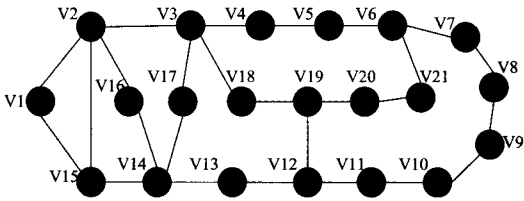


图 6 APRA 网络拓扑结构图

从表 3 可以看出, 本节提出的基于互信息的评估算法得出的最重要节点是 v_{14} , 而其它的几种方法得出的结果是 v_3 或 v_{14} 。从图 6 中可见 v_3 和 v_{14} 两个节点局部的拓扑结构很相似, 它们都是网络中度数最大的节点, 不同方法得出不同结果的原因是不同方法关注的角度不同。基于互信息的评估方法与度的相关性较强, 但是与度数法相比, 本节的方法不仅考虑了局部网络的结构特性, 也考虑了邻居的重要性反馈。另外在度数、介数、生成树指标中, v_7 、 v_8 、 v_9 、 v_{10} 、 v_{11} 这 5 个节点的重要性相同, 而在本节提出的方法中, 节点 v_8 、 v_9 、 v_{10} 相对于 v_7 、 v_{11} 更为重要。观察可以发现: 移除 v_7 节点后, v_{12} 节点的负载加重; 而移除 v_9 节点可能同时造成 v_6 节点和 v_{12} 节点的拥塞, 对网络造成的潜在破坏更大。以上分析表明, 本节提出的互信息方法可以明确地分辨出这些差别。

表 3 APRA 网络节点重要程度对照表

编号	度数	介数	生成树	PageRank	互信息
V1	2	0.000	0.6262	0.036	-1.099
V2	4	0.025	0.9721	0.067	1.674
V3	4	0.030	0.9930	0.070	2.079
V4	2	0.005	0.8387	0.040	-0.693
V5	2	0.005	0.8387	0.041	-0.405
V6	3	0.015	0.9836	0.061	1.216
V7	2	0.005	0.8797	0.043	-0.405
V8	2	0.005	0.8797	0.045	0.000
V9	2	0.005	0.8797	0.045	0.000
V10	2	0.005	0.8797	0.044	0.000
V11	2	0.005	0.8797	0.042	-0.405
V12	3	0.015	0.9780	0.057	0.811
V13	2	0.005	0.8051	0.038	-1.099
V14	4	0.030	0.9864	0.069	2.367
V15	3	0.010	0.8787	0.051	-0.170
V16	2	0.005	0.6639	0.036	-1.386
V17	2	0.005	0.6977	0.037	-1.386
V18	2	0.005	0.7701	0.038	-1.099
V19	3	0.015	0.9671	0.057	0.811
V20	2	0.005	0.8279	0.041	-0.405
V21	2	0.005	0.8279	0.042	-0.405

图 7 是根据各种指标挖掘出的 APRA 中的网络头 5 个

重要节点(用灰色标记)。可见虽然各种方法出发点不一样,关注的角度不同,但是得出的重要节点是类似的。

这种基于互信息的复杂网络节点重要性评价方法在一定程度上克服了传统方法的缺陷,能够更准确地发掘出复杂网络中的重要节点,并且简单易算,显示了该方法的优越性,使得复杂网络节点重要性的评价更加合理。

