

金融网络频繁链路发现算法

吕芳, 汤丰赫, 黄俊恒, 王佰玲

(哈尔滨工业大学(威海)计算机科学与技术学院, 山东 威海 264209)

摘要: 随着各种非法金融活动的泛滥, 从金融网络中发现犯罪线索的分析研究越来越引起学者的重视。对银行账户交易数据的特点进行了详细分析, 建立了银行账户交易网络通用模型。在此基础上, 为解决金融实体之间关系强度的评估问题, 提出了双向活跃边搜索计算方法。为了还原犯罪组织的资金流动方式, 提出了深度可控的广度优先频繁链路发现方法。在真实银行数据上的实验证明, 上述方法能有效解决同伙预测和资金追踪问题。

关键词: 双向活跃边; 频繁链路; 同伙预测; 资金追踪

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2019050

Frequent path discovery algorithm for financial network

LYU Fang, TANG Fenghe, HUANG Junheng, WANG Bailing

School of Computer Science and Technology, Harbin Institute of Technology(weihai), Weihai 264209, China

Abstract: With the proliferation of various illegal financial activities, more and more attention is paid to the research of finding criminal cues in financial network by scholars. The characteristics of the transaction data generated by bank accounts are analyzed in detail, and a general model of bank account transaction network is established. On this basis, a two-direction active edge searching method is proposed to solve the problem of evaluating the relationship strength between financial entities. And then, a breadth-first frequent path discovery algorithm with depth controlled is presented,

收稿日期: 2018-11-28; 修回日期: 2019-02-10

通信作者: 王佰玲, wbl@hit.edu.cn

基金项目: 国家重点研发计划重点专项基金资助项目 (No.2018YFB2004201, No.2017YFB0801804); 前沿科技创新专项基金资助项目 (No.2016QY05X1002-2); 国家区域创新中心科技专项基金资助项目 (No.2017QYCX14); 山东省重点研发计划基金资助项目 (No.2017CXGC0706); 中央高校基本科研业务费专项资金资助项目 (No.HIT.NSRIF.2020098); 2017 威海市大学共建基金资助项目

Foundation Items: The National Key Research and Development Program of China (No.2018YFB2004201, No.2017YFB0801804), Frontier Science and Technology in Notation of China (No.2016QY05X1002-2), National Regional Innovation Center Science and Technology Special Project of China (No.2017QYCX14), Key Research and Development Program of Shandong Province (No.2017CXGC0706), The Fundamental Research Funds for the Central Universities (No.HIT.NSRIF.2020098), 2017 University Co-construction Project in Weihai City

论文引用格式: 吕芳, 汤丰赫, 黄俊恒, 等. 金融网络频繁链路发现算法[J]. 网络与信息安全学报, 2019, 5(5): 48-55.

LYU F, TANG F H, HUANG J H, et al. Frequent path discovery algorithm for financial network[J]. Chinese Journal of Network and Information Security, 2019, 5(5): 48-55.

with which the way how the financial flows is restored. Experiment results on the real bank data show that the above two methods are effective in solving the problem of peer prediction and financial tracking respectively.

Key words: two-direction active edge, frequent path, peer prediction, financial tracking

1 引言

多年来, 非法传销、非法集资、洗钱和诈骗等金融犯罪组织或犯罪活动屡禁不止。这些组织或活动在进行资金吸纳和资金运作时, 离不开银行账户间的资金交易。随着我国社会经济的飞速发展, 银行账户的开户数量以及账户间的交易数目和交易金额大幅增长。交易的非现金支付占整个支付系统比例超过 70%。非现金支付和 DT (data technology) 时代的到来, 使金融数据呈爆炸式增长。美国波士顿咨询公司 2015 年发布的《互联网金融生态系统 2020 系列报告之大数据篇》中指出, 银行业每创收 1×10^6 美元, 平均产生 820 GB 的数据。金融犯罪组织的资金运转线索隐藏在这庞大的银行数据中。从银行海量交易中挖掘账户实体的关系强度、挖掘非法资金的交易链路, 对打击经济犯罪活动有直接的指导意义。

