



(12) 发明专利

(10) 授权公告号 CN 112825177 B

(45) 授权公告日 2024.03.29

(21) 申请号 201911142901.3

(22) 申请日 2019.11.20

(65) 同一申请的已公布的文献号

申请公布号 CN 112825177 A

(43) 申请公布日 2021.05.21

(73) 专利权人 北京宸瑞科技股份有限公司

地址 100190 北京市海淀区中关村北二条

13号46幢1层102

(72) 发明人 孙二林 王丽薇 王凯

(74) 专利代理机构 北京康思博达知识产权代理

事务所(普通合伙) 11426

专利代理师 范国锋 刘冬梅

(51) Int. Cl.

G06Q 30/0201 (2023.01)

G06Q 40/04 (2012.01)

(56) 对比文件

CN 105913235 A, 2016.08.31

CN 107480852 A, 2017.12.15

CN 107832964 A, 2018.03.23

CN 110400220 A, 2019.11.01

JP 2015088037 A, 2015.05.07

审查员 黄恒琪

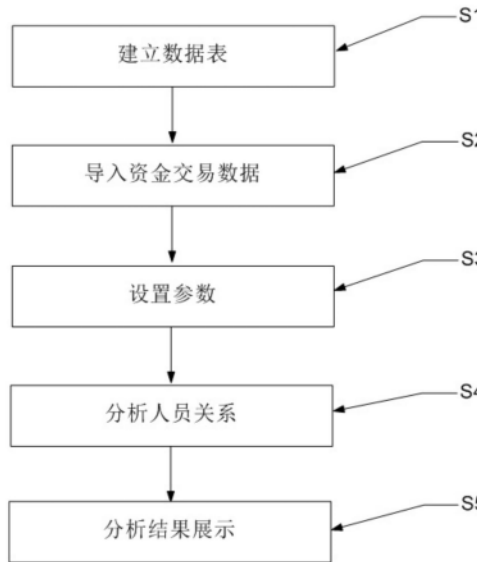
权利要求书2页 说明书16页 附图2页

(54) 发明名称

基于资金交易数据分析人员关系的方法及系统

(57) 摘要

本发明提供了一种基于资金交易数据分析人员关系的方法和系统,包括建立数据表、导入资金交易数据、设置参数、分析人员关系和分析结果展示。所述基于资金交易数据分析人员关系的方法和系统具有能够快速处理海量数据、综合分析资金交易数据中人员之间亲密度、确定人员之间的团体关系和上下游关系等优点。



1.一种基于资金交易数据分析人员关系的方法,其特征在于,综合多个维度对人员之间的关系进行分析,包括以下步骤:

- S1、建立数据表;
- S2、导入资金交易数据;
- S3、设置参数;
- S4、分析人员关系;
- S5、分析结果展示;

在步骤S1中,所述建立数据表,是在数据库中建立表,所述表包括资金交易表、人员表、关系表和参数表;

在步骤S2中,所述导入资金交易数据,是指将银行、反洗钱中心的相关资金交易数据导入到步骤S1建立的数据表中;

在步骤S3中,所述设置参数,为设置或修改参数表中的参数,包括设置资金交易最小金额、净值比总值的阈值、共同交易对手最小个数、同乡亲密度、同姓亲密度、同乡且同姓亲密度、共同交易对手亲密度、成员亲密度和上下游亲密度参数;

所述净值比总值的阈值通过如下方法得到:

S31、预估取值范围;

S32、真实样本验证,选取真实资金交易数据,针对某一组已查明人员关系的样本,将预估值逐个代入真实资金交易数据中,确定代入结果与已查明的人员关系结果吻合度最高的预估值,将此预估值做为净值比总值的阈值;

在步骤S4中,所述分析人员关系,包括如下子步骤:

- S41、提取人员信息;
- S42、确定资金比值;
- S43、判断人员关系;
- S44、确定共同交易对手个数;
- S45、判断是否同乡同姓;
- S46、确定亲密度;

步骤S41中所述提取人员信息,包括从资金交易表中提取姓名、身份证号和银行账号信息,所述姓名、身份证号和银行账号信息从资金交易表的付款方账号、付款方名称、付款方证件号码、收款方账号、收款方名称和收款方证件号码中提取;

在步骤S42中,所述确定资金比值包括确定两人之间的资金交易总值、资金交易净值和净值与总值的比值;

在步骤S43中,所述判断人员关系,为判断关系表中每条数据的主体和客体之间的关系,若关系表数据中的净值比总值的绝对值小于参数表中的净值比总值的阈值,则认为关系表中的主体ID和客体ID是成员关系,若关系表数据中的净值比总值的绝对值大于参数表中的净值比总值的阈值,则认为主体ID和客体ID是上下游关系;

在步骤S44中,所述确定共同交易对手个数为确定两人之间的共同交易对手个数,通过将人员表中的所有人员进行两两组合,进而得出共同交易对手个数;

在步骤S45中,所述判断是否同乡同姓为判断关系表中的主体和客体之间是否存在同乡、同姓关系;

在步骤S46中,所述亲密度为能够反应主体和客体之间关系紧密程度的数值,所述亲密度由多个能够反应主体和客体之间关系的维度构成,所述维度包括乡亲亲密度、共同交易对手亲密度和资金亲密度。

2.根据权利要求1所述的分析人员关系的方法,其特征在于,

所述银行、反洗钱中心的相关资金交易数据为多个Excel文件,每个Excel文件中具有多个表格,将所有表格中的每条数据都导入到资金交易表中。

3.根据权利要求1所述的分析人员关系的方法,其特征在于,

所述步骤S44包括以下子步骤:

S441、统计既与人员A保持“成员:成员”关系,也与人员B保持“成员:成员”关系的人员个数;

S442、统计既与人员A保持“上游:下游”或“下游:上游”关系,也与人员B保持“上游:下游”或“下游:上游”关系,且人员A和人员B均位于“上游”一端的人员个数;

S443、统计既与人员A保持“上游:下游”或“下游:上游”关系,也与人员B保持“上游:下游”或“下游:上游”关系,且人员A和人员B均位于“下游”的人员个数;