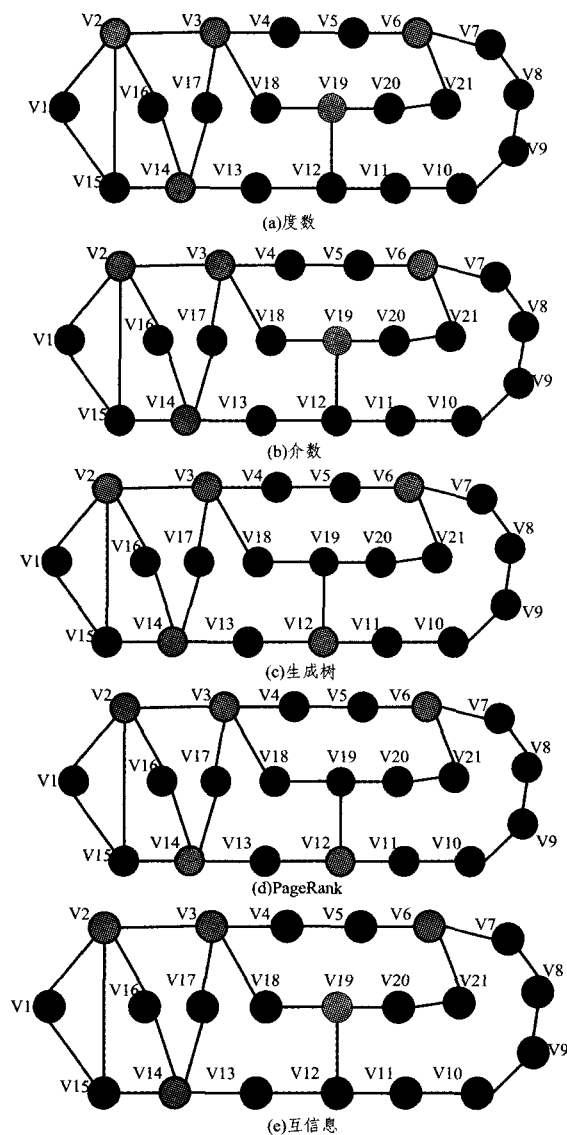


图7 APRA网络中各种评估指标挖掘出的头5个重要节点(用灰色标记)

4 基于数据场的复杂网络聚类算法

在这种基于数据场的复杂网络聚类算法中,复杂网络拓扑中不同节点的影响力有差别,应用节点的重要性来代表其吸引力,重要节点在网络中影响范围较大。利用这种影响力的大小,来划分复杂网络的簇结构;我们把复杂网络看成一个势场,按照实际的需求来决定划分的粒度,将第3节中的方法计算出的前 N 个重要节点作为复杂网络中各个簇的核,通过计算不同核的势来决定将该节点划分到哪一个簇中^[10]。

算法描述如下:

Step1 利用第3节的方法计算网络节点的重要性,根据需求选取前 N 个重要的节点作为各个簇的核;

Step2 对于非核节点 i 按照式(1)计算每个核对于节点

的势,将节点 i 加入势最大的那个核所在的簇中。

这种基于数据场的复杂网络聚类算法是一种启发式的聚类算法,主要通过 Step2 来进行聚类,首先挖掘重要节点,然后将节点划分到不同的簇中,算法的时间复杂度为 $O(n^2)$ 。

5 仿真实验与分析

Zachary 空手道俱乐部网络是 Zachary 通过对一个美国大学空手道俱乐部的长期观测而构建的^[11]。该网络是在复杂网络社团结构分析中经常用到的一个网络,用来测试聚类算法的划分精度。该网络中的节点代表俱乐部成员,边代表成员之间的关系,包含 34 个节点和 78 条边。在观察过程中发现该俱乐部由于在是否提高俱乐部的收费上产生分歧,最终分裂为以俱乐部主管和俱乐部教练为中心的两个子集团(如图8所示),图中节点1和节点34分别是主管和教练,灰色方形和白色圆形节点分别代表分裂后的两个子集团的成员。

