



(12)发明专利

(10)授权公告号 CN 105426365 B

(45)授权公告日 2018.11.02

(21)申请号 201410377277.6

(56)对比文件

(22)申请日 2014.08.01

CN 1262949 C, 2006.07.05,

(65)同一申请的已公布的文献号

CN 102591995 A, 2012.07.18,

申请公布号 CN 105426365 A

CN 103365969 A, 2013.10.23,

(43)申请公布日 2016.03.23

US 2011/0206198 A1, 2011.08.25,

(73)专利权人 阿里巴巴集团控股有限公司

US 7251645 B1, 2007.07.31,

地址 英属开曼群岛大开曼资本大厦一座四
层847号邮箱

CN 103440199 A, 2013.12.11,

(72)发明人 包妮娜

审查员 王一

(74)专利代理机构 北京博浩百睿知识产权代理
有限责任公司 11134

代理人 宋子良

(51)Int.Cl.

G06F 17/30(2006.01)

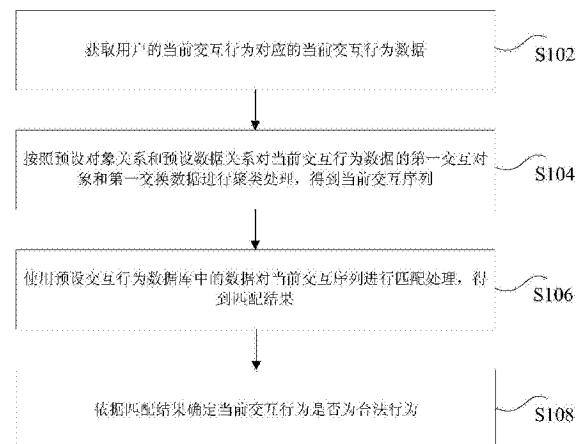
权利要求书3页 说明书11页 附图2页

(54)发明名称

区分交互行为的方法及装置

(57)摘要

本申请公开了一种区分交互行为的方法及装置。其中，该方法包括：获取用户的当前交互行为对应的当前交互行为数据；按照预设对象关系和预设数据关系对当前交互行为数据的第一交互对象和第一交换数据进行聚类处理，得到当前交互序列，其中，第一交换数据表示当前交互行为交换的数值；使用预设交互行为数据库中的数据对当前交互序列进行匹配处理，得到匹配结果；依据匹配结果确定当前交互行为是否为合法行为。通过本申请，解决了现有技术中对用户交互数据的分析不准确的问题，实现了准确分析用户的交互行为是否合法的效果。



1.一种区分交互行为的方法,其特征在于,包括:

获取用户的当前交互行为对应的当前交互行为数据;

按照预设对象关系和预设数据关系对所述当前交互行为数据的第一交互对象和第一交换数据进行聚类处理,得到当前交互序列,其中,所述第一交换数据表示所述当前交互行为交换的数值;

使用预设交互行为数据库中的数据对所述当前交互序列进行匹配处理,得到匹配结果;

依据所述匹配结果确定所述当前交互行为是否为合法行为;

其中,按照预设对象关系和预设数据关系对所述当前交互行为数据的第一交互对象和第一交换数据进行聚类处理,得到当前交互序列包括:

使用所述预设对象关系查询与所述第一交互对象对应的第四交互对象;

使用所述预设数据关系查询与所述第一交换数据对应的第四交换数据;

使用所述第四交互对象替换所述第一交互对象,使用所述第四交换数据替换所述第一交互对象,得到所述当前交互序列。

2.根据权利要求1所述的方法,其特征在于,在使用预设交互行为数据库对所述当前交互序列进行匹配处理,得到匹配结果之前,所述方法还包括:

采集所述用户的历史交互行为数据,得到历史交互行为集合;

从所述历史交互行为集合中提取发生频率符合预设阈值的历史交互行为数据,得到合法行为项集;

将所述合法行为项集中的所述历史交互行为数据转换为历史时间序列,生成所述预设交互行为数据库。

3.根据权利要求2所述的方法,其特征在于,从所述历史交互行为集合中提取发生频率符合预设阈值的历史交互行为数据,得到合法行为项集包括:

对所述历史交互行为数据中的第二交互对象和第二交换数据进行降维处理得到第三交互对象和第三交换数据,得到更新后的历史交互行为数据;

计算对应同一个所述第三交互对象的所述更新后的历史交互行为数据的发生频率;

将所述发生频率大于所述预设阈值的所述更新后的历史交互行为数据生成所述合法行为项集。

4.根据权利要求3所述的方法,其特征在于,对所述历史交互行为数据中的所述第二交互对象和所述第二交换数据进行降维处理得到第三交互对象和第三交换数据,得到更新后的历史交互行为数据包括:

对所述历史交互行为数据中的所述第二交互对象进行降维处理得到所述第三交互对象包括:将对应同一个所述第二交互对象的所述历史交互行为数据进行分组得到多个交互对象集合;对每个所述交互对象集合中的所述历史交互行为数据的所述第二交换数据和交互类型进行离散处理得到每个所述历史交互行为数据的交互分数;查询得到对应所述交互分数的所述第三交互对象;

对所述历史交互行为数据中的所述第二交换数据进行降维处理得到所述第三交换数据;查询得到与所述第二交换数据对应的所述第三交换数据;

