



(12)发明专利申请

(10)申请公布号 CN 107832964 A

(43)申请公布日 2018.03.23

(21)申请号 201711165335.9

(22)申请日 2017.11.21

(71)申请人 江苏神威云数据科技有限公司

地址 214000 江苏省无锡市滨湖区蠡园开发区吟白路1号研创大厦1902室

(72)发明人 尤福宝 徐文渊 黄云辉 汤成辉

(74)专利代理机构 无锡市汇诚永信专利代理事务所(普通合伙) 32260

代理人 张欢勇

(51)Int.Cl.

G06Q 10/06(2012.01)

G06Q 40/02(2012.01)

G06Q 50/00(2012.01)

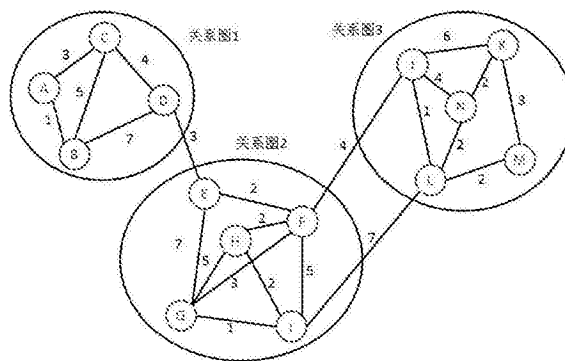
权利要求书3页 说明书7页 附图2页

(54)发明名称

银行客户关系圈分析方法及系统

(57)摘要

本发明涉及金融服务技术领域,尤其涉及一种银行客户关系圈分析方法及系统,该分析方法的具体步骤为:对客户在本行的资金流水信息进行收集整理,形成资金流水数据;将数据中的一个客户用一个节点表示,客户与客户之间因交易形成的关系用一条无向边表示,形成客户资金关系网络,基于Modularity增益的社区划分理论,获得划分好的关系圈;在每个关系圈中获得客户的重要度指标;在获得每个关系圈中所有客户的重要度指标后,选出该关系圈中重要度最高的客户作为中心客户,根据其资金流向,标记其上下游客户,形成客户链,研究分析客户链中客户的资金及资金关系情况,即可制定出客户营销计划及风险防控计划。本方法具有分析参考指标多,分析结果准确度高的优点。



1. 一种银行客户关系圈分析方法,具体步骤包括:

A. 对客户在本行的资金流水信息进行收集整理,合并该客户本行不同账户下的所有资金流水信息,获得一个客户只有一个收支主体、一笔转账必有一个交易对手的资金流水数据;

B. 对资金流水数据进行统计分布处理,消除极大值、极小值;

C. 将一个客户用一个节点表示,客户与客户之间因交易形成的关系用一条无向边表示,客户与客户之间的关系紧密程度用一个与转账金额、转账频率、最近一次转账距今时间相关的权重表示,形成客户资金关系网络;

D. 基于Modularity增益的社区划分理论,首先将每个节点初始化为单一节点的关系圈,每个节点根据模块度增益决定该节点是否与相邻节点所属的关系圈合并以及和哪个相邻节点所属的节点合并,然后重新构建网络,将上一步的每一个关系圈合并为一个节点,重复进行以上步骤直到整个网络的模块度不再增加或者增加幅度低于某一指定阈值为止,最终获得划分好的关系圈;

E. 在每个关系圈中获得客户的重要度指标,客户的重要度根据其转账行为进行评判,先将每个关系圈中的每个客户的转账关系还原为有向有权转账关系,再根据PageRank理论,建立基于客户转账行为的客户重要度计算方法,计算出每个客户的重要度指标;

F. 在获得每个关系圈中所有客户的重要度指标后,选出该关系圈中重要度最高的客户作为中心客户,根据其资金流向,标记其上下游客户;

G. 将步骤F获得的中心客户及其上下游客户作为客户链,研究分析客户链中每个客户的资金情况及各个客户之间的资金关系情况,根据研究分析结果对关系圈中的客户制定营销计划及风险防控计划。

2. 根据权利要求1所述的银行客户关系圈分析方法,其特征在于:所述步骤A中资金流水信息不包括客户本行不同账户之间和跨行同名账户之间的资金流水信息、交易对手信息缺失的资金流水信息,客户向税务部门、社保部门、法院等公共事业部门缴费的资金流水信息,客户与会计师事务所、律师事务所、信托公司等服务支持类客户的资金流水信息、客户向银行缴纳手续费、利息收支等资金流水信息。