S444、将步骤S441~S443中得到的人员个数加和,得到人员A和人员B的共同交易对手个数,根据两人的共同交易对手个数,确定两人是否同一团体。

基于资金交易数据分析人员关系的方法及系统

技术领域

[0001] 本发明涉及一种分析人员关系的方法,特别涉及一种基于资金交易数据分析人员关系的方法,属于情报分析领域。

背景技术

[0002] 随着经济的日益发展,资金交易数据快速增长,互联网、金融、公安、安全、海关缉私等行业越来越多地需要从资金交易数据中分析人员之间的关系。

[0003] 从资金交易数据中分析人员关系是一项新事物,目前业务人员的分析方法是基于Excel等电子表格进行手工分析,分析的速度慢、效率低,只能分析一些简单的、表面的关系。而且手工分析的方法因人而异,缺乏标准规范,不适于分析海量数据。

[0004] 因此,亟待设计一种能够快速、高效,并且能够深层次分析资金交易与人员之间关系的方法。

发明内容

[0005] 为了解决上述问题,本发明人进行了锐意研究,通过资金比例、人员关系、共同交易对手个数、是否同乡同姓、亲密度等多个维度对人员进行分析,从而完成了本发明。

[0006] 本发明的目的在于提供以下方面:

[0007] 一方面,本发明提供了一种基于资金交易数据分析人员关系的方法,其特征在于,包括以下步骤:

[0008] S1、建立数据表;

[0009] S2、导入资金交易数据;

[0010] S3、设置参数;

[0011] S4、分析人员关系;

[0012] S5、分析结果展示。

[0013] 在步骤1中,所述建立数据表,是在数据库中建立表,包括资金交易表、人员表、关系表、参数表。

[0014] 在步骤S2中,所述导入资金交易数据,是指将银行、反洗钱中心等单位的相关资金交易数据导入到步骤S1建立的数据表中,

[0015] 所述银行、反洗钱中心等单位的相关资金交易数据为多个Excel文件,每个Excel文件中具有多个表格,将所有表格中的每条数据都导入到资金交易表中。

[0016] 在步骤S3中,所述设置参数,为设置或修改参数表中的参数,包括设置资金交易最小金额、净值比总值的阈值、共同交易对手最小个数、同乡亲密度、同姓亲密度、同乡且同姓亲密度、共同交易对手亲密度、成员亲密度、上下游亲密度中的一种或多种参数。

[0017] 在步骤S4中,所述分析人员关系,包括如下子步骤:

[0018] S41、提取人员信息;

[0019] S42、确定资金比值;

- [0020] S43、判断人员关系；
- [0021] S44、确定共同交易对手个数；
- [0022] S45、判断是否同乡同姓；
- [0023] S46、确定亲密度。
- [0024] 步骤S41中所述提取人员信息是指，从资金交易表中提取姓名、身份证号和银行账号信息，所述姓名、身份证号和银行账号信息从资金交易表的付款方账号、付款方名称、付款方证件号码、收款方账号、收款方名称和收款方证件号码中提取；
- [0025] 在步骤S42中，所述确定资金比值为确定两人之间的资金交易总值、资金交易净值、净值与总值的比值；
- [0026] 在步骤S43中，所述判断人员关系，为判断关系表中每条数据的主体和客体之间的关系，若关系表数据中的净值比总值的绝对值小于参数表中的净值比总值的阈值，则认为关系表中的主体ID和客体ID是成员关系，若关系表数据中的净值比总值的绝对值大于参数表中的净值比总值的阈值，则认为主体ID和客体ID是上下游关系；
- [0027] 在步骤S44中，所述确定共同交易对手个数为确定两人（假设是A和B）之间的共同交易对手个数，通过将人员表中的所有人员进行两两组合，进而得出共同交易对手个数；
- [0028] 在步骤S45中，所述判断是否同乡同姓为判断关系表中的主体和客体之间是否存在同乡、同姓关系；
- [0029] 在步骤S46中，所述亲密度为能够反应主体和客体之间关系紧密程度的数值，所述亲密度由多个能够反应主体和客体之间关系的维度构成，所述维度包括乡亲亲密度、共同交易对手亲密度和资金亲密度。
- [0030] 其中，所述步骤S44包括以下子步骤：
- [0031] S441、统计既与A保持“成员：成员”关系，也与B保持“成员：成员”关系的人员个数；
- [0032] S442、统计既与A保持“上游：下游”或“下游：上游”关系，也与B保持“上游：下游”或“下游：上游”关系，且A和B均位于“上游”一端的人员个数；
- [0033] S443、统计既与A保持“上游：下游”或“下游：上游”关系，也与B保持“上游：下游”或“下游：上游”关系，且A和B均位于“下游”的人员个数；
- [0034] S444、将步骤S441～S443中得到的人员个数加和，得到A和B的共同交易对手个数，根据两人的共同交易对手个数，确定两人是否同一团体。
- [0035] 在步骤S5中，所述分析结果展示，为通过展示不同人员之间的亲密度，表现不同人员之间的关系。
- [0036] 本发明另一方面还提供了一种基于资金交易数据分析人员关系的系统，该系统包括数据模块61、参数设置模块62、分析模块63和展示模块64。
- [0037] 所述数据模块61，用于建立数据表，并将相关资金交易数据导入并存储到数据表中；
- [0038] 所述参数设置模块62，能够设置或修改数据模块61数据表中参数表内的相关参数；
- [0039] 所述分析模块63，能够分析不同人员之间的同乡、同姓关系，上下游关系、团体关系以及亲密程度，进而体现不同人员之间的关系；
- [0040] 所述展示模块64，能够提取数据模块61中关系表内的数据，并展示给使用者。

[0041] 根据本发明提供一种基于资金交易数据分析人员关系的方法和系统,具有以下有益效果:

[0042] (1) 将多个Excel中的资金交易数据合并到数据库的一个表中,可便于统一处理,进而达到快速处理海量数据的效果;

[0043] (2) 能够分析资金交易数据中的人员之间的同乡、同姓、资金交易比例、共同交易对手等多种关系,综合得出两人的亲密度;

[0044] (3) 可同时得出人员之间的团体关系和上下游关系;

[0045] (4) 有清晰明确的流程和标准规范,可实现自动化分析;