使用所述第三交互对象替换所述第二交互对象,并使用所述第三交换数据替换所述第

二交换数据,得到所述更新后的历史交互行为数据。

5. 根据权利要求1至4中任意一项所述的方法,其特征在于,

使用预设交互行为数据库中的数据对所述当前交互序列进行匹配处理,得到匹配结果包括:将所述当前交互序列与所述预设交互行为数据库中的数据逐个进行匹配处理,得到所述预设交互行为数据库中存在与所述当前交互序列相匹配的数据的第一匹配结果,或得到所述预设交互行为数据库中不存在与所述当前交互序列相匹配的数据的第二匹配结果;

依据所述匹配结果确定所述当前交互行为是否为合法行为包括:若得到所述第一匹配结果确定所述当前交互行为为所述合法行为;若得到所述第二匹配结果确定所述当前交互行为不为所述合法行为。

6. 一种区分交互行为的装置,其特征在于,包括:

获取模块,用于获取用户的当前交互行为对应的当前交互行为数据;

聚类处理模块,用于按照预设对象关系和预设数据关系对所述当前交互行为数据的第一交互对象和第一交换数据进行聚类处理,得到当前交互序列,其中,所述第一交换数据表示所述当前交互行为交换的数值;

匹配模块,用于使用预设交互行为数据库中的数据对所述当前交互序列进行匹配处理,得到匹配结果;

确定模块,用于依据所述匹配结果确定所述当前交互行为是否为合法行为;

其中,所述聚类处理模块包括:

第三查询模块,用于使用所述预设对象关系查询与所述第一交互对象对应的第四交互对象;

第四查询模块,用于使用所述预设数据关系查询与所述第一交换数据对应的第四交换数据;

第二替换模块,用于使用所述第四交互对象替换所述第一交互对象,使用所述第四交换数据替换所述第一交互对象,得到所述当前交互序列。

7. 根据权利要求6所述的装置,其特征在于,所述装置还包括:

采集模块,用于在使用预设交互行为数据库对所述当前交互序列进行匹配处理,得到匹配结果之前,采集所述用户的历史交互行为数据,得到历史交互行为集合;

提取模块,用于从所述历史交互行为集合中提取发生频率符合预设阈值的历史交互行为数据,得到合法行为项集;

转换模块,用于将所述合法行为项集中的所述历史交互行为数据转换为历史时间序列,生成所述预设交互行为数据库。

8. 根据权利要求7所述的装置,其特征在于,所述提取模块包括:

更新模块,用于对所述历史交互行为数据中的第二交互对象和第二交换数据进行降维处理得到第三交互对象和第三交换数据,得到更新后的历史交互行为数据;

第一计算模块,用于计算对应同一个所述第三交互对象的所述更新后的历史交互行为数据的发生频率;

生成模块,用于将所述发生频率大于所述预设阈值的所述更新后的历史交互行为数据,生成所述合法行为项集。

9. 根据权利要求8所述的装置,其特征在于,所述更新模块包括:

第一处理模块，用于对所述历史交互行为数据中的所述第二交互对象进行降维处理得到所述第三交互对象，所述第一处理模块包括：分组模块，用于将对应同一个所述第二交互对象的所述历史交互行为数据进行分组得到多个交互对象集合；离散处理模块，用于对每个所述交互对象集合中的所述历史交互行为数据的所述第二交换数据和交互类型进行离散处理得到每个所述历史交互行为数据的交互分数；第一查询模块，用于查询得到对应所述交互分数的所述第三交互对象；

第二处理模块，用于对所述历史交互行为数据中的所述第二交换数据进行降维处理得到所述第三交换数据；第二查询模块，用于查询得到与所述第二交换数据对应的所述第三交换数据；

第一替换模块，用于使用所述第三交互对象替换所述第二交互对象，并使用所述第三交换数据替换所述第二交换数据，得到所述更新后的历史交互行为数据。

10. 根据权利要求6至9中任意一项所述的装置，其特征在于，

所述匹配模块包括：匹配子模块，用于将所述当前交互序列与所述预设交互行为数据库中的数据逐个进行匹配处理，得到所述预设交互行为数据库中存在与所述当前交互序列相匹配的数据的第一匹配结果，或得到所述预设交互行为数据库中不存在与所述当前交互序列相匹配的数据的第二匹配结果；

所述确定模块包括：第一确定子模块，用于若得到所述第一匹配结果确定所述当前交互行为为所述合法行为；第二确定子模块，用于若得到所述第二匹配结果确定所述当前交互行为不为所述合法行为。

区分交互行为的方法及装置

技术领域

[0001] 本申请涉及数据处理领域,具体而言,涉及一种区分交互行为的方法及装置。

背景技术

[0002] 现有技术中在分析用户的交易行为时,一般会考虑到事件发生时涉及的人物、时间、类型和内容的综合关系,具体地,一般只从某些维度考虑行为的异常性。

[0003] 针对现有技术中对用户交互数据的分析不准确的问题,目前尚未提出有效的解决方案。

发明内容

[0004] 针对相关技术中对用户交互数据的分析不准确的问题,目前尚未提出有效的解决方案,为此,本申请的主要目的在于提供一种区分交互行为的方法及装置,以解决上述问题。