3. 根据权利要求1所述的银行客户关系圈分析方法,其特征在于:所述步骤B中消除极大值的方法为:用样本均值加上3倍标准差获得的值替换所有极大值,消除极小值的方法为:用样本均值减去3倍标准差获得的值替换所有极小值。

4. 根据权利要求1所述的银行客户关系圈分析方法,其特征在于:所述步骤C中权重的计算方法为:

C1. 对转账金额、转账频率和最近一次转账距今时间均采用Z-Score标准化法进行归一化处理,使得变量数据落在 $[0, 1]$ 区间之内;

C2. 归一化处理后的转账金额、转账频率、最近一次转账距今时间按照4:4:2的比例计算客户关系权重值。

5. 根据权利要求1所述的银行客户关系圈分析方法,其特征在于:所述步骤D中关系圈模块度的计算基于以下公式:

$$Q = \frac{1}{2m} \sum_{i,j} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j) \quad \text{式 (1)}$$

式(1)中, i, j 表示关系圈编号, A_{ij} 表示边权重, 根据两个节点间转账金额、转账频率、最近一次转账距今时间加权计算并做分箱处理获得1-10级评分, k_i, k_j 分别表示关系圈*i*、关系圈*j*的出度, 即关系圈与外部节点相连的边数, $\delta(c_i, c_j)$ 表示关系圈*i*、关系圈*j*是否处于同一关系圈的标志变量, 如果关系圈*i*、关系圈*j*属于同一关系圈, 则该标志变量取值为1, 否则取值为0。

6. 根据权利要求1所述的银行客户关系圈分析方法, 其特征在于: 所述步骤E中重要度的计算方法基于以下公式:

$$PR(p_i) = \frac{1 - \alpha}{N} + \alpha \sum_{p_j \in M(p_i)} \frac{PR(p_j)}{L(p_j)} \quad \text{式 (2)}$$

式(2)中 $i = 1, 2, \dots, n$, p_1, p_2, \dots, p_n 表示*N*个节点, $M(p_i)$ 表示转入*p_i*节点的客户加权集合, $L(p_j)$ 表示转出*p_j*节点的客户加权集合, N 表示关系圈中的总客户数, α 表示预设的随机转出参数, PR 值的计算, 采用幂迭代法、特征值法或者代数法求解。

7. 基于权利要求1所述的银行客户关系圈分析方法的银行客户关系圈分析系统, 其特征在于: 包括数据采集单元、数据整理单元、数据加工单元、数据分析处理单元、目标数据寻找单元、数据展示单元和数据导出单元;

所述数据采集单元用于采集客户在本行的资金流水信息;

所述数据整理单元对采集到的客户在本行的资金流水信息进行整理, 合并该客户本行不同账户下的所有资金流水信息, 获得一个客户只有一个收支主体、一笔转账必有一个交易对手的资金流水数据;

所述数据加工单元用于通过资金流水数据建立客户资金关系网络;

所述数据分析处理单元用于对客户资金关系网络进行分析, 获得划分完成的关系圈;

所述目标数据寻找单元用于在每个关系圈中以转账行为为重要度评判标准, 寻找出重要度最高的客户, 并以该客户作为中心客户, 根据其资金流向, 标记其上下游客户;

所述数据展示单元用于展示中心客户及其上下游客户形成的客户链, 所述客户链包含客户及客户的资金流向信息;

所述数据导出单元用于将数据展示单元展示的数据导出到现有的银行管理系统中。

8. 根据权利要求7所述的银行客户关系圈分析系统, 其特征在于: 所述数据整理单元还用于对资金流水数据进行统计分布处理, 消除极大值、极小值。

9. 根据权利要求7所述的银行客户关系圈分析系统, 其特征在于: 所述数据加工单元具体用于将一个客户用一个节点表示, 客户与客户之间因交易形成的关系用一条无向边表示, 客户与客户之间的关系紧密程度用一个与转账金额、转账频率、最近一次转账距今时间相关的权重表示, 形成客户资金关系网络。

10. 根据权利要求8所述的银行客户关系圈分析系统, 其特征在于: 所述数据分析处理单元具体用于基于Modularity增益的社区划分理论, 首先将每个节点初始化为单一节点的关系圈, 每个节点根据模块度增益决定该节点是否与相邻节点所属的关系圈合并以及与哪个相邻节点所属的节点合并, 然后重新构建网络, 将上一步的每一个关系圈合并为一个节