目前, 由于银行数据的保密性质, 针对资金交易数据的分析研究还处于初级阶段, 有关研究主要体现在下几个方面。

1) 在欺诈检测方面: Yu 等^[1]将随机游走算法应用于欺诈检测中, 以路径是否经过认证节点为依据, 判定节点是否可疑。Tran 等^[2]在文献[1]的基础上, 提出一种在随机游走算法上结合广度优先搜索的改进系统。不同于上述从正常节点出发预测可疑节点的策略, Yang 等^[3]提出一种从已知可疑节点出发搜索可疑节点的算法; 刘泉等^[4]提出一种利用概率图检测可疑节点的方法。此外, 结合规则库的设计, 丁濛濛^[5]提出一个基于规则引擎的反欺诈模型, 研究了规则匹配过程的优化方法。

2) 在反洗钱研究方面: 张成虎等^[6]基于 AI 技术设计了一种反洗钱系统。喻炜等^[7]基于交易

网络特征向量中心度量, 提出了一种可疑洗钱行为检测系统。孙景等^[8]提出利用复杂网络理论研究反洗钱的思路和方法。刘丽芳等^[9]利用拓扑机构分析工具, 分析了洗钱关联账户之间的资金流转关系。

3) 在非法传销研究方面: Wang 等^[10]提出了利用决策树理论识别可疑客户的框架。Liu 等^[11]采用线性判别和中心图发现技术, 建立了传销网络核心人物和同伙判定模型。李艳丽等^[12]从用户社交行为数据中识别传销网络模型, 分别建立了正常、传销等不同性质用户的“自我中心网络”, 进而根据网络结构特性分析了传销用户的行为特征。

以上这些研究均针对特定的金融犯罪活动, 根据犯罪活动的不同模式和特点, 设计启发式、机器学习等方法检测异常个体或异常组织。

在非法传销、非法集资、洗钱和虚开发票等需要多人协作完成犯罪的非法金融活动中, 普遍存在初始资金通过多个账户中转最终到达汇集账户的现象。如何快速、准确地挖掘出进行上述非法资金转移的关系账户和交易链路, 对打击非法金融组织有直接的指导意义。

本文通过分析银行账户交易数据特点, 首先构建了金融交易数据通用网络模型; 其次, 在对银行账户交易数据特点分析的基础上, 分别提出了评估账户实体关系强度的“双向活跃边”搜索计算方法和还原资金流向的深度可控广度优先“频繁链路”发现算法。在真实银行数据上的实验结果证明, “双向活跃边”算法能有效预测传销同伙, “频繁链路”发现算法能有效追踪传销资金的去向。

2 相关工作

2.1 银行账户交易网络构建

如果把每一个银行账户表示为网络中的节

点,账户间的交易关系表示为节点之间的有向边,两个节点之间交易的时间、金额、次数等信息表示为有向边的权重,则银行账户之间的交易构成一个有向加权金融交易网络。根据金融交易网络的特点,本文定义银行账户交易网络数学模型如下。

定义 1 一个银行账户交易网络可表示为 $G = (V, E, W)$, 其中, $V = \{v_1, v_2, \dots, v_n\}$ 为账户集合, $E = \{e_{ij}\}$, $E \subset V \times V$, $1 \leq i \leq n$, $1 \leq j \leq n$ 为边集合, v_i 与 v_j 之间存在有向边当且仅当 v_i 对 v_j 至有一笔交易,该有向边记作 $e_{ij} = \langle v_i, v_j \rangle$ 。 $W = \{w_{ij}\}$ 为有向边权重集合, $W \in (R^+)^{Nt}$, $w_{ij} = (r_{ij}^{(1)}, r_{ij}^{(2)}, \dots, r_{ij}^{(Nt)})$, 其中, $r_{ij}^{(i)}$ 和 Nt 分别表示节点 v_i 与 v_j 间第 i 次交易的金额和总交易次数。

2.2 双向活跃边搜索计算

在银行账户交易网络 G 中,节点之间的有向边信息不但反映了账户之间的交易情况,还隐含了用户实体间的关系强度。节点对之间双向交易的频繁程度反映了账户之间的关系强度和关系类型。两用户在犯罪组织中的角色不同,导致账户在相对关系强度和有向关系类型方面均有显著差异。因此,节点对之间双向交易的频繁程度可用于发现可疑交易节点、同伙和组织。基于此,本文提出了“双向活跃边”概念和双向活跃边搜索计算方法。双向活跃边的定义如下。

定义 2 设 $G = (V, E, W)$ 是符合定义 1 的一个银行账户交易网络, $\forall e_{ij} \in E \wedge e_{ji} \in E$, 如果 $\|w_{ij}\| \geq \alpha \wedge \|w_{ji}\| \geq \alpha$, 则称 (v_i, v_j) 为双向活跃边, 其中, $\|w_{ij}\|$ 是权重向量 w_{ij} 的长度, α 是阈值。