[0046] (5) 参数表的设置,使得本方法具有使用灵活、兼容性高、针对性强等优点。

附图说明

[0047] 图1示出一种优选实施方式中基于资金交易数据分析人员关系的流程图;

[0048] 图2示出一种优选实施方式分析人员关系的流程图。

具体实施方式

[0049] 下面通过对本发明进行详细说明,本发明的特点和优点将随着这些示例性说明而变得更为清楚、明确。

[0050] 在这里专用的词“示例性”意为“用作例子、实施例或说明性”。这里作为“示例性”所说明的任何实施例不必解释为优于或好于其它实施例。尽管在附图中示出了实施例的各种方面,但是除非特别指出,不必按比例绘制附图。

[0051] 本发明适用于互联网、金融、公安、安全、海关缉私等领域的数据分析及情报分析工作,具有高速高效的优点,能够适应复杂、深层次的人员分析工作。

[0052] 一方面,本发明提供了一种基于资金交易数据分析人员关系的方法,该方法包括如下步骤:

[0053] S1、建立数据表;

[0054] S2、导入资金交易数据;

[0055] S3、设置参数;

[0056] S4、分析人员关系;

[0057] S5、分析结果展示。

[0058] 具体的,在步骤1中,所述建立数据表,是在数据库中建立表,包括资金交易表、人员表、关系表、参数表,

[0059] 数据表的应用,不仅使得海量数据可以统一处理,极大的提升了分析速度,提高了分析效率,还为分析过程中的中间数据提供存储空间,使得分析过程更加清晰明确,便于对分析结果进行复检。

[0060] 进一步地,所述资金交易表,用于存储资金交易数据,包括交易时间、交易方式、付款方账号、付款方名称、付款方证件号码、付款方银行、收款方账号、收款方名称、收款方证件号码、收款方银行、货币名称、原币金额、交易记录ID、报告机构、交易发生地、用途、涉外收支分类、代办人名称、代办人证件号码,优选地,所述资金交易表还设置有全局序号、来源文件、来源表。

[0061] 所述人员表,用于存储人员信息,包括人员ID、姓名、性别、出生日期、出生地行政区划、身份证号、银行账号、备注等内容。

[0062] 所述关系表,用于存储人员与人员之间关系的表,包括关系ID、主体ID、客体ID、资金交易净值、资金交易总值、净值比总值、资金关系、共同交易对手个数、乡亲关系、亲密度、备注等内容。

[0063] 所述参数表,用于存储本方法所分析过程中用到的相关参数,参数表的设置,使得本方法具有使用灵活、兼容性高、针对性强等优点。

[0064] 在步骤S2中,所述导入资金交易数据,是指将银行、反洗钱中心等单位的相关资金交易数据导入到步骤S1建立的数据表中,

[0065] 进一步地,所述银行、反洗钱中心等单位的相关资金交易数据为多个Excel文件,每个Excel文件中具有多个表格,将所有表格中的每条数据都导入到资金交易表中。

[0066] 在本发明中,对资金交易数据的导入的方法不做特别限定,任意能够实现数据导入的方法均可,例如使用SQL Server,通过Java、Python编程等实现数据导入等。

[0067] 优选地,在导入资金交易数据时,Excel文件中每条数据生成一个全局序号,所述全局序号不重复,更优选从1开始累加得到,使得数据的调用和查找更为方便,以提升分析速度。

[0068] 在一个优选的实施方式中,在导入资金交易数据时,在资金交易表中的来源文件中记录每条数据来源的Excel文件名称,在来源表中记录每条数据来源的Excel文件表格的名称。

[0069] 将多个Excel中的资金交易数据合并到数据库的一个表中,可便于统一处理,进而达到快速处理海量数据的效果。

[0070] 在步骤S3中,所述设置参数,为设置或修改参数表中的参数,所述参数为能够表征人员之间关系的数据,

[0071] 在本发明中,通过将人员关系程度在不同维度下数字化,并按一定的比重将不同维度下数字化的人员关系综合,进而量化的展现不同人员之间的关系以及密切程度

[0072] 所述维度包括人员出生背景维度、资金交易量维度、上下游关系维度和行业关系维度。

[0073] 根据本发明,针对资金交易量维度,在参数中设置有“资金交易最小金额”参数,当不同人员之间交易总额超过一定数量时,才对其进行人员关系分析,该参数的设置,有效的排除了相互之间只有较弱联系的人群,例如超市与普通顾客、饭店与食客等,使得分析速度显著提高,