点,重复进行以上步骤直到整个网络的模块度不再增加或者增加幅度低于某一指定阈值为止,最终获得划分好的关系圈。

银行客户关系圈分析方法及系统

技术领域

[0001] 本发明涉及金融服务技术领域,尤其涉及一种银行客户关系圈分析方法及系统。

背景技术

[0002] 随着利率市场化改革的全面实现及直接融资的加速发展,使得国内银行同业竞争日趋激烈,再加上国内经济增速放缓,互联网金融企业等非银行金融机构以及非金融机构的参与,使得银行获得稳定的客户群体更加困难。对银行来说,如果想要获得稳定的客户群体,如果能够通过对客户转账数据来分析客户在产业链中的地位及其与上下游客户的关联性,全产业链留存客户与资金等数据,就能通过这些数据获得客户的稳定性、成长性等指标,还可以获得客户上下游关联客户的情况,不仅能够帮助银行规避客户资金风险,还可以定向发展上下游关联客户,拓展银行客户群,但目前的银行金融系统不具备这些功能。

[0003] 针对这一问题,目前银行业从业者开发了一些客户转账关系分析系统,例如中国专利公开号为CN105913235A的发明专利申请就公开了一种客户转账关系分析方法及系统,其具体方法为:(1)采集客户在银行的转账交易数据;(2)对所述转账交易数据进行数据加工,生成客户的交易向量,所述交易向量包括:客户号、交易权重和交易对手号;(3)以发生转账交易的客户为节点,客户之间的交易权重为边,生成转账关系网络;(4)并行计算转账关系网络中每个节点的度;(5)对节点的度进行排序,输出节点的度大于预设阈值的节点;(6)为转账关系网络中每个节点赋予一个第一无二的初始标签,表示迭代开始前,每个节点自身为一个独立的社区;(7)对于转账关系网络中的每一个节点,遍历节点的邻居节点信息,选取节点中次数最多的标签,将其设定为该节点的新标签;(8)重复上述步骤,直至所有节点的标签不再改变或者达到预先设定的迭代次数;(9)将具有相同标签的几点合并到同一社区。其中在步骤(2)的具体方法为:对所述转账交易数据进行加载、清洗和转换,生成客户的交易向量;步骤(3)的具体方法为:用无向图的节点表示转账关系网络中的成员,用无向图的边表示成员与成员件的关系,建了客户转账关系的邻接矩阵,有转账关系的节点将其对应的位置为转账次数,无转账关系矩阵对应的位置置为0。

[0004] 虽然上述方法能够实现客户转账关系的分析,但其仍然存在一些缺陷,例如:在社区划分的方法上,只选取邻居节点出现次数最多的节点进行合并,达到迭代次数后终止,这就意味着在评价社区内客户关系的时候,仅考虑了转账次数这一影响因子,并不对转账时间、转账金额等影响因子进行分析;在社区划分上,仅从节点出发去考虑,如果迭代次数不加以限制,会造成有的社区规模会非常大,有的社区规模则非常小,导致样本整体信息量损失较大。因此,亟待开发一种新的银行客户关系圈分析方法及系统,以改善现有分析方法中存在的问题。

发明内容

[0005] 针对现有技术中的问题,本发明提供一种银行客户关系圈分析方法及系统。

[0006] 为实现以上技术目的,本发明的技术方案是:

[0007] 一种银行客户关系圈分析方法,具体步骤包括:

[0008] A.对客户在本行的资金流水信息进行收集整理,合并该客户本行不同账户下的所有资金流水信息,获得一个客户只有一个收支主体、一笔转账必有一个交易对手的资金流水数据;

[0009] B.对资金流水数据进行统计分布处理,消除极大值、极小值;

[0010] C.将一个客户用一个节点表示,客户与客户之间因交易形成的关系用一条无向边表示,客户与客户之间的关系紧密程度用一个与转账金额、转账频率、最近一次转账距今时间相关的权重表示,形成客户资金关系网络;

[0011] D.基于Modularity增益的社区划分理论,首先将每个节点初始化为单一节点的关系圈,每个节点根据模块度增益决定该节点是否与相邻节点所属的关系圈合并以及与哪个相邻节点所属的节点合并,然后重新构建网络,将上一步的每一个关系圈合并为一个节点,重复进行以上步骤直到整个网络的模块度不再增加或者增加幅度低于某一指定阈值为止,最终获得划分好的关系圈;

[0012] E.在每个关系圈中获得客户的重要度指标,客户的重要度根据其转账行为进行评判,先将每个关系圈中的每个客户的转账关系还原为有向有权转账关系,再根据PageRank理论,建立基于客户转账行为的客户重要度计算方法,计算出每个客户的重要度指标;