根据定义 2, 本文提出了一种双向活跃边搜索计算方法。

1) 双向活跃边搜索计算方法 1

在定义 1 和定义 2 的基础上,本节设计了一种双向活跃边搜索计算方法,输入金融网络 $G = (V, E, W)$, 集合 E_{ac} 记录双向活跃边,遍历图 G 的有向边集合 $\langle v_i, v_j \rangle \in E \wedge \langle v_j, v_i \rangle \in E$, 对于给定

的阈值 α , 判断 $\|w_{ij}\| \geq \alpha \wedge \|w_{ji}\| \geq \alpha$ 是否成立,若成立,则 $E_{ac} = E_{ac} \cup \{(v_i, v_j, w_{ij}, w_{ji})\}$ 。详见算法 1。

算法 1 双向活跃边搜索计算方法 1

输入 银行账户交易网络 $G = (V, E, W)$, 阈值 α //参见定义 1

输出 双向活跃边集合 E_{ac} //记录双向活跃边

- 1) $E_{ac} = \emptyset$
- 2) for $\langle v_i, v_j \rangle \in E \wedge \langle v_j, v_i \rangle \in E$ do //遍历边
- 3) if $\|w_{ij}\| \geq \alpha \wedge \|w_{ji}\| \geq \alpha$
- 4) $E_{ac} = E_{ac} \cup \{(v_i, v_j, w_{ij}, w_{ji})\}$
- 5) end if
- 6) end for

算法 1 的时间复杂性为 $O(e^2)$, 为了提高算

法的效率,本文利用金融字典的方法对算法 1 进行改进,提出了双向活跃边搜索计算方法 2。

2) 双向活跃边搜索计算方法 2

算法 1 需要遍历图 G 中的每条有向边 $\langle v_i, v_j \rangle$, 再判断是否存在 $\langle v_j, v_i \rangle \in E$, 这个过程的时间复杂性是 $O(e^2)$, 由于 G 中有向边集合 E 数量庞大,时间耗费太大。本文通过构建金融字典,提出了双向活跃边搜索计算方法 2,有效降低了运算时间。

算法 2 利用分治的思想,把每个节点的关联有向边进行分类,即将 $\forall v_i \in V$ 的关联边集合记为 A_i , 字典项记为 $d_i = v_i: A_i$ 。给定 $G = (V, E, W)$, 节点集合 $V = \{v_1, v_2, \dots, v_n\}$ 的关联有向边集合可表示为 $A = \{A_1, A_2, \dots, A_n\}$, 金融字典可表示为 $D = \{d_1, d_2, \dots, d_n\}$ 。通过金融字典 D 的构建,把 $\langle v_j, v_i \rangle \in E$ 的搜索范围限定在集合 A_j 中。“双向活跃边”搜索计算方法 2 的算法设计如算法 2 所示。

算法 2 双向活跃边搜索计算方法 2

输入 银行账户交易网络 $G = (V, E, W)$, 阈值 α //见定义 1

输出 双向活跃边集合 E_{ac} //记录双向活跃边

- 1) 初始化: $E_{ac} = \emptyset$

```

2) for  $\forall v_i \in V$  do
3)  $A_i = \text{get}(v_i) // d_i = v_i; A_i \in D$ 
4) for  $j \in A_i$  do
5) if  $\|w_{ij}\| \geq \alpha \wedge \|w_{ji}\| \geq \alpha$ 
6)  $E_{ac} = E_{ac} \cup \{(v_i, v_j, w_{ij}, w_{ji})\}$  //记录双向
活跃边信息
7) end if
8) end for
9) end for

```

算法 2 将时间复杂性降为 $O\left(\frac{e^2}{n}\right)$ 。可见，构建金融字典有效提高了算法的速度。

3 频繁链路发现

在非法传销、非法集资、洗钱和诈骗等涉众型非法金融活动中，资金的流向、流通方式均受违法者操控，即初始资金往往通过多个账户周转到最终获利账户，而且中转账户在特定场景下是固定的。可见，在一段时间内，被操控的非法资金会多次由固定的账户顺次进行流通。本文把这种多次发生的交易账户路径称作频繁链路。