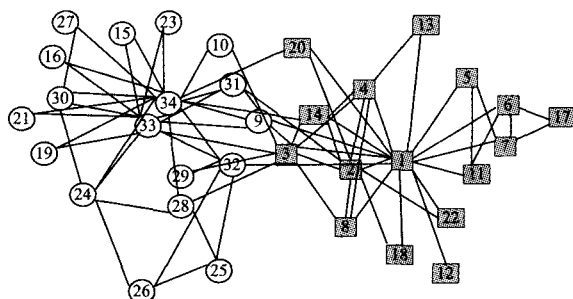


图8 Zachary 空手道俱乐部网络

利用基于数据场的复杂网络聚类算法对其进行分析,首先计算网络中的重要节点,计算出的头5个重要节点是 34、1、33、2、3,对应的互信息指标值分别是 28.5447、23.5251、14.1273、6.04115、5.99615。在该网络中,由于先验条件说明分裂成两个集团,因此选取 34 和 1 两个节点作为簇的核心,然后按照数据场方法划分每个节点所在的簇(取 2),这里计算最短路径用的是 Dijkstra 算法。利用基于数据场的复杂网络聚类算法划分出的 Zachary 空手道俱乐部网络簇结构如图9所示,灰色方形和白色圆形节点分别代表数据场方法划分后的两个子集团的成员。可见绝大部分节点处在正确的子集团中,只有成员 3 被错误地划分到教练的集团中,这可能是由于成员 3 处在两个子集团的边界,与每个子集团都有 5 条边相连,本身就具有一定的歧义性。因此可以得出,基于数据场的复杂网络聚类算法准确地划分了 Zachary 空手道俱乐部网络的簇结构。

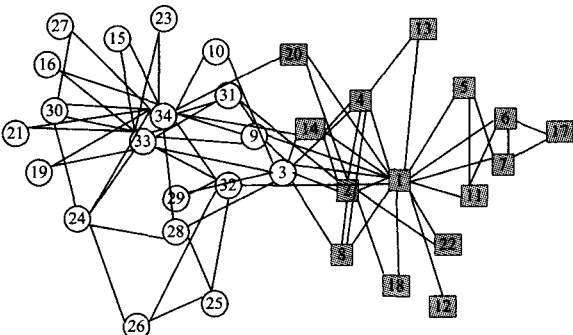


图9 基于数据场的方法划分 Zachary 网络簇结构

(下转第 93 页)

点矩阵能够取得近 100% 的准确率。可以得出,热点矩阵方法全面优于已有的贝叶斯方法。

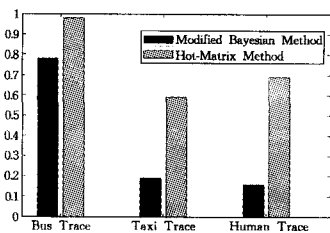


图 5 β 对准确率的影响

结束语 本文主要讨论了异步参考信息对匿名移动路径信息中节点身份的隐私攻击。我们认为攻击者即使只掌握了异步参考信息,也能够比较准确地识别出目标用户在匿名公开数据集中的 ID。我们提出了基于移动位置分布特性的热点矩阵识别算法,并且在 3 个真实的数据集中对其准确率进行了验证。实验结果表明,热点矩阵识别算法在出租车、人类和公交车数据集上分别能够达到 60%、70% 和 100% 的识别准确率,远高于已有的方法。

参考文献

- [1] Lba counts in apple store and android marketplace [OL]. <http://www.skyhookwireless.com/locationapps/>, 2010
- [2] Quercia D, Lathia N, Calabrese F, et al. Recommending social events from mobile phone location data[C]//Proceedings of the 2010 IEEE International Conference on Data Mining (ICDM' 10). Washington, DC, USA, 2010. IEEE Computer Society, 2010:971-976
- [3] Sohn T, Li K A, Lee G, et al. Place-its: A study of location-based reminders on mobile phones[C]//Fifth International Conference on Ubiquitous Computing (UbiComp' 05). Berlin, Heidelberg, 2005. Springer Verlag, 2005:232-250
- [4] Gaonkar S, Li J, Choudhury R R, et al. Micro-blog: sharing and querying content through mobile phones and social participation

[C]//Proceedings of the 6th international conference on Mobile systems, applications, and services (MobiSys' 08). New York, NY, USA, 2008. ACM, 2008:174-186