[0013] F.在获得每个关系圈中所有客户的重要度指标后,选出该关系圈中重要度最高的客户作为中心客户,根据其资金流向,标记其上下游客户;

[0014] G.将步骤F获得的中心客户及其上下游客户作为客户链,研究分析客户链中每个客户的资金情况及各个客户之间的资金关系情况,根据研究分析结果对关系圈中的客户制定营销计划及风险控制计划。

[0015] 作为优选,所述步骤A中资金流水信息不包括客户本行不同账户之间和跨行同名账户之间的资金流水信息、交易对手信息缺失的资金流水信息,客户向税务部门、社保部门、法院等公共事业部门缴费的资金流水信息,客户与会计师事务所、律师事务所、信托公司等服务支持类客户的资金流水信息、客户向银行缴纳手续费、利息收支等资金流水信息。

[0016] 作为优选,所述步骤B中消除极大值的方法为:用样本均值加上3倍标准差获得的值替换所有极大值,消除极小值的方法为:用样本均值减去3倍标准差获得的值替换所有极小值。

[0017] 作为优选,所述步骤C中权重的计算方法为:

[0018] C1.对转账金额、转账频率和最近一次转账距今时间均采用Z-Score标准化法进行归一化处理,使得变量数据落在[0,1]区间之内;

[0019] C2.归一化处理后的转账金额、转账频率、最近一次转账距今时间按照4:4:2的比例计算客户关系权重值。

[0020] 作为优选,所述步骤D中关系圈模块度的计算基于以下公式:

$$[0021] \quad Q = \frac{1}{2m} \sum_{i,j} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j) \quad \text{式(1)}$$

[0022] 式(1)中, i, j 表示关系圈编号, A_{ij} 表示边权重,根据两个节点间转账金额、转账频率、最近一次转账距今时间加权计算并做分箱处理获得1-10级评分, k_i, k_j 分别表示关系圈*i*、关系圈*j*的出度,即关系圈与外部节点相连的边数, $\delta(c_i, c_j)$ 表示关系圈*i*、关系圈*j*是否处

于同一关系圈的标志变量,如果关系圈*i*、关系圈*j*属于同一关系圈,则该标志变量取值为1,否则取值为0。

[0023] 作为优选,所述步骤E中重要度的计算方法基于以下公式:

$$[0024] \quad PR(p_i) = \frac{1 - \alpha}{N} + \alpha \sum_{p_j \in M(p_i)} \frac{PR(p_j)}{L(p_j)} \quad \text{式 (2)}$$

[0025] 式(2)中*i*=1,2,⋯,n,*p*₁,*p*₂,⋯,*p*_n表示*N*个节点,*M*(*p*_{*i*})表示转入*p*_{*i*}节点的客户加权集合,*L*(*p*_{*j*})表示转出*p*_{*j*}节点的客户加权集合,*N*表示关系圈中的总客户数,*α*表示预设的随机转出参数,*PR*值的计算,采用幂迭代法、特征值法或者代数法求解。

[0026] 基于上述的银行客户关系圈分析方法,建立一种银行客户关系圈分析系统,包括数据采集单元、数据整理单元、数据加工单元、数据分析处理单元、目标数据寻找单元、数据展示单元和数据导出单元;

[0027] 所述数据采集单元用于采集客户在本行的资金流水信息;

[0028] 所述数据整理单元对采集到的客户在本行的资金流水信息进行整理,合并该客户本行不同账户下的所有资金流水信息,获得一个客户只有一个收支主体、一笔转账必有一个交易对手的资金流水数据;

[0029] 所述数据加工单元用于通过资金流水数据建立客户资金关系网络;

[0030] 所述数据分析处理单元用于对客户资金关系网络进行分析,获得划分完成的关系圈;

[0031] 所述目标数据寻找单元用于在每个关系圈中以转账行为为重要度评判标准,寻找出重要度最高的客户,并以该客户作为中心客户,根据其资金流向,标记其上下游客户;

[0032] 所述数据展示单元用于展示中心客户及其上下游客户形成的客户链,所述客户链包含客户及客户的资金流向信息;

[0033] 所述数据导出单元用于将数据展示单元展示的数据导出到现有的银行管理系统中。

[0034] 作为改进,所述数据整理单元还用于对资金流水数据进行统计分布处理,消除极大值、极小值。

[0035] 作为优选,所述数据加工单元具体用于将一个客户用一个节点表示,客户与客户之间因交易形成的关系用一条无向边表示,客户与客户之间的关系紧密程度用一个与转账金额、转账频率、最近一次转账距今时间相关的权重表示,形成客户资金关系网络。