定义 3 给定银行账户交易网络 $G=(V, E, W)$ ，其中 $E \subset V \times V$ 为有向边集合，对任意的节点序列 $L=(v_1, v_2, \dots, v_l)$ ，如果 L 中两个相邻节点 v_i, v_j 满足 $\langle v_i, v_j \rangle \in E$ 且 $\|w_{ij}\| > \beta$ ，则称 L 为一条频繁链路。其中， l 表示链路深度， β 为单向活跃边阈值。

根据定义 3，本文首先提出一种深度可控的广度优先频繁链路发现算法，该算法首先以任意节点 v 为起点，利用广度优先搜索方式查找与 v 的有向交易次数大于阈值 β 的节点集合 S_v ；然后以 S_v 中的任意节点 r 为起点重复该过程，当节点 r 相对 v 的搜索深度大于阈值 l 时停止在这条路径上的搜索；最后输出节点 v 在 G 中的广度优先生成树 T_v 。其中， T_v 以 v 为根节点，从 v 到每个叶节点的路径都是一条频繁链路。详见算法 3。

算法 3 深度可控的广度优先频繁链路发现算法

输入 银行账户交易网络 $G=(V, E, W)$ ，资金链路长度 l ，初始种子节点 v ，阈值 β

输出 以 v 为根节点的广度优先频繁链路生成树 T_v

```

1) 设置辅助队列  $Q$ ，深度控制变量  $len = 0$ ， $T_v = \emptyset$  //条件初始化
2)  $Q.add(v, len)$  //队列初始化
3) while ( $!IsEmpty(Q)$ ) do
4)  $pop(v', len)$  //出队列
5) if ( $len < l$ )
6) for all  $u$  in  $V$  do
7) if  $(v', u) \in E \wedge \|w_{v'u}\| > \beta$  //单向“活跃边”
8)  $T_v = T_v \cup (v', u); Q.add(u, ++len)$  //递归构建  $T_v$ 
9) end if
10) end for
11) end if
12) end while

```

对于任意节点 v ，算法 3 生成了以节点 v 为根节点的频繁链路生成树 T_v ，其中从 v 到每个叶节点的路径都为频繁链路。算法 3 的时间复杂度为 $O(E^{len})$ ，其中， E 为边的条数， len 为链路搜索的深度。

4 实验和结果分析

4.1 实验环境

Inter(R) Core i7-7700HQ CPU@ 2.80 GHz，内存(RAM) 16 GB。软件环境为：Python 语言，Windows 7 操作系统。

4.2 数据集

实验数据来自某经侦部门经过脱敏处理、包含某大型线下传销组织的长期资金交易的银行账户交易数据，包含 15 685 个交易账户和 227 231 条

交易信息,其中传销账户为1305个,可疑账户对之间的交易记录为18549条。

4.3 算法分析

账号对之间的直接交易反映了账号的亲密关系,多个账号之间的频繁资金流动反映了组织的资金流动模式。本节首先说明了线下传销组织的资金交易特点,然后分析了资金交易网络中双向活跃边的存在情况,进而验证了双向活跃边搜索计算算法在预测传销同伙应用中的准确性。针对频繁链路挖掘算法,本文首先分析了传销组织交易频繁环路的存在情况,然后给出了该算法用于资金追踪的有效性。

1) 线下传销组织资金交易特点

线下传销组织的成员关系结构呈金字塔型,自顶向下的等级明确且不可逾越,且上下级之间是一对多的所属关系。上级以拉人头的形式进行会员扩张,下级以缴纳会员费的形式加入组织,该过程中上级的收益方式为依据自身层级等因素获得相应比例的返利提成。资金频繁地从多个下级账户汇集到塔尖的最大获利账户。因此,传销组织在资金流动上也呈现出一定的定向、环路现象。

典型的资金流动方式如图1所示。传销资金由低层节点逐层自顶向上汇集到账户 T 中,由资金汇集方向组成的网络呈现树形结构,如 T_1 、 T_2 所示。组织中有 A 、 B 两个控制资金交易的主要账户,会员费在 T_1 、 T_2 、 A 、 T 之间是单向交易的,且越靠近树根路径上的交易越频繁。账户 B 是返利资金的分发账户,其定期从账户 T 收到组织成员所有的返利资金,然后分发给树形结构中非叶子节点。返利金在 T 、 B 、 T_1 、 T_2 之间也是单向交易,且呈星形网状结构。需要注意的是,传销组织的真实资金流动关系复杂多样, A 、 B 、 T 三个或任意两个角色可能由同一个账户负责。但是,稳定的线下传销组织以上述方式完成传销