[0005] 为了实现上述目的,根据本申请的一个方面,提供了一种区分交互行为的方法,该方法包括:获取用户的当前交互行为对应的当前交互行为数据;按照预设对象关系和预设数据关系对当前交互行为数据的第一交互对象和第一交换数据进行聚类处理,得到当前交互序列,其中,第一交换数据表示当前交互行为交换的数值;使用预设交互行为数据库中的数据对当前交互序列进行匹配处理,得到匹配结果;依据匹配结果确定当前交互行为是否为合法行为。

[0006] 进一步地,在使用预设交互行为数据库对当前交互序列进行匹配处理,得到匹配结果之前,方法还包括:采集用户的历史交互行为数据,得到历史交互行为集合;从历史交互行为集合中提取发生频率符合预设阈值的历史交互行为数据,得到合法行为项集;将合法行为项集中的历史交互行为数据转换为历史时间序列,生成预设交互行为数据库。

[0007] 进一步地,从历史交互行为集合中提取发生频率符合预设阈值的历史交互行为数据,得到合法行为项集包括:对历史交互行为数据中的第二交互对象和第二交换数据进行降维处理得到第三交互对象和第三交换数据,得到更新后的历史交互行为数据;计算对应同一个第三交互对象的更新后的历史交互行为数据的发生频率;将发生频率大于预设阈值的更新后的历史交互行为数据生成合法行为项集。

[0008] 进一步地,对历史交互行为数据中的第二交互对象和第二交换数据进行降维处理得到第三交互对象和第三交换数据,得到更新后的历史交互行为数据包括:对历史交互行为数据中的第二交互对象进行降维处理得到第三交互对象包括:将对应同一个第二交互对象的历史交互行为数据进行分组得到多个交互对象集合;对每个交互对象集合中的历史交互行为数据的第二交换数据和交互类型进行离散处理得到每个历史交互行为数据的交互分数;查询得到对应交互分数的第三交互对象;对历史交互行为数据中的第二交换数据进行降维处理得到第三交换数据:查询得到与第二交换数据对应的第三交换数据;使用第三交互对象替换第二交互对象,并使用第三交换数据替换第二交换数据,得到更新后的历史

交互行为数据。

[0009] 进一步地,按照预设对象关系和预设数据关系对当前交互行为数据的第一交互对象和第一交换数据进行聚类处理,得到当前交互序列包括:使用预设对象关系查询与第一交互对象对应的第四交互对象;使用预设数据关系查询与第一交换数据对应的第四交换数据;使用第四交互对象替换第一交互对象,使用第四交换数据替换第一交互对象,得到当前交互序列。

[0010] 进一步地,使用预设交互行为数据库中的数据对当前交互序列进行匹配处理,得到匹配结果包括:将当前交互序列与预设交互行为数据库中的数据逐个进行匹配处理,得到预设交互行为数据库中存在与当前交互序列相匹配的数据的第一匹配结果,或得到预设交互行为数据库中不存在与当前交互序列相匹配的数据的第二匹配结果;依据匹配结果确定当前交互行为是否为合法行为包括:若得到第一匹配结果确定当前交互行为为合法行为;若得到第二匹配结果确定当前交互行为不为合法行为。

[0011] 为了实现上述目的,根据本申请的另一方面,提供了一种区分交互行为的装置,该装置包括:获取模块,用于获取用户的当前交互行为对应的当前交互行为数据;聚类处理模块,用于按照预设对象关系和预设数据关系对当前交互行为数据的第一交互对象和第一交换数据进行聚类处理,得到当前交互序列,其中,第一交换数据表示当前交互行为交换的数值;匹配模块,用于使用预设交互行为数据库中的数据对当前交互序列进行匹配处理,得到匹配结果;确定模块,用于依据匹配结果确定当前交互行为是否为合法行为。

[0012] 进一步地,装置还包括:采集模块,用于在使用预设交互行为数据库对当前交互序列进行匹配处理,得到匹配结果之前,采集用户的历史交互行为数据,得到历史交互行为集合;提取模块,用于从历史交互行为集合中提取发生频率符合预设阈值的历史交互行为数据,得到合法行为项集;转换模块,用于将合法行为项集中的历史交互行为数据转换为历史时间序列,生成预设交互行为数据库。

[0013] 进一步地,提取模块包括:更新模块,用于对历史交互行为数据中的第二交互对象和第二交换数据进行降维处理得到第三交互对象和第三交换数据,得到更新后的历史交互行为数据;第一计算模块,用于计算对应同一个第三交互对象的更新后的历史交互行为数据的发生频率;生成模块,用于将发生频率大于预设阈值的更新后的历史交互行为数据,生成合法行为项集。

[0014] 进一步地,更新模块包括:第一处理模块,用于对历史交互行为数据中的第二交互对象进行降维处理得到第三交互对象,第一处理模块包括:分组模块,用于将对应同一个第二交互对象的历史交互行为数据进行分组得到多个交互对象集合;离散处理模块,用于对每个交互对象集合中的历史交互行为数据的第二交换数据和交互类型进行离散处理得到每个历史交互行为数据的交互分数;第一查询模块,用于查询得到对应交互分数的第三交互对象;第二处理模块,用于对历史交互行为数据中的第二交换数据进行降维处理得到第三交换数据;第二查询模块,用于查询得到与第二交换数据对应的第三交换数据;第一替换模块,用于使用第三交互对象替换第二交互对象,并使用第三交换数据替换第二交换数据,得到更新后的历史交互行为数据。