[0036] 作为优选,所述数据分析处理单元具体用于基于Modularity增益的社区划分理论,首先将每个节点初始化为单一节点的关系圈,每个节点根据模块度增益决定该节点是否与相邻节点所属的关系圈合并以及与哪个相邻节点所属的节点合并,然后重新构建网络,将上一步的每一个关系圈合并为一个节点,重复进行以上步骤直到整个网络的模块度不再增加或者增加幅度低于某一指定阈值为止,最终获得划分好的关系圈。

[0037] 从以上描述可以看出,本发明具备以下优点:本发明基于银行客户的资金流水信息,针对客户构建关系圈、关系链方法,对客户进行分析处理,获得客户之间的关系紧密度、单个客户的重要度等指标,再根据获得的指标对客户进行定向业务营销和客户资金风险防控,大大提高了银行对现有客户的风险防控有效性及对潜在客户、业务往来相关客户的业务开展有效性,与现有的客户关系圈分析方法相比,本发明具有分析参考指标多,分析过程

样本整体信息量损失小、分析结果准确度高的优点。

附图说明

- [0038] 图1是本发明实施例1的节点分析示意图；
 [0039] 图2是本发明实施例1关系圈的分析示意图；
 [0040] 图3是本发明实施例1中心客户的分析示意图；
 [0041] 图4是本发明实施例2的系统框图。

具体实施方式

[0042] 结合图1至图3,详细说明本发明的实施例1,但不对本发明的权利要求做任何限定。

[0043] 如图1所示,一种银行客户关系圈分析方法,具体步骤包括:

[0044] A.对客户在本行的资金流水信息进行收集整理,合并该客户本行不同账户下的所有资金流水信息,获得一个客户只有一个收支主体、一笔转账必有一个交易对手的资金流水数据;其中资金流水信息可以剔除客户本行不同账户之间和跨行同名账户之间的资金流水信息、交易对手信息缺失的资金流水信息、客户向税务部门、社保部门、法院等公共事业部门缴费的资金流水信息、客户与会计事务所、律师事务所、信托公司等服务支持类客户的资金流水信息、客户向银行缴纳手续费、利息收支等资金流水信息等分析价值较低的资金流水信息,以提高资金流水信息的分析有效性。

[0045] B.对资金流水数据进行统计分布处理,消除极大值、极小值,消除极大值的具体方法为:用样本均值加上3倍标准差获得的值替换所有极大值,消除极小值的具体方法为:用样本均值减去3倍标准差获得的值替换所有极小值。

[0046] C.将一个客户用一个节点表示,客户与客户之间因交易形成的关系用一条无向边表示,客户与客户之间的关系紧密程度用一个与转账金额、转账频率、最近一次转账距今时间相关的权重表示,形成客户资金关系网络;其中权重的具体计算方法为:先对转账金额、转账频率和最近一次转账距今时间均采用Z-Score标准化法进行归一化处理,使得变量数据落在[0,1]区间之内;再对归一化处理后的转账金额、转账频率、最近一次转账距今时间按照4:4:2的比例计算客户关系权重值,即转账金额、转账频率、最近一次转账距今时间三个要素值分别乘以0.4,0.4,0.2,得到一个[0,3]的加总值,然后根据加总值做从小到大排序,按照等深分箱打分0-9(即1-10级评分),从而得到权重值。

[0047] D.基于Modularity增益的社区划分理论,首先将每个节点初始化为单一节点的关系圈,每个节点根据模块度增益决定该节点是否与相邻节点所属的关系圈合并以及与哪个相邻节点所属的节点合并;然后重新构建网络,将上一步的每一个关系圈合并为一个节点,重复进行以上步骤直到整个网络的模块度不再增加或者增加幅度低于某一指定阈值为止,最终获得如图2所示的划分好的关系圈;

[0048] 其中,关系圈模块度的计算基于以下公式:

$$Q = \frac{1}{2m} \sum_{i,j} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j) \quad \text{式(1)}$$

[0050] 式(1)中,i,j表示关系圈编号, A_{ij} 表示边权重,根据两个节点间转账金额、转账频

率、最近一次转账距今时间加权计算并做分箱处理获得1-10级评分, (加权计算并做分箱处理获得1-10级评分的方法, 可以选用步骤C中的权重值计算方法) k_i 、 k_j 分别表示关系圈i、关系圈j的出度, 即关系圈与外部节点相连的边数, $\delta(c_i, c_j)$ 表示关系圈i、关系圈j是否处于同一关系圈的标志变量, 如果关系圈i、关系圈j属于同一关系圈, 则该标志变量取值为1, 否则取值为0。