[0074] 优选地,所述资金交易最小金额范围为50000~200000,以排除正常生活消费等。

[0075] 根据本发明,针对上下游关系维度,在参数中设置有“净值比总值的阈值”参数,

[0076] A和B两人之间的净值是指,在A和B所有交易中,若A收入B的资金则将金额转为正值,A给付B的金额则将金额转为负值,并将所有资金累加,

[0077] A和B两人之间的总值是指,A和B所有交易金额累加,

[0078] 两人之间交易的净值与总值的比值,能够反应出两人之间的关系,当两人为上下游关系时,净值与总值的比值远高于两人为同一团体成员时的比值。

[0079] 优选地,所述净值比总值的阈值设置为0.4~0.8,更优选为0.6。

- [0080] 在一个优选的实施方式中,所述净值比总值的阈值通过如下方法得到:
- [0081] S31、预估取值范围,
- [0082] 根据工作经验,对净值比总值的阈值的范围进行预估,例如预估值为0.3、0.4、0.5、……、0.9。
- [0083] S32、真实样本验证,
- [0084] 选取真实资金交易数据,针对某一组已查明人员关系的样本,按照本发明提供的方法,将预估值逐个代入真实资金交易数据中,确定代入结果与已查明的人员关系结果吻合度最高的预估值,将此预估值做为净值比总值的阈值。
- [0085] 优选地,所述真实资金交易数据大于100万条数据量。
- [0086] 在一个优选的实施方式中,针对多组已查明人员关系的样本,重复步骤S32,得到多个预估值,将得到的多个预估值平均,作为净值比总值的阈值。
- [0087] 根据本发明,针对行业关系维度,在参数中设置有“共同交易对手最小个数”参数,通过对两人之间是否有相同的交易对象,能够准确的判定两人是否在同一行业,
- [0088] 优选地,共同交易对手最小个数设置为3。
- [0089] 在一个优选的实施方式中,所述共同交易对手最小个数通过如下方法得到:
- [0090] S33、预估取值范围,
- [0091] 根据工作经验,对共同交易对手最小个数进行预估,例如预估值为1、2、3、4、5。
- [0092] S34、真实样本验证,
- [0093] 选取真实资金交易数据,针对某一组已查明人员关系的样本,按照本发明提供的方法,将预估值逐个代入真实资金交易数据中,确定代入结果与已查明的人员关系结果吻合度最高的预估值,将此预估值做为共同交易对手最小个数。
- [0094] 优选地,所述真实资金交易数据大于100万条数据量。
- [0095] 在一个优选的实施方式中,针对多组已查明人员关系的样本,重复步骤S32,得到多个预估值,将得到的多个预估值平均,作为共同交易对手最小个数。
- [0096] 在本发明中,所述设置参数,还包括设置不同维度在综合分析中的权重,具体的,包括“同乡亲密度”权重、“同姓亲密度”权重、“同乡且同姓亲密度”权重、“共同交易对手亲密度”权重、“成员亲密度”权重和“上下游亲密度”权重。
- [0097] 优选地,不同维度的权重如表1所示,其中,同乡亲密度为0.1~0.3,优选为0.2;同姓亲密度为0.1~0.3,优选为0.2;同乡且同姓亲密度为0.2~0.6,优选为0.4;共同交易对手亲密度为0.1~0.3,优选为0.1;成员亲密度为0.2~0.6,优选为0.4;上下游亲密度为0.1~0.3,优选为0.1。
- [0098] 在一个优选的实施方式中,所述不同维度的权重通过如下方法得到:
- [0099] S33、预估取值范围,
- [0100] 根据工作经验,对不同维度的权重进行预估。
- [0101] S34、真实样本验证,
- [0102] 选取真实资金交易数据,针对某一组已查明人员关系的样本,按照本发明提供的方法,将预估值逐个代入真实资金交易数据中,确定代入结果与已查明的人员关系结果吻合度最高的预估值,将此预估值做为不同维度的权重。
- [0103] 优选地,所述真实资金交易数据大于100万条数据量。

[0104] 在一个优选的实施方式中,针对多组已查明人员关系的样本,重复步骤S32,得到多组预估值,将得到的多组预估值平均,作为不同维度的权重。

[0105] 表1

参数名	参数值
资金交易最小金额	100000
净值比总值的阈值	0.6
共同交易对手最小个数	3
同乡亲密度	0.2
同姓亲密度	0.2
同乡且同姓亲密度	0.4
共同交易对手亲密度	0.1
成员亲密度	0.4
上下游亲密度	0.1

[0107] 在步骤S4中,所述分析人员关系,包括如下子步骤:

[0108] S41、提取人员信息;

[0109] S42、确定资金比值;

[0110] S43、判断人员关系;

[0111] S44、确定共同交易对手个数;

[0112] S45、判断是否同乡同姓;

[0113] S46、确定亲密度。

[0114] 根据本发明,步骤S41中所述提取人员信息是指,从资金交易表中提取姓名、身份证号和银行账号信息。

[0115] 具体的,所述姓名、身份证号和银行账号信息从资金交易表的付款方账号、付款方名称、付款方证件号码、收款方账号、收款方名称和收款方证件号码中提取。

[0116] 若身份证号是15位,则按照15位身份证号转18位的规则统一转化为18位。

[0117] 进一步地,从身份证号码中提取人员相关信息:从身份证号码中提取前6位,生成出生地行政区划;从身份证号码的7-14位提取出生日期;从第17位判断性别,单数为男、双数为女。

[0118] 更进一步地,将提取到的姓名、身份证号、银行账号、出生地行政区划、出生日期、性别等信息存储在人员表中,并生成人员ID。每个身份证号具有一个人员ID,若出现一个身份证具有相同的两个银行账号的,把两个账号合到一起,中间用“,”分隔。

[0119] 在步骤S42中,所述确定资金比值为确定两人(假设是A和B)之间的资金交易总值、资金交易净值、净值与总值的比值。

[0120] 交易资金总值的多少以及交易的净值有效的反应了人员之间的密切程度以及人员之间的关系,交易总值尤其是净值与总值的比值越高,则人员之间的关系越密切。

[0121] 具体的,通过将人员表中的所有人员进行两两组合,结合资金交易表中对应人员的原币金额,进而得出资金比值,

[0122] 其中,A和B两人之间的资金交易总值为,将资金交易表中A和B之间所有不重复的资金交易数据中的原币金额乘以原币对人民币的汇率得到人民币金额,然后将人民币金额

累加起来得到；

[0123] A和B两人之间的资金交易净值为,将资金交易表中A和B之间所有不重复的资金交易数据中的“原币金额”乘以原币对人民币的汇率得到人民币金额,如果是A收入的金额就保留金额为正值,如果是A付出的金额就将金额转为负值,然后将人民币金额累加起来得到,所述资金交易净值可能是正值也可能是负值；

[0124] A和B两人之间的净值与总值的比值为,将A和B两人之间的资金交易净值除以A和B两人之间的资金交易总值得到。

[0125] 进一步地,若A和B两人之间的资金交易总值大于或等于参数表中的资金交易最小金额,则将人员表中A的人员ID、B的人员ID以及A和B两人之间的资金交易净值、A和B两人之间的资金交易总值、A和B两人之间的净值与总值的比值分别存储到关系表中的主体ID、客体ID、资金交易净值、资金交易总值、净值比总值中。

[0126] 在步骤S43中,所述判断人员关系,为判断关系表中每条数据的主体和客体之间的关系。

[0127] 根据两人之间的资金交易净值、总值以及净值与总值的比例,可以分析出两人是上下游关系还是为同一团体的成员关系。

[0128] 成员关系的确定,使得人员之间的关系更加明晰,极大提高了分析人员关系的准确性。

[0129] 在本发明中,若关系表数据中的净值比总值的绝对值小于参数表中的净值比总值的阈值,即两人之间的交易未出现资金明显流向一方的现象,则认为关系表中的主体ID和客体ID是成员关系,将关系表中的该条数据的资金关系存储为“成员:成员”；