[0015] 进一步地,聚类处理模块包括:第三查询模块,用于使用预设对象关系查询与第一交互对象对应的第四交互对象;第四查询模块,用于使用预设数据关系查询与第一交换数

据对应的第四交换数据；第二替换模块，用于使用第四交互对象替换第一交互对象，使用第四交换数据替换第一交互对象，得到当前交互序列。

[0016] 进一步地，匹配模块包括：匹配子模块，用于将当前交互序列与预设交互行为数据库中的数据逐个进行匹配处理，得到预设交互行为数据库中存在与当前交互序列相匹配的数据的第一匹配结果，或得到预设交互行为数据库中不存在与当前交互序列相匹配的数据的第二匹配结果；确定模块包括：第一确定子模块，用于若得到第一匹配结果确定当前交互行为为合法行为；第二确定子模块，用于若得到第二匹配结果确定当前交互行为不为合法行为。

[0017] 采用本申请实施例，在获取用户的当前交互行为的当前交互行为数据之后，可以按照预设对象关系和预设数据关系对当前交互行为数据的第一交互对象和第一交换数据进行聚类处理，得到当前交互序列，并根据当前交互序列确定当前交互行为是否为合法行为。在该实施例中，不仅仅是通过交易金额(即上述实施例中的交换数据)判断当前交互行为是否为合法行为，同时考虑到了当前交互行为的交互对象、交互时间，通过多个维度的对当前交互行为数据进行处理，判断交互行为是否为合法行为的准确度更高，解决了现有技术中对用户交互数据的分析不准确的问题，实现了准确分析用户的交互行为是否合法的效果。

附图说明

[0018] 此处所说明的附图用来提供对本申请的进一步理解，构成本申请的一部分，本申请的示意性实施例及其说明用于解释本申请，并不构成对本申请的不当限定。在附图中：

[0019] 图1是根据本申请实施例的交互行为数据的处理方法的流程图；

[0020] 图2是根据本申请实施例的交互行为数据的处理方法的框图；以及

[0021] 图3是根据本申请实施例的交互行为数据的处理装置的示意图。

具体实施方式

[0022] 首先，在对本申请实施例进行描述的过程中出现的部分名词或术语适用于如下解释：

[0023] 项集：由若干项组成的集合。设非空集合 $I = (i_1, i_2, \dots, i_n)$ 表示所有项的集合，其中 $i_k (1 \leq k \leq n)$ 称为项。

[0024] 时间序列：由若干项集组成的有序集合。序列 $S = (s_1, s_2, s_m)$ ，其中 $s_j (1 \leq j \leq m)$ 是项集，也称为序列的元素。

[0025] 多维时间序列：序列 $S = (s_1, s_2, s_m)$ 的元素包含的项 $>= 2$ ，该序列就称为多维时间序列。

[0026] 为了使本技术领域的人员更好地理解本申请方案，下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本申请一部分的实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都应当属于本申请保护的范围。

[0027] 需要说明的是，本申请的说明书和权利要求书及上述附图中的术语“第一”、“第

二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本申请的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0028] 图1是根据本申请实施例的交互行为数据的处理方法的流程图。如图1所示,该方法可以通过如下步骤实现:

[0029] 步骤S102:获取用户的当前交互行为对应的当前交互行为数据。

[0030] 步骤S104:按照预设对象关系和预设数据关系对当前交互行为数据的第一交互对象和第一交换数据进行聚类处理,得到当前交互序列,其中,第一交换数据表示当前交互行为交换的数值。

[0031] 步骤S106:使用预设交互行为数据库中的数据对当前交互序列进行匹配处理,得到匹配结果。

[0032] 步骤S108:依据匹配结果确定当前交互行为是否为合法行为。

[0033] 采用本申请实施例,在获取用户的当前交互行为的当前交互行为数据之后,可以按照预设对象关系和预设数据关系对当前交互行为数据的第一交互对象和第一交换数据进行聚类处理,得到当前交互序列,并根据当前交互序列确定当前交互行为是否为合法行为。在该实施例中,不仅仅是通过交易金额(即上述实施例中的交换数据)判断当前交互行为是否为合法行为,同时考虑到了当前交互行为的交互对象、交互时间,通过多个维度的对当前交互行为数据进行处理,判断交互行为是否为合法行为的准确度更高,解决了现有技术中对用户交互数据的分析不准确的问题,实现了准确分析用户的交互行为是否合法的效果。

[0034] 在本申请上述实施例中的交互行为可以是通过网络平台(如支付宝、淘宝等)进行虚拟资源或者有形资源的交易的行为。例如,交互行为交换数据(即交易金额)为3000的转账交易,在本实施例中根据交易对象(即上述实施例中的交互对象)及交互时间进行区分对待,如果交易对象是朋友或者家人,那风险系数就很低。此外本实施例中还考虑到交互时间的因素,例如给陌生人对象转账的时间,如果交易时间具有周期性,例如每月的1号发生那该事件风险系数会比同类事件在偶然时间发生的风险系数低。事件发生时刻各因素综合考虑使得用户异常行为的识别更准确和全面。