[0051] E. 在每个关系圈中获得客户的重要度指标, 客户的重要度根据其转账行为进行评判, 先将每个关系圈中的每个客户的转账关系还原为有向有权转账关系 (以图2中所示的关系圈2为例, 关系圈2还原后的有向有权转账关系如图3所示), 其中权重值仍然以步骤C获得的权重值为准, 再根据PageRank理论, 建立基于客户转账行为的客户重要度计算方法, 计算出每个客户的重要度指标;

[0052] 其中, 重要度的计算方法基于以下公式:

$$[0053] \quad PR(p_i) = \frac{1 - \alpha}{N} + \alpha \sum_{p_j \in M(p_i)} \frac{PR(p_j)}{L(p_j)} \quad \text{式 (2)}$$

[0054] 式(2)中 $i=1, 2, \dots, n$, p_1, p_2, \dots, p_n 表示N个节点, $M(p_i)$ 表示转入 p_i 节点的客户加权集合, $L(p_j)$ 表示转出 p_j 节点的客户加权集合, N 表示关系圈中的总客户数, α 表示预设的随机转出参数, PR值的计算, 采用幂迭代法、特征值法或者代数法求解。以图3为例, 采用以上方法通过计算机程序计算得到的结果下表所示。

[0055]

客户	重要度分值
G	0.296
E	0.114
H	0.114
F	0.313
I	0.162

[0056] F. 在获得每个关系圈中所有客户的重要度指标后, 选出该关系圈中重要度最高的客户作为中心客户, 根据其资金流向, 标记其上下游客户。以步骤E中获得的数据表为例, 如果以H客户为中心, 进行标记, 可以获得如下表所示的上下游客户分类, 下表中上下游客户标记根据实际资金流向获得。

[0057]

客户	上中下游
G	上游
E	上游
H	上游
F	中心
I	下游

[0058] 以现实中的企业进行举例说明, 如房地产企业, 上游客户主要是建筑材料企业、装修公司等, 下游客户主要是房地产销售代理公司、广告公司等, 上游客户和下游客户基本上不会出现中心转账的情况; 而如汽车生产厂, 其上游是零配件制造商, 下游主要是4S店、汽车经销商等, 会存在上游厂商直接向下游4S店直接转账的情况, 所以, 在判断客户与中心客

户的关系时,需要在资金关系基础上结合行业知识进行具体分析。由于图1至图3的关系圈划分过程中,并未给出关系圈中各个客户的性质,所以上表中的分析结果仅作为示例参考,并无实际意义。

[0059] G.将步骤F获得的中心客户及其上下游客户作为客户链,研究分析客户链中每个客户的资金情况及各个客户之间的资金关系情况,根据研究分析结果即可对关系圈中的客户制定营销计划及风险控制计划,具体包括:

[0060] (1)根据中心客户在本行的资金情况,向其上下游客户进行银行业务营销,将单个客户拓展到产业链客户群;

[0061] (2)根据上下游客户经营情况,验证、预测中心客户经营情况,防控中心客户资金风险,确定客户稳定性和成长性。

[0062] 从以上描述可以看出,本发明具备以下优点:本发明基于银行客户的资金流水信息,针对客户构建关系圈、关系链方法,对客户进行分析处理,获得客户之间的关系紧密度、单个客户的重要度等指标,再根据获得的指标对客户进行定向业务营销和客户资金风险控制,大大提高了银行对现有客户的风险控制有效性及对潜在客户、业务往来相关客户的业务开展有效性,与现有的客户关系圈分析方法相比,本发明具有分析参考指标多,分析过程样本整体信息量损失小、分析结果准确度高的优点。

[0063] 结合图4,详细说明本发明的实施例2,但不对本发明的权利要求做任何限定。

[0064] 基于实施例1所述的银行客户关系圈分析方法,建立一种如图4所示的银行客户关系圈分析系统,包括数据采集单元、数据整理单元、数据加工单元、数据分析处理单元、目标数据寻找单元、数据展示单元和数据导出单元;

[0065] 数据采集单元用于采集客户在本行的资金流水信息,最好不采集客户本行不同账户之间和跨行同名账户之间的资金流水信息、交易对手信息缺失的资金流水信息、客户向税务部门、社保部门、法院等公共事业部门缴费的资金流水信息、客户与会计事务所、律师事务所、信托公司等服务支持类客户的资金流水信息、客户向银行缴纳手续费、利息收支等资金流水信息等分析价值较低的资金流水信息,以提高资金流水信息的分析有效性;