照等级、会员数量获得返利金。因此,传销资金交易网络结构中,存在双向活跃边、频繁链路、环形结构。另外,线下传销多以好友、亲戚、同事等形式发展会员,因此,会员账户之间也存在因非传销活动产生的资金往来。这一现象提升了网络中双向活跃边存在的可能性。

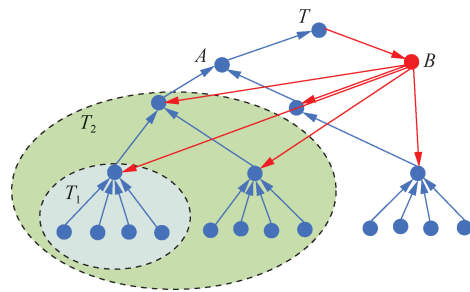


图1 线下传销资金流动方式

2) 双向活跃边实验分析

本文对真实资金交易网络中“双向活跃边”的存在情况进行测试。阈值 α 是判断存在有向交易的节点对是否为“活跃”的依据。如图2所示,随着 α 值的增长,双向活跃边的数量逐渐减少。当 $\alpha \leq 2$ 时,双向活跃边的数量为1442,表示网络中存在双向交易的节点,分别至少产生过两次交易。当 $2 < \alpha < 7$ 时,双向活跃边的数量以较快的速度下降, $\alpha = 7$ 时,双向活跃边的数量小于200;当 $\alpha > 7$ 时,双向活跃边的数量下降较为平缓。当 $\alpha = 11$ 时,双向活跃边的数量小于100, $\alpha = 35$ 时,双向活跃边的数量小于10。由此可见,资金交易网络中的“双向活跃边”能反映现实世界中用户之间的亲密关系。

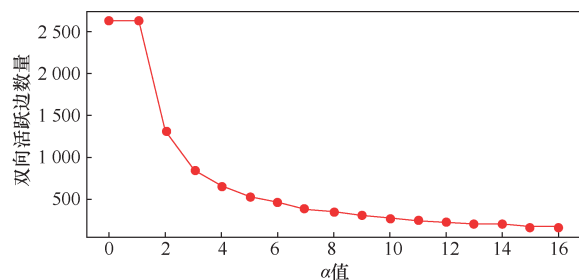


图2 双向活跃边数量随阈值 α 的变化

3) 利用双向活跃边预测传销同伙

从传销账户样本集合中随机选择 m 个账户，组成传销账户集合 $S_1 = \{v_1, v_2, \dots, v_m\}$ ，与集合 S_1 中账户双向活跃边大于 α 的关联账户集合 $S_2^a \subseteq S_1$ 。显然，随着阈值 α 值的增加， S_2^a 中的账户与集合 S_1 中账户的关系越加密切，当 α 大于一定值时，有理由怀疑 S_2^a 中的账户为 S_1 中账户的传销同伙。

实验中，本文采用传销样本集合为 1 305 个可疑账户，随机选择 $m = 100$ 个传销账户集合 S_1 ，搜索计算得到这些账户的双向活跃边账户集合 S_2^a 。在双向活跃边账户集合 S_2^a 中既可能存在可疑账户节点也可能存在正常交易账户节点，图 3 为进行 50 次随机采样得到 α 值与预测准确率的变化关系。随着阈值 α 的增加， S_2^a 账户集合中传销账户的预测准确率逐渐增加。当 $\alpha \geq 12$ 时， S_2^a 中账户为传销账户的预测准确率达到 70%，故可以推断出可疑账户之间的密切度随着 α 值增大而更紧密， S_2^a 中账户的正常节点极有可能为疑节点。

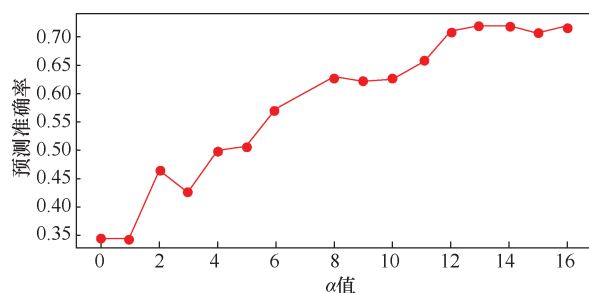


图3 α 值对传销账户预测准确率的影响

4) 传销网络中频繁环路分析

由图 1 可知，传销组织之间的资金交易关系存在稳定的环结构。去掉资金交易发生的时间顺序，将交易网络视为静态网络，则网络中由可疑账户产生的最大长度为 8 的交易链路共有 233 764 835 条。对真实网络中存在于可疑账户之间的环结构的大小及相应数量分析如表 1 所示。