[0130] 若关系表数据中的净值比总值的绝对值大于参数表中的净值比总值的阈值,即两人之间的交易出现资金明显流向一方,则认为主体ID和客体ID是上下游关系。进一步地,若资金交易净值是正值,则主体ID是上游,客体ID是下游,将该条数据的资金关系存储为“上游:下游”；若资金交易净值是负值,则主体ID是下游,客体ID是上游,将该条数据的资金关系存储为“下游:上游”。

[0131] 在步骤S44中,所述确定共同交易对手个数为确定两人(假设是A和B)之间的共同交易对手个数。

[0132] 具体的,通过将人员表中的所有人员进行两两组合,进而得出共同交易对手个数,包括以下子步骤:

[0133] S441、统计既与A保持“成员:成员”关系,也与B保持“成员:成员”关系的人员个数；

[0134] S442、统计既与A保持“上游:下游”或“下游:上游”关系,也与B保持“上游:下游”或“下游:上游”关系,且A和B均位于“上游”一端的人员个数；

[0135] S443、统计既与A保持“上游:下游”或“下游:上游”关系,也与B保持“上游:下游”或“下游:上游”关系,且A和B均位于“下游”的人员个数；

[0136] S444、将步骤S441~S443中得到的人员个数加和,得到A和B的共同交易对手个数,由于同行之间的交易对手会有较大重叠,根据两人的共同交易对手个数,可以分析出两人是否在同一行业。

[0137] 进一步地,若共同交易对手个数大于或等于参数表中的“共同交易对手最小个数”,则判断关系表中是否已存储有A为主体、B为客体或A为客体、B为主体的数据。

- [0138] 如果已存储有,则将共同交易对手个数存储到“共同交易对手个数”字段;
- [0139] 如果未存储有,则在关系表新增一条数据,以存储A的人员ID、B的人员ID、共同交易对手个数。
- [0140] 根据本发明,在步骤S45中,所述判断是否同乡同姓为判断关系表中的主体和客体之间是否存在同乡、同姓关系。
- [0141] 具体的,根据两人姓名中的第一个汉字是否相同,判断两人是否存在同姓关系,根据两人的出生地行政区划是否完全相同,判断两人是否存在同乡关系,所述同乡,是指出生地为同一个市辖区、县级市或旗。
- [0142] 若两人是同姓关系,则在该条数据的乡亲关系中存储“同姓”;
- [0143] 若两人是同乡关系,则在该条数据的乡亲关系中存储“同乡”;
- [0144] 若两人既是同乡又是同姓关系,则在该条数据的乡亲关系中存储“同乡且同姓”。
- [0145] 在步骤S46中,所述亲密度为能够反应主体和客体之间关系紧密程度的数值,所述亲密度优选由多个能够反应主体和客体之间关系的维度构成,所述维度包括人员出生背景维度、资金交易量维度、上下游关系维度和行业关系维度。
- [0146] 进一步地,所述乡亲亲密度、共同交易对手亲密度和资金亲密度根据参数表和关系表得到,
- [0147] 若乡亲关系为“同乡”,则乡亲亲密度等于参数表中的“同乡亲密度”对应的权重值;若乡亲关系为“同姓”,则乡亲亲密度等于参数表中的“同姓亲密度”对应的权重值;若乡亲关系为“同乡且同姓”,则乡亲亲密度等于参数表中的“同乡且同姓亲密度”对应的权重值,通过权重的分配,量化的展现了两人之间人员出生背景维度的关系;
- [0148] 若资金关系为“成员:成员”,则资金亲密度等于参数表中的“成员亲密度”对应的权重值。如果资金关系为“上游:下游”或“下游:上游”,则资金亲密度等于参数表中的“上下游亲密度”对应的权重值;
- [0149] 共同交易对手亲密度等于关系表中的共同交易对手个数乘以参数表中的共同交易对手亲密度对应的权重值,两人的共同交易对手个数越多,则代表两人的业务重叠性越高,两人的关系也就越密切;
- [0150] 更进一步地,所述亲密度由上述不同维度下的代表人员关系的数值加和而成,
- [0151] 在一个优选的实施方式中,所述亲密度为乡亲亲密度、共同交易对手亲密度与资金亲密度的加和。
- [0152] 在步骤S5中,所述分析结果展示,为通过展示不同人员之间的亲密度,表现不同人员之间的关系,优选地,以与被分析人的亲密度从高到低的顺序进行展示,
- [0153] 在本发明中,对展示的方式不做特别限制,可以是表格形式、图表形式或其它能够表现亲密关系的形式,优选以表格的形式进行展示。
- [0154] 另一方面,本发明提供了一种基于资金交易数据分析人员关系的系统,该系统包括数据模块61、参数设置模块62、分析模块63和展示模块64。
- [0155] 所述数据模块61,用于建立数据表,并将相关资金交易数据导入并存储到数据表中。
- [0156] 所述数据表,包括资金交易表、人员表、关系表、参数表。
- [0157] 所述将相关资金交易数据导入为通过Python编程等实现数据导入。

[0158] 所述参数设置模块62,能够设置或修改数据模块61数据表中参数表内的相关参数,所述参数包括资金交易最小金额、净值比总值的阈值、共同交易对手最小个数、同乡亲密度、同姓亲密度、同乡且同姓亲密度、共同交易对手亲密度、成员亲密度、上下游亲密度。

[0159] 所述分析模块63,能够分析不同人员之间的同乡、同姓关系,上下游关系、团体关系以及亲密程度,进而体现不同人员之间的关系,

[0160] 进一步地,所述分析模块63包括人员信息子模块631、资金比例子模块632、人员关系子模块633、共同交易对手模块634、同乡同姓模块635、亲密度模块636。

[0161] 所述人员信息子模块631能够根据上述步骤S41的方法,提取数据模块61中资金交易表内的姓名、身份证号和银行账号信息,并存储在数据模块61中的人员表内。

[0162] 进一步地,所述人员信息子模块631还能够根据上述步骤S45的方法,提取数据模块61中的人员表内的信息,判断是否存在同乡、同姓关系,并将结果存储在数据模块61中的关系表内。