[0066] 数据整理单元对采集到的客户在本行的资金流水信息进行整理,合并该客户本行不同账户下的所有资金流水信息,获得一个客户只有一个收支主体、一笔转账必有一个交易对手的资金流水数据,整理过程中最好进行统计分布处理,消除极大值、极小值;

[0067] 数据加工单元用于通过资金流水数据建立客户资金关系网络,具体为:将一个客户用一个节点表示,客户与客户之间因交易形成的关系用一条无向边表示,客户与客户之间的关系紧密程度用一个与转账金额、转账频率、最近一次转账距今时间相关的权重表示,形成客户资金关系网络;

[0068] 数据分析处理单元用于对客户资金关系网络进行分析,获得划分完成的关系圈,具体为:基于Modularity增益的社区划分理论,首先将每个节点初始化为单一节点的关系圈,每个节点根据模块度增益决定该节点是否与相邻节点所属的关系圈合并以及与哪个相邻节点所属的节点合并,然后重新构建网络,将上一步的每一个关系圈合并为一个节点,重复进行以上步骤直到整个网络的模块度不再增加或者增加幅度低于某一指定阈值为止,最终获得划分好的关系圈;

[0069] 目标数据寻找单元用于在每个关系圈中以转账行为为重要度评判标准,寻找出重

要度最高的客户,并以该客户作为中心客户,根据其资金流向,标记其上下游客户;

[0070] 数据展示单元用于展示中心客户及其上下游客户形成的客户链,所述客户链包含客户及客户的资金流向信息;

[0071] 数据导出单元用于将数据展示单元展示的数据导出到现有的银行管理系统中。

[0072] 综上所述,本发明具有以下优点:本发明基于银行客户的资金流水信息,针对客户构建关系圈、关系链方法,对客户进行分析处理,获得客户之间的关系紧密度、单个客户的重要度等指标,再根据获得的指标对客户进行定向业务营销和客户资金风险防控,大大提高了银行对现有客户的风险防控有效性及对潜在客户、业务往来相关客户的业务开展有效性,与现有的客户关系圈分析方法相比,本发明具有分析参考指标多,分析过程样本整体信息量损失小、分析结果准确度高的优点。

[0073] 可以理解的是,以上关于本发明的具体描述,仅用于说明本发明而并非受限于本发明实施例所描述的技术方案。本领域的普通技术人员应当理解,仍然可以对本发明进行修改或等同替换,以达到相同的技术效果;只要满足使用需要,都在本发明的保护范围之内。