[0035] 根据本申请的上述实施例,在使用预设交互行为数据库对当前交互序列进行匹配处理,得到匹配结果之前,方法还可以包括:采集用户的历史交互行为数据,得到历史交互行为集合;从历史交互行为集合中提取发生频率符合预设阈值的历史交互行为数据,得到合法行为项集;将合法行为项集中的历史交互行为数据转换为历史时间序列,生成预设交互行为数据库。

[0036] 具体地,可以获取用户的历史交易行为(即历史交互行为)的历史交互行为集合,如图2所示的事件库 $X = (a_1, a_2, \dots, a_m)$,其中, a_i 表示历史交互行为数据。

[0037] 在本申请的上述实施例中,从历史交互行为集合中提取发生频率符合预设阈值的历史交互行为数据,得到合法行为项集可以包括:对历史交互行为数据中的第二交互对象

和第二交换数据进行降维处理得到第三交互对象和第三交换数据,得到更新后的历史交互行为数据;计算对应同一个第三交互对象的更新后的历史交互行为数据的发生频率;将发生频率大于预设阈值的更新后的历史交互行为数据生成合法行为项集。

[0038] 其中,上述实施例中的聚类分析指将物理或抽象对象的集合分组成为由类似对象组成的多个类的分析过程。具体到本申请实施例中,聚类处理可以为将第二交互对象分入不同的组中的操作,不同的组使用第三交互对象来表示。具体地,聚类处理的算法可以采用k-均值或者k-中心等聚类算法。上述实施例中的降维处理即为降低数据维数的处理,可以通过线性降维或者非线性降维,通过该种处理方法可以降低数据的复杂度。

[0039] 通过上述实施例,可以从历史交互行为集合中的数据中提取合法行为项集,在提取合法行为项集之前,可以对其中的历史交互行为数据进行维度离散化处理。具体地,通过对事件维度数据(即上述实施例中的历史交互行为数据)进行聚类处理(具体地可以通过降维处理实现聚类处理),去除细节干扰,保留事件(即历史交互行为)的主要形态,从而可以有效地降低挖掘算法的存储和计算代价,提高挖掘的效率和准确性。

[0040] 例如,在一个月内,用户A的交易对象(即上述实施例中的第二交互对象)可以有15个,其中,包括A1-A15,通过上述的聚类处理可以将15个第二交互对象分为如下几组:亲人、朋友和银行。

[0041] 上述的亲人、朋友和银行即为第三交互对象,这样将15个第二交互对象降低维数得到三个第三交互对象,降低了数据的复杂度。

[0042] 需要进一步说明的是,对历史交互行为数据中的第二交互对象和第二交换数据进行降维处理得到第三交互对象和第三交换数据,得到更新后的历史交互行为数据可以包括:对历史交互行为数据中的第二交互对象进行降维处理得到第三交互对象包括:将对应同一个第二交互对象的历史交互行为数据进行分组得到多个交互对象集合;对每个交互对象集合中的历史交互行为数据的第二交换数据和交互类型进行离散处理得到每个历史交互行为数据的交互分数;查询得到对应交互分数的第三交互对象;对历史交互行为数据中的第二交换数据进行降维处理得到第三交换数据:查询得到与第二交换数据对应的第三交换数据;使用第三交互对象替换第二交互对象,并使用第三交换数据替换第二交换数据,得到更新后的历史交互行为数据。

[0043] 上述实施例中,对第二交互对象进行降维处理和对第二交换数据的降维处理可以是并行进行的。

[0044] 具体地,对第二交互对象进行降维处理可以通过如下方法实现:统计相对于同一个第二交互对象的历史交互行为的发生频率,对每个历史交互行为数据中的交互类型和发生频率查表评分,得到每个历史交互行为数据的交互对象评分,获取与该交互对象评分对应的第一交互对象。

[0045] 如表1所示的用户A的2014年上半年历史交互行为(表1仅作示例性说明,本申请上述实施例不受该应用场景的限制):

[0046] 表1

[0047]

第二交互对象	交互类型	发生频率	第三交互对象
A1	养老转账	0.5	亲人

A2	购物	0.2	卖家
A3	购物	0.2	卖家
A4	养老转账	0.5	亲人
A5	话费	0.2	通信运行商
A6	话费	0.2	通信运行商
A7	养老转账	0.5	亲人
A8	饭费	0.1	其他
A9	养老转账	0.5	亲人
A10	养老转账	0.5	亲人

[0048] 如表1所示,用户A的交易对象(即上述实施例中的第二交互对象)可以有10个,其中包括A1-A10,通过上述的统计,可以对其发生频率和交互类型进行查表评分,具体地,可以通过计算交互类型与评分表中的关键字的相似度,将相似度大于预设相似度阈值的关键字的评分作为所述交互类型的评分,并查询交互对象表中对应该评分的第三交互对象。通过上述处理将10个第二交互对象分为了四类(也可以称之为四组):亲友、卖家、通信运营商和其他。