由表 1 可知，交易网络中可疑账户之间存在

的 2 节点环结构数量为 210 个，可见，传销组织中双向的存在比例较低，且由实验 3) 可知，当双向边的活跃度提高到一定阈值时，2 节点环结构对传销组织的覆盖率可达到 70%。随着环结构的增大，网络中环结构的数量急剧增加，且长度为 7 时达到最大。可见，可疑账户存在长度为 7 的环结构的概率极高。因此，验证了图 1 中资金交易的网络关系。

表 1 环结构存在情况分析

大小	数量/个
2	210
3	530
4	1 679
5	6 319
6	35 918
7	115 069

5) 频繁链路挖掘算法分析

数据集中的账户和交易关系组成一个资金交易网络，以网络中任意节点为起点，采用深度可控的广度优先频繁路径算法挖掘网络中的频繁链路。本节分别对频繁链路定义 3 中的单向活跃边阈值 β 和链路长度 l 进行测试，分析资金交易网络中的资金流动情况。

图 4 为 $l = \{3, 4, 5\}$ ，即路径长度分别表示为 threenodelinks、fournodelinks、fivenodelinks 时，资金交易网络中频繁链路数随 β 的变化情况。

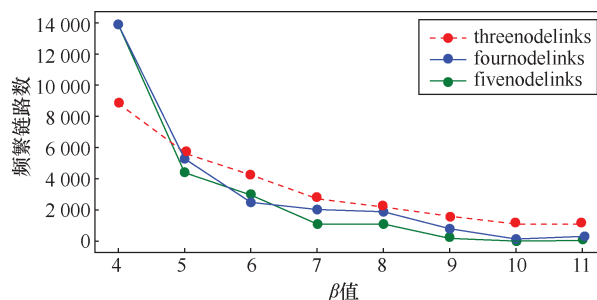


图4 频繁链路数随阈值 β 值变化

如图 4 所示, 横坐标 β 的取值区间为 $[4, 11]$, 其中, 曲线“threenodelinks”上的数值 (4, 13 925) 表示当 β 取值为 4 或者 5 时, 网络存在 13 925 条频繁链路。可见, 随着 β 取值的增加, 网络中的频繁链路越来越少, 且当 $\beta = 5$ 时, 网络中的 4 频繁链路与 5 频繁链路数量趋于一致。可见, 当 β 大于一定值之后, 4 频繁链路集合能代表网络中稳定存在的频繁链路。

此外, 实验还分析了频繁链路传销组织资金交易链路的重合情况, 如图 5 所示。

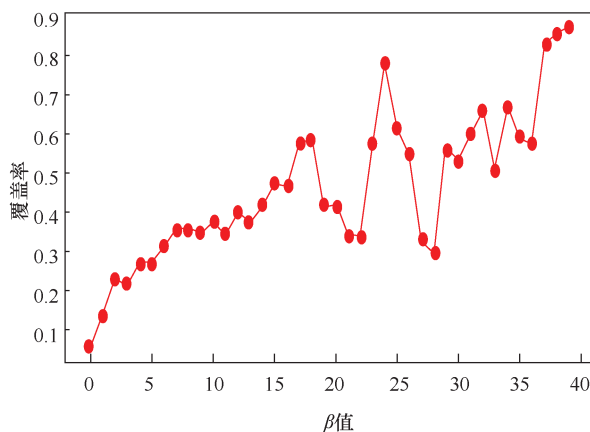


图 5 频繁链路对传销节点的覆盖率

由图 5 可知, 随着阈值 β 的增大, 频繁链路对传销节点的覆盖率呈现增加的趋势, 当 $\beta > 40$ 时, 4 频繁链路传销网络的重合度达到 90% 以上。这充分说明, 可以通过频繁链路的挖掘还原传销组织的资金流向。

5 结束语

针对金融网络实体关系强度计算问题, 本文提出了双向活跃边的概念和搜索计算方法, 该算法在传销同伙的预测方面取得了良好的效果。在实际应用中发现, 两个账号除具有直接关系外, 还具有其他间接关系, 为提高预测准确率, 在下一步的研究中将加入对间接关系的计算。针对涉众型金融犯罪的资金追踪问题, 本文提出了频繁