[0163] 所述资金比例子模块632,能够根据上述步骤S42的方法,提取数据模块61中的人员表和资金交易表内的信息,得到资金交易总值、资金交易净值、净值与总值的比值并存储在数据模块61中的关系表内。

[0164] 所述人员关系子模块633,能够提取数据模块61中关系表内的信息,根据上述步骤S43的方法,得到关系表中每条数据的主体和客体之间的关系,并将结果存储在关系表中。

[0165] 所述共同交易对手模块634,能够提取数据模块61中人员内的信息,根据上述步骤S44的方法,得到共同交易对手个数信息,并存储在数据模块61中关系表内。

[0166] 所述亲密度模块636,能够提取数据模块61中关系表和参数表内的信息,根据上述步骤S46的方法,得到亲密度信息,并存储在数据模块61的关系表内。

[0167] 所述展示模块64,能够提取数据模块61中关系表内的数据,并展示给使用者,优选地,所述展示模块64具有交互功能,能够接受使用者输入的检索条件,并根据检索条件显示数据模块61中关系表内的相关信息。

[0168] 实施例

[0169] 实施例1

[0170] 海关一次稽查中发现张三具有走私行为,使用该方法对张三的关系进行分析。

[0171] 在数据库中建立4个表:资金交易表、人员表、关系表、参数表。

[0172] 资金交易表的数据结构如表2所示:

[0173] 表2

[0174]

字段名称	数据类型	备注
序号	整数型	原始数据每个表格中的数据标识,从1开始。
交易时间	字符串型	14位,年月日时分秒
交易方式	字符串型	
付款方账号	字符串型	
付款方名称	字符串型	
付款方证件号码	字符串型	
付款方银行	字符串型	
收款方账号	字符串型	

收款方名称	字符串型	
收款方证件号码	字符串型	
收款方银行	字符串型	
货币名称	字符串型	
原币金额	数值型	正值
交易记录ID	字符串型	
报告机构	字符串型	
交易发生地	字符串型	
用途	字符串型	
涉外收支分类	字符串型	
代办人名称	字符串型	
代办人证件号码	字符串型	
全局序号	整数型	新加字段,合并后数据的唯一标识,从1开始
来源文件	字符串型	新加字段,记录了数据所在的文件名称
来源表	字符串型	新加字段,记录了数据所在的表名称,即人员姓名

[0175] 人员表的数据结构如表3所示:

[0176] 表3

字段名	数据类型	备注
人员 ID	整数型	主键, 从 1 开始增加

[0177]

姓名	字符串型	
性别	字符串型	男、女
出生日期	字符串型	8 位数字代表年月日
出生地行政区划	字符串型	6 位数字代表出生地的省市区县
身份证号	字符串型	18 位, 15 位身份证号也转为 18 位
银行账号	字符串型	如有多个银行账号, 用 “,” 分隔
备注	字符串型	

[0178]

[0179] 关系表的数据结构如表4所示:

[0180] 表4

字段名	数据类型	备注
关系 ID	整数型	主键, 从 1 开始增加
主体 ID	整数型	人员 ID
客体 ID	整数型	人员 ID
资金交易净值	数值型	两人之间的资金交易净值 (主体收入减去主体支出), 可以是负值
资金交易总值	数值型	两人之间的资金交易总值 (主体收入加上主体支出), 正值
净值比总值	数值型	净值与总值的比值, 可以是负值
资金关系	字符串型	主体在前客体在后, 用汉字冒号分隔, 如“成员: 成员”、“上游: 下游”、“下游: 上游”
共同交易对手个数	整数型	
乡亲关系	字符串型	“同乡”、“同姓”、“同乡且同姓”
亲密度	数值型	正值
备注	字符串型	

[0182] 参数表的数据结构如表5所示:

[0183] 表5

[0184]

参数名	参数值
资金交易最小金额	100000
净值比总值的阈值	0.6
共同交易对手最小个数	3
同乡亲密度	0.2
同姓亲密度	0.1
同乡且同姓亲密度	0.4
共同交易对手亲密度	0.1
成员亲密度	0.4
上下游亲密度	0.1

[0185] 从银行拿到资金交易数据, 导入到资金交易表中, 得到表6:

[0186] 表6

[0187]

序号	012356	01535165	03521546
交易时间	20180814	20180914	201801001
交易方式	转账	转账	转账
付款方账号	43922675* ***6583	43922675*** *6583	53592897****2 683
付款方名称	张三	张三	李四
付款方证件号码	440605198 60520***6	44060519860 520***6	4406051986092 5***3
付款方银行	中国邮政	中国邮政	MASTER
收款方账号	53592897* ***2683	53592897*** *2683	43922675****6 583
收款方名称	李四	李四	张三
收款方证件号码	440605198 60925***3	44060519860 925***3	4406051986052 0***6
收款方银行	MASTER	MASTER	中国邮政
货币名称	人民币	美元	人民币
原币金额	100000	10000	130000
交易记录ID	3652152	3652152	3652152
报告机构	中国邮政	中国邮政	中国邮政
交易发生地	广州	广州	广州
用途	--	--	--
涉外收支分类	--	--	--
代办人名称	李*	李*	李*
代办人证件号码	440*****	440*****	440*****
全局序号	021563	021563	021563
来源文件	中国邮政	中国邮政	中国邮政
来源表	大厅转账表	大厅转账表	大厅转账表

[0188] 根据表6,提取人员信息,得到人员表,如表7所示:

[0189] 表7

[0190]

人员 ID	01356	16523	12653
姓名	张三	李四	王五
性别	男	男	男
出生日期	19860520	19860925	19900122
出生地 行政区划	440605	440605	110119
身份证	440605198	44060519860925	11012019900122

[0191]	号	60520***6	***3	***5	
	银行账号	43922675* ***6583	53592897****26 83	62258813****26 83	
	备注				

[0192] 根据表6,确定资金比例,其中,张三(人员ID01235)和李四(人员ID016523)之间有3笔交易,张三给李四转了两笔,第一笔是100000元人民币,第二笔是10000元美金,美元对人民币的汇率是7;李四给张三转了一笔130000元人民币。张三和李四的资金交易总值=100000*1+10000*7+130000=300000元人民币。张三和李四的资金交易净值=100000-100000-70000=-40000元人民币。净值比总值=-40000/300000=-0.13。资金交易总值300000>资金交易最小金额100000,得到如表8所示关系表,同样方法得到张三与王五(人员ID12653)之间资金比例:

[0193] 表8

[0194]	关系 ID	01235	01563
	主体 ID	01356	01356
	客体 ID	16523	12653
	资金交易净值	-40000	90000
	资金交易总值	300000	90000
	净值比总值	-0.13	1
	资金关系		
	共同交易对手个数		
	乡亲关系		
	亲密度		
	备注		

[0195] 根据表8中的数据,判断主体和客体之间的关系,张三和李四的净值比总值为-0.13,绝对值为0.13,小于净值比总值的阈值为0.6,因此张三(人员ID01235)和李四(人员ID016523)是成员关系,同样方法得到张三与王五(人员ID12653)、张三与陈七(人员ID15365)、张三与杨八(人员ID55223)之间关系均为下游:上游;

[0196] 同时,根据上述方法得到李四与王五、陈七、杨八之间的关系也为下游:上游,则张三与李四的共同交易对手个数为3,如表9所示:

[0197] 表9

[0198]	关系 ID	01235	01563	01563	01563
	主体 ID	01356	01356	01356	01356
	客体 ID	16523	12653	15365	55223
	资金交易净值	-40000	-90000	-80000	-90000
	资金交易总值	300000	90000	100000	100000
	净值比总值	-0.13	-1	-0.8	-0.9
	资金关系	成员： 成员	下游：上 游	下游：上 游	下游：上 游
	共同交易对手个数	3	0	0	0
	乡亲关系
	亲密度
	备注

[0199] 根据表7和表9,可判断张三和李四、王五、陈七、杨八之间的同乡、同姓关系,由于姓名中的第一个汉字都不同,张三与其他人都不是同姓关系,再根据出生地行政区划是否相同,判定是否为同乡,结果如表10所示:

[0200] 表10

[0201]	关系 ID	01235	01563	01563	01563
	主体 ID	01356	01356	01356	01356
	客体 ID	16523	12653	15365	55223
	资金交易净值	-40000	-90000	-80000	-90000
	资金交易总值	300000	90000	100000	100000
[0202]	净值比总值	-0.13	-1	-0.8	-0.9
	资金关系	成员： 成员	下游：上 游	下游：上 游	下游：上 游
	共同交易对手个数	3
	乡亲关系	同乡	--	--	--
	亲密度
	备注

[0203] 根据表10中的数据与表5的参数相结合,可得到张三与李四、王五、陈七、杨八之间的亲密度,其中,张三与李四的关系为同乡,则相亲亲密度为0.2,共同交易对手个数为3,则共同交易对手亲密度为 $3*0.1$,资金关系为“成员:成员”,则资金亲密度为0.4,综合得到张三与李四的亲密度= $0.2+3*0.1+0.4=0.9$,同样方法得到张三与王五、陈七、杨八之间的亲密度,结果如表11所示:

[0204] 表11

	关系 ID	01235	01563	01563	01563
	主体 ID	01356	01356	01356	01356
	客体 ID	16523	12653	15365	55223
	资金交易净值	-40000	-90000	-80000	-90000
	资金交易总值	300000	90000	100000	100000
[0205]	净值比总值	-0.13	-1	-0.8	-0.9
	资金关系	成员： 成员	下游：上 游	下游：上 游	下游：上 游
	共同交易对手个数	3	0	0	0
	乡亲关系	同乡	--	--	--
	亲密度	0.9	0.5	0.7	0.3
	备注

[0206] 将与张三相关的信息按照亲密度关系进行展示,结果如表12所示:

[0207]	被分析人: 张三
	被分析人证件号: 44060519860520***6

	相关 人员	亲 密 度	证件号	资 金 交 易 净 值	资 金 交 易 总 值
[0208]	李四	0.9	44060519860925***3	-40000	300000
	陈七	0.7	11011919900307***8	-80000	100000
	王五	0.5	11010120090507***5	-90000	90000
	杨八	0.3	31011420020313***0	-90000	100000

[0209] 通过本发明所述方法及系统,能够在10s以内快速分析展示出人员之间的关系,并按照亲密程度对与其有关的人员进行展示,极大提高了分析速率,提高了分析准确率,海关警方可根据分析出的人员关系对相关人员进行调查,以查明更多有用证据。

[0210] 实验例

[0211] 实验例1

[0212] 在某次案情分析过程中,按照本发明所示的方法进行分析,与业务人员原来手工的分析结果进行比较,对比结果如表13所示,

[0213] 表13

	分析时间	分析结果
[0214] 本发明	7S	发现23人与被分析人有关，其中关系亲密度按高到底依次排列，并能够查询其与被分析人是否为同乡、是否为同一团体、是否为上下游关系等
传统方法	2个工作日	发现5人与被分析人有关，但关系程度未知

[0215] 实验例2

[0216] 本发明的参数设置如表1所示，其中的参数值经过100万真实的资金交易数据的测试检验。在某次案情分析过程中，按照本发明的参数设置进行分析，与其他参数设置的分析进行比较，对比结果如表14所示，

[0217] 表14

		分析结果（发现与被分析人亲密度较高、且为客户关心的目标人员的人数）
本发明默认的参数设置		23
[0218] 其他参数设置	资金交易最小金额改为10000	10
	净值比总值的阈值改为0.5	20
	共同交易对手最小个数改为2	15
	同乡亲密度改为0.3	18
	同姓亲密度改为0.3	17
	同乡且同姓亲密度改为 0.3	14
	共同交易对手亲密度改为 0.2	15
	成员亲密度改为 0.5	16
上下游亲密度改为 0.2	19	

[0219] 本发明提供的参数，能够筛选出更多的与被分析人亲密度高、且客户更为关心的目标人员数。

[0220] 以上结合具体实施方式和范例性实例对本发明进行了详细说明，不过这些说明并不能理解为对本发明的限制。本领域技术人员理解，在不偏离本发明精神和范围的情况下，可以对本发明技术方案及其实施方式进行多种等价替换、修饰或改进，这些均落入本发明的范围内。本发明的保护范围以所附权利要求为准。