[0049] 又如,上述实施例中,如表2所示用户A发生的历史交互行为的交换数据的降维处理。

[0050] 表2

	第二交换数据	第三交换数据
[0051]	1000	高
	378	中
	500	中

[0052]	1000	高
	200	低
	200	低
	1000	高
	2000	高
	1000	高
	1000	高

[0053] 对历史交互行为数据中的第二交换数据进行降维处理得到第三交换数据:查询得到与第二交换数据对应的第三交换数据(如表2中的高、中、低)。具体地,在第二交换数据符合第一范围时,第三交换数据为高;在第二交换数据符合第二范围时,第三交换数据为中;在第二交换数据符合第三范围时,第三交换数据为低。其中,第一范围可以为 $[1000, \infty)$;第

二范围可以为[300,1000)；第三范围可以为(0,300)。

[0054] 在上述实施例中,第一范围、第二范围以及第三范围共同构成了上述的预设数据关系。

[0055] 在完成对第二交换数据和第二交互对象的降维处理之后,使用第三交换数据替换第二交换数据,并使用第三交互对象替换第二交互对象,得到更新后的历史交互行为数据,如表3所示。

[0056] 表3

交互时间	第三交换数据	交互类型	发生频率	第三交互对象
1月6日	高	养老转账	0.5	亲人
1月18日	中	购物	0.2	卖家
2月1日	中	购物	0.2	卖家
2月6日	高	养老转账	0.5	亲人
2月15日	低	话费	0.2	通信运行商
3月5日	低	话费	0.2	通信运行商
3月6日	高	养老转账	0.5	亲人
3月7日	高	饭费	0.1	其他
4月6日	高	养老转账	0.5	亲人

[0058]	5月6日	高	养老转账	0.5	亲人
--------	------	---	------	-----	----

[0059] 通过上述实施例,获取到的更新第二交互对象和第二交换数据的历史交互行为集合,然后从该更新后的历史交互行为集合中提取发生频率符合预设阈值的历史交互行为数据,得到合法行为项集。具体到上述的表中,可以提取发生频率大于0.1的历史交互行为数据生成合法行为项集,然后将合法行为项集中的历史交互行为数据转换为历史时间序列,生成预设交互行为数据库,如表4所示的合法行为项集中的历史交互行为数据可以使用六个描述数据来表示,表中的每一行用序列的形式表示即为历史时间序列,如第二行中的历史时间序列为:(2014,1,6,高,养老转账,亲人)。

[0060] 表4

[0061]

年	月	日	第三交换数据	交互类型	第三交互对象
2014	1	6	高	养老转账	亲人
2014	1	18	中	购物	卖家
2014	2	1	中	购物	卖家
2014	2	6	高	养老转账	亲人
2014	2	15	低	话费	通信运行商

2014	3	5	低	话费	通信运行商
2014	3	6	高	养老转账	亲人
2014	4	6	高	养老转账	亲人
2014	5	6	高	养老转账	亲人

[0062] 需要进一步说明的是,获取历史时间序列之后,还可以对其进行进一步地聚类分析,也即,可以将得到的合法行为项集中的历史时间序列进行合并处理得到预设交互行为数据库。

[0063] 具体地,将得到的合法行为项集中的历史时间序列进行合并处理包括:将得到的历史时间序列中至少三个属性数据相同的数据合并为一条历史时间序列,将其中不同的属性数据使用字符“*”代替。

[0064] 如,将上述的第二条和第三条的历史时间序列合并为(*,*,*中,购物,卖家)。

[0065] 在预设交互行为数据库中还可以保存有上述进行降维处理过程中第二交互对象与第三交互对象的映射关系将其作为预设对象关系。

[0066] 根据本申请的上述实施例,如图2所示,可以按照预设的交互亲密度模型和交换数据的模型对历史交互行为数据进行聚类处理。

[0067] 具体地,交互亲密度模型可以根据交易对象(即上述实施例中的第二交互对象)与用户的历史交互数据对交易对象进行分类;交换数据的模型可以根据用户的交易金额进行聚类,实现连续变量的离散化处理。

[0068] 上述实施例中的获取合法行为项集的过程即为项集频繁项目集挖掘的过程。在该实施例中,执行该步骤是为了发现用户的频繁交易行为,该步骤不会考虑到事件的时间因素影响。

[0069] 需要进一步说明的是,在该实施例中可以将用户的一次网上交易行为作为一个事务(即作为一条历史交互行为数据),伴随事务发生往往涉及多维度信息:第二交互时间、第二交互对象、交互类型(如转账、汇款等)以及交易金额(即上述实施例中的交换数据),通过上述实施例,可以挖掘出用户的历史交易行为模式(即上述实施例中的历史交互行为的模式),并建立相应的用户正常交易行为规则库(即上述实施例中的预设交互行为数据库),通过当前交互行为数据与预设交互行为数据库作匹配,以识别用户的风险交易行为。

[0070] 具体地,假设I是项的集合,给定一个交易数据库D(即上述实施例中的历史交互行为集合),其中每个事务(Transaction)t(即上述实施例中的历史交互行为数据)是I的非空子集,即每一个交易都与一个唯一的标识符TID(Transaction ID)对应。上述实施例中的频繁指某一项集出现的概率相对于所有记录而言必须达到某一水平(即达到预设概率阈值)。项集出现的概率称为支持度(Support),以一个包含A与B两个项的2-itemset为例,事务同时包含A和B的百分比即是{A,B}项的支持度,若支持度大于等于所设定的最小支持度(Minimum Support)门槛值时,则{A,B}称为频繁项集。一个满足最小支持度的k-itemset,则称为高频k-项集(Frequent k-itemset),一般表示为Large k或Frequent k。