- [5] Motani M, Srinivasan V, Nuggehalli P S. Peoplenet: engineering a wireless virtual social network[C]//Proceedings of the 11th annual international conference on Mobile computing and networking (MobiCom' 05). New York, NY, USA, 2005. ACM, 2005:243-257
- [6] Narayanan A, Shmatikov V. De-anonymizing social networks[C]//Proceedings of the 2009 30th IEEE Symposium on Security and Privacy (SP' 09). Washington, DC, USA, 2009. IEEE Computer Society, 2009:173-187
- [7] Sala A, Zhao X, Wilson C, et al. Sharing graphs using differentially private graph models[C]//Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference (IMC' 11). New York, NY, USA, 2011. ACM, 2011:81-98
- [8] The foursquare website[OL]. <https://foursquare.com/>, 2012
- [9] Piorkowski M, Sarafijanovic-Djukic N, Grossglauser M. CRAWDAD data set epfl/mobility [OL]. <http://crawdad.cs.dartmouth.edu/epfl/mobility>, 2009-02-24
- [10] Eagle N, Pentland A. Reality mining: sensing complex social systems[J]. Personal and Ubiquitous Computing, 2006, 10(4):255-268
- [11] University S J. Suvnet-trace data[OL]. <http://wirelesslab.sjtu.edu.cn>
- [12] Ma C Y, Yau D K, Yip N K, et al. Privacy vulnerability of published anonymous mobility traces[C]//Proceedings of the sixteenth annual international conference on Mobile computing and networking (MobiCom' 10). New York, NY, USA, 2010. ACM, 2010:185-196
- [13] Gonzalez M C, Hidalgo C A, Barabasi A-L. Understanding individual human mobility patterns[J]. Nature, 2008, 453:779-782
- [14] Eagle N, Pentland A S. CRAWDAD data set mit/reality[OL]. <http://crawdad.cs.dartmouth.edu/mit/reality>, 2005-07-01

(上接第 73 页)

结束语 复杂网络的聚类分析是一个很重要的问题,在很多领域都有广泛的应用。本文提出了一种基于数据场的复杂网络聚类算法,它将物理学的数据场概念引入复杂网络中,通过一种基于互信息的复杂网络节点重要计算方法计算出节点的重要性,然后根据节点的势来划分网络的簇结构。本文提出的聚类算法有效地对复杂网络的簇结构进行划分,并且可以根据实际的需求决定划分的粒度。基于数据场的复杂网络聚类算法虽然在计算精确度和计算复杂度上具有一定的优势,但是过分依赖事先给定的先验条件。下一步应该研究能根据网络的实际情况智能地决定划分条件、不需要人为地给出参数并且有待进一步推广到大型复杂网络中的聚类算法。

参考文献

- [1] Brandes U, Delling D, et al. On modularity clustering[J]. IEEE Transaction on Knowledge and Data Engineering, 2008, 20(2):172-188
- [2] 杨博,刘大有,等. 复杂网络聚类方法[J]. 软件学报, 2009, 20(1):54-66

- [3] 孙吉贵,刘杰,赵连宇. 聚类算法研究[J]. 软件学报, 2008, 19(1):48-61
- [4] 淦文燕,李德毅,王建民. 一种基于数据场的层次聚类方法[J]. 电子学报, 2006, 34(2):258-262
- [5] 戴晓军,淦文燕,李德毅. 基于数据场的图像数据挖掘研究[J]. 计算机工程与应用, 2004(26)
- [6] 陈勇,胡爱群,胡啸. 通信网中节点重要性的评价方法[J]. 通信学报, 2004, 25(8)
- [7] Page L, Brin S. The PageRank Citation Ranking: Bringing Order to the Web[C]//Stanford Digital Libraries Working Paper, 1998
- [8] Wasserman S, Faust K. Social network analysis: methods and applications [M]. Cambridge: Cambridge University Press, 1994:218
- [9] 傅祖芸. 信息论-基础理论与应用[M]. 北京:电子工业出版社, 2001
- [10] 胡钢锋,李德毅,陈桂生,等. 一种网络化数据挖掘方法研究[J]. 微电子学与计算机, 2006, 23(9):126-128
- [11] Zachary W W. An information flow model for conflict and fission in small groups[J]. Journal of Anthropological Research, 1977, 33(4):452-473