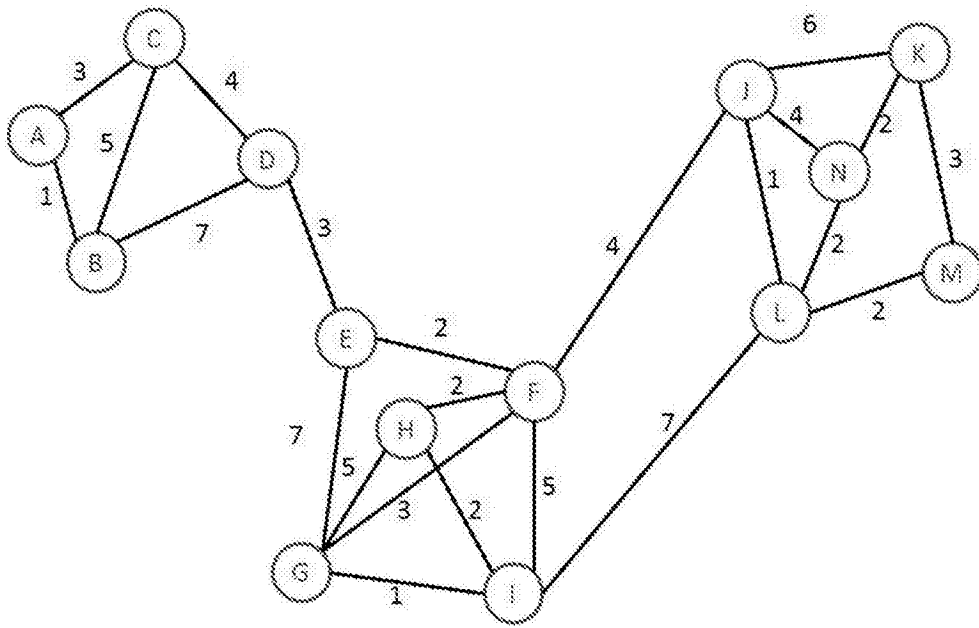


图1

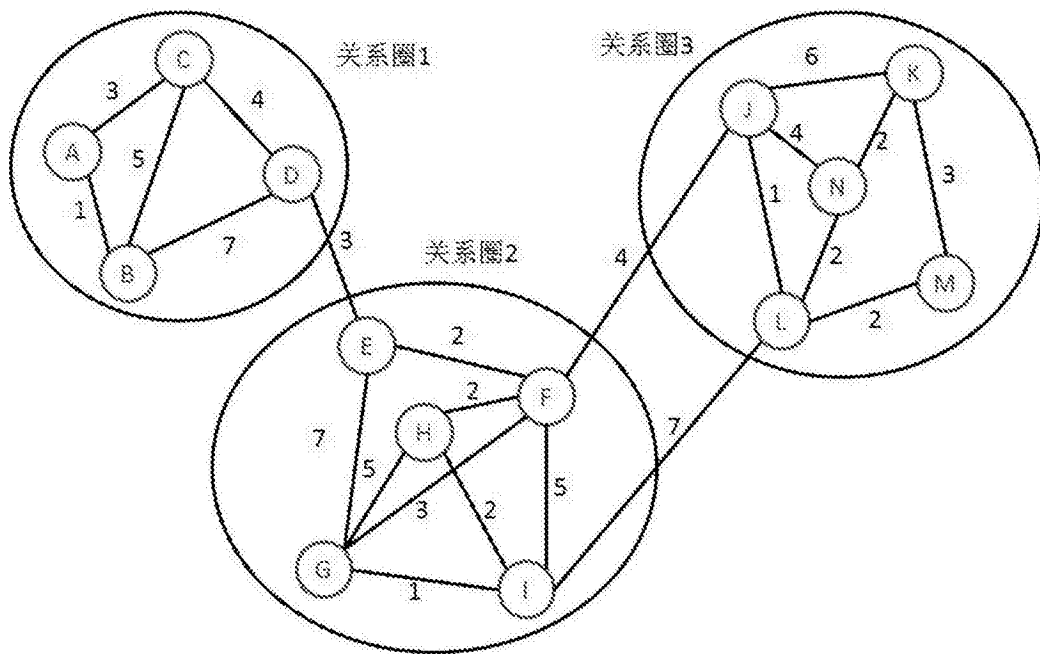


图2

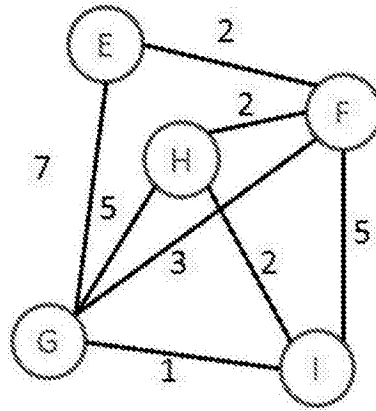


图3

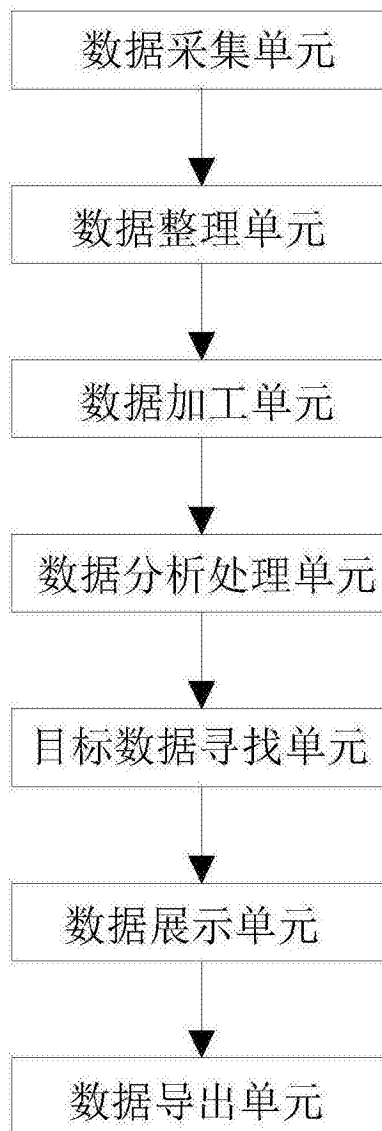


图4