[0071] 通过上述实施例得到合法行为项集之后,可以将合法行为项集中的历史交互行为数据转换为历史时间序列,生成预设交互行为数据库。具体地,获取用户频繁交易项集I(即上述实施例中的合法行为项集),项ik=(object,amount,trade_type),其中的object为第二交互对象,amount为交换数据,trade_type为交互类型,在该历史交互行为数据中并未包

含时间维度的数据,可选地,可以将频繁交易项集转换为多维时间序列(即上述实施例中的历史时间序列) $S = (s_1, s_2, \dots, s_n)$, 其中 $s_j = (\text{year}, \text{month}, \text{day}, i_k)$, 也即历史时间序列中增加了时间参数 $s_j = (\text{年}, \text{月}, \text{日}, \text{时})$ 。

[0072] 通过上述实施例,得到的历史时间序列可以是如下的表现形式“1个月以前购买奔腾PC的客户很可能在一个月内订购新的CPU芯片”。

[0073] 在得到的历史时间序列中可以包括历史交互行为数据的六个属性,该六个属性可以为:年、月、日、第三交互对象、第三交换数据以及交互类型。

[0074] 例如,用户在2014年3月15日向招商银行信用卡还款1500;在2014年4月15日向招商银行信用卡还款1000;在2014年5月15日向招商银行信用卡还款2500;……。对用户的该历史交互行为进行分析之后,该类交互行为即可成为一个合法行为项集,也即X月的15日向招商银行还款M。

[0075] 根据本申请的上述实施例,按照预设对象关系和预设数据关系对当前交互行为数据的第一交互对象和第一交换数据进行聚类处理,得到当前交互序列包括:使用预设对象关系查询与第一交互对象对应的第四交互对象;使用预设数据关系查询与第一交换数据对应的第四交换数据;使用第四交互对象替换第一交互对象,使用第四交换数据替换第一交互对象,得到当前交互序列。

[0076] 需要进一步说明的是,使用预设交互行为数据库中的数据对当前交互序列进行匹配处理,得到匹配结果包括:将当前交互序列与预设交互行为数据库中的数据逐个进行匹配处理,得到预设交互行为数据库中存在与当前交互序列相匹配的数据的第一匹配结果,或得到预设交互行为数据库中不存在与当前交互序列相匹配的数据的第二匹配结果;依据匹配结果确定当前交互行为是否为合法行为包括:若得到第一匹配结果确定当前交互行为为合法行为;若得到第二匹配结果确定当前交互行为不为合法行为。

[0077] 在通过上述实施例确定当前交互行为不为合法行为的情况下,如图2所示可以进行后续的风险评估。

[0078] 例如,如果当前交互行为数据表示的交互行为为:8月15日向招商银行信用卡还款5000。通过在预设交互行为数据库中作匹配,匹配到上述的合法行为项集可以如表5所示:

[0079] 表5

[0080]

年	月	日	交互对象	交互类型	交换数据
*	*	15	招商银行	信用卡还款	*
2014	8	15	招商银行	信用卡还款	5000

[0081] 如表5所示,上述的合法行为项集中的历史时间序列与当前交互序列相匹配,得到第一匹配结果。

[0082] 本申请的一种可选的实施例中,还可以通过计算历史时间序列与当前交互序列的相似度,通过相似度得到匹配分数。例如,在历史时间序列中提供的三个属性与当前交互序列完全匹配,则可以查表得到匹配分数为100,判断匹配分数是否大于预设分数阈值75,100 > 75,则确定当前交互行为数据对应的交互行为为合法行为;如果不符合则确定当前交互行为数据对应的交互行为不为合法行为,则向用户发出报警信号。

[0083] 需要说明的是,在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的

计算机系统中执行，并且，虽然在流程图中示出了逻辑顺序，但是在某些情况下，可以以不同于此处的顺序执行所示出或描述的步骤。

[0084] 图3是根据本申请实施例的交互行为数据的处理装置的示意图。如图3所示，该装置可以包括：获取模块10、聚类处理模块20、匹配模块30以及确定模块40。

[0085] 其中，获取模块，用于获取用户的当前交互行为对应的当前交互行为数据。

[0086] 聚类处理模块，用于按照预设对象关系和预设数据关系对当前交互行为数据的第一交互对象和第一交换数据进行聚类处理，得到当前交互序列，其中，第一交换数据表示当前交互行为交换的数值。

[0087] 匹配模块，用于使用预设交互行为数据库中的数据对当前交互序列进行匹配处理，得到匹配结果。

[0088] 确定模块，用于依据匹配结果确定当前交互行为是否为合法行为。

[0089] 采用本申请实施例，在获取用户的当前交互行为的当前交互行为数据之后，可以按照预设对象关系和预设数据关系对当前交互行为数据的第一交互对象和第一交换数据进行聚类处理，得到当前交互序列，并根据当前交互序列确定当前交互行为是否为合法行为。在该实施例中，不仅仅是通过交易金额(即上述实施例中的交换数据)判断当前交互行为是否为合法行为，同时考虑到了当前交互行为的交互对象、交互时间，通过多个维度的对当前交互行为数据进行处理，判断交互行为是否为合法行为的准确度更高，解决了现有技术中对用户交互数据的分析不准确的问题，实现了准确分析用户的交互行为是否合法的效果。