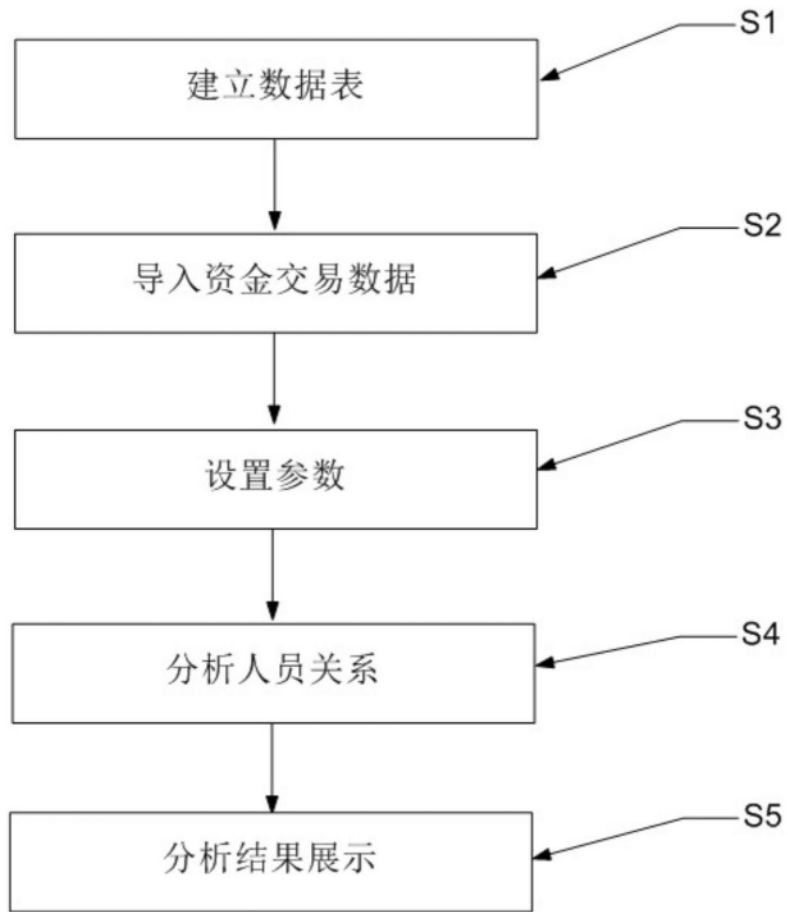


图1

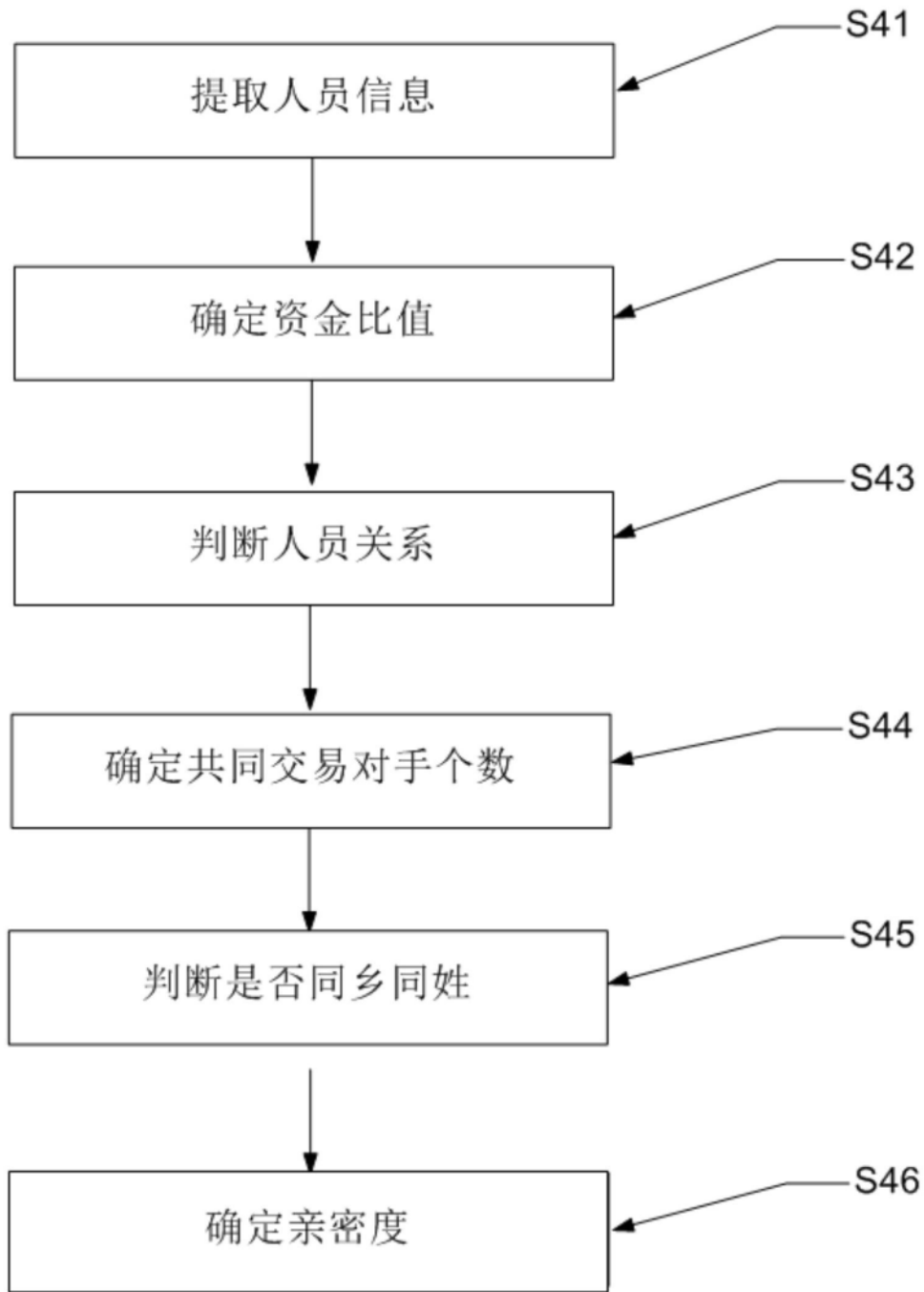


图2