链路的概念和一种深度可控的广度优先频繁链路挖掘算法, 在真实的传销网络的资金追踪应用中取得了很好的效果。

以上两种方法在面对海量数据量时容易遇到性能瓶颈。近年来, 遗传、蚁群等仿生算法在很多领域取得了很好的效果, 接下来将开展利用仿生算法解决金融问题的分析研究。

参考文献:

- [1] YU H, KAMINSKY M, Gibbons P B, et al. SybilGuard: defending against sybil attacks via social networks[J]. IEEE/ACM Transactions on Networking, 2008, 16(3):576-589.
- [2] TRAN N, LI J, SUBRAMANIAN L, et al. Optimal Sybil-resilient node admission control[C]// IEEE Infocom. 2015.
- [3] YANG C, HARKREADER R, ZHANG J, et al. Analyzing spammers' social networks for fun and profit: a case study of cyber criminal ecosystem on twitter[C]//International Conference on World Wide Web. 2012.
- [4] 刘泉, 王晓国. 基于概率图的银行电信诈骗检测方法[J]. 计算机科学, 2018, 45(7):122-128.
LIU X, WANG X G. Probabilistic graphical model based approach for bank telecommunication fraud detection[J]. Computer Science, 2018, 45(7): 122-128.
- [5] 丁濛濛. 基于规则引擎的互联网金融反欺诈研究[J]. 电脑知识与技术, 2018, 14(1):1-3.
DING M M. Internet finance anti-fraud research based on rule engine[J]. Computer Knowledge and Technology, 2018, 14(1):1-3.
- [6] 张成虎, 李时. 基于 AI 技术的反洗钱系统设计[J]. 中国金融电脑, 2005, (3):44-47.
ZHANG C H, LI S. Design of anti-money laundering system based on AI technology[J]. Financial Computer of China, 2005, (3): 44-47.
- [7] 喻炜, 王建东. 基于交易网络特征向量中心度的可疑洗钱识别系统[J]. 计算机应用, 2009, 29(9):2581-2585.
YU Y, WANG J D. Suspicious money laundering detection system based on eigenvector centrality measure of transaction network[J]. Journal of Computer Applications, 2009, 29(9): 2581-2585.
- [8] 孙景, 陈婧, 万红. 基于复杂网络的可疑金融交易识别研究[J]. 数字技术与应用, 2013, 4(149):206-207.
SUN J, CHEN J, WAN H. Research on suspicious financial transaction identification based on complex network[J]. Digital Technology and Application, 2013, 4(149):206-207.
- [9] 刘丽芳, 陶文立, 陈延妙. 拓扑工具在反洗钱关联账户资金流分

析中的运用[J]. 福建金融, 2013, 2:39-44.

LIU L F, TAO W L, CHEN Y M. Application of topology tools in analysis of fund flow of anti-money laundering related accounts[J]. Fujian Finance, 2013, 2:39-44.

- [10] WANG S N, YANG J G. A money laundering risk evaluation method based on decision tree[C]//International Conference on Machine Learning and Cybernetics. 2007.

- [11] LIU Y. Based on social network crime organization relation mining and central figure determining[C]//IEEE 3rd International Conference on Software Engineering and Service Science. Beijing, China, 2012.

- [12] 李艳丽, 刘阳, 谢文波, 等. 大数据发现非法传销网络[J]. 大数据, 2017, 3(5):106-112.

LI Y L, LIU Y, XIE W B, et al. Detecting illegal pyramid scheme network in big data[J]. Big Data, 2017, 3(5):106-112.

[作者简介]



吕芳(1990-), 女, 山东阳谷人, 哈尔滨工业大学(威海)博士生, 主要研究方向为复杂网络、信息内容安全、数据挖掘。



汤丰赫(1998-), 男, 满族, 内蒙古呼和浩特人, 主要研究方向为信息内容安全。



黄俊恒(1966-), 男, 河南新乡人, 哈尔滨工业大学(威海)副教授, 主要研究方向为数据挖掘、人工智能。



王佰玲(1978-), 男, 黑龙江哈尔滨人, 哈尔滨工业大学教授、博士生导师, 主要研究方向为信息对抗、信息安全、信息搜索、移动网络、金融安全。