[0090] 在本申请上述实施例中的交互行为可以是通过网络平台(如支付宝、淘宝等)进行虚拟资源或者有形资源的交易的行为。例如，交互行为交换数据(即交易金额)为3000的转账交易，在本实施例中根据交易对象(即上述实施例中的交互对象)及交互时间进行区分对待，如果交易对象是朋友或者家人，那风险系数就很低。此外本实施例中还考虑到交互时间的因素，例如给陌生人对象转账的时间，如果交易时间具有周期性，例如每月的1号发生那该事件风险系数会比同类事件在偶然时间发生的风险系数低。事件发生时刻各因素综合考虑使得用户异常行为的识别更准确和全面。

[0091] 在本申请的上述实施例中，装置还可以包括：采集模块，用于在使用预设交互行为数据库对当前交互序列进行匹配处理，得到匹配结果之前，采集用户的历史交互行为数据，得到历史交互行为集合；提取模块，用于从历史交互行为集合中提取发生频率符合预设阈值的历史交互行为数据，得到合法行为项集；转换模块，用于将合法行为项集中的历史交互行为数据转换为历史时间序列，生成预设交互行为数据库。

[0092] 具体地，提取模块可以包括：更新模块，用于对历史交互行为数据中的第二交互对象和第二交换数据进行降维处理得到第三交互对象和第三交换数据，得到更新后的历史交互行为数据；第一计算模块，用于计算对应同一个第三交互对象的更新后的历史交互行为数据的发生频率；生成模块，用于将发生频率大于预设阈值的更新后的历史交互行为数据，生成合法行为项集。

[0093] 更进一步地，更新模块可以包括：第一处理模块，用于对历史交互行为数据中的第二交互对象进行降维处理得到第三交互对象，第一处理模块包括：分组模块，用于将对应同一个第二交互对象的历史交互行为数据进行分组得到多个交互对象集合；离散处理模块，

用于对每个交互对象集合中的历史交互行为数据的第二交换数据和交互类型进行离散处理得到每个历史交互行为数据的交互分数；第一查询模块，用于查询得到对应交互分数的第三交互对象；第二处理模块，用于对历史交互行为数据中的第二交换数据进行降维处理得到第三交换数据；第二查询模块，用于查询得到与第二交换数据对应的第三交换数据；第一替换模块，用于使用第三交互对象替换第二交互对象，并使用第三交换数据替换第二交换数据，得到更新后的历史交互行为数据。

[0094] 在本申请的上述实施例中，聚类处理模块可以包括：第三查询模块，用于使用预设对象关系查询与第一交互对象对应的第四交互对象；第四查询模块，用于使用预设数据关系查询与第一交换数据对应的第四交换数据；第二替换模块，用于使用第四交互对象替换第一交互对象，使用第四交换数据替换第一交互对象，得到当前交互序列。

[0095] 需要进一步说明的是，在本申请一个可选的实施例中，匹配模块包括：匹配子模块，用于将当前交互序列与预设交互行为数据库中的数据逐个进行匹配处理，得到预设交互行为数据库中存在与当前交互序列相匹配的数据的第一匹配结果，或得到预设交互行为数据库中不存在与当前交互序列相匹配的数据的第二匹配结果；确定模块包括：第一确定子模块，用于若得到第一匹配结果确定当前交互行为为合法行为；第二确定子模块，用于若得到第二匹配结果确定当前交互行为不为合法行为。

[0096] 本实施例中所提供的各个模块与方法实施例对应步骤所提供的使用方法相同、应用场景也可以相同。当然，需要注意的是，上述模块涉及的方案可以不限于上述实施例一中的内容和场景，且上述模块可以运行在计算机终端或移动终端，可以通过软件或硬件实现。

[0097] 从以上的描述中，可以看出，本申请实现了如下技术效果：

[0098] 采用本申请实施例，在获取用户的当前交互行为的当前交互行为数据之后，可以按照预设对象关系和预设数据关系对当前交互行为数据的第一交互对象和第一交换数据进行聚类处理，得到当前交互序列，并根据当前交互序列确定当前交互行为是否为合法行为。在该实施例中，不仅仅是通过交易金额(即上述实施例中的交换数据)判断当前交互行为是否为合法行为，同时考虑到了当前交互行为的交互对象、交互时间，通过多个维度的对当前交互行为数据进行处理，判断交互行为是否为合法行为的准确度更高，解决了现有技术中对用户交互数据的分析不准确的问题，实现了准确分析用户的交互行为是否合法的效果。

[0099] 显然，本领域的技术人员应该明白，上述的本申请的各模块或各步骤可以用通用的计算装置来实现，它们可以集中在单个的计算装置上，或者分布在多个计算装置所组成的网络上，可选地，它们可以用计算装置可执行的程序代码来实现，从而，可以将它们存储在存储装置中由计算装置来执行，或者将它们分别制作成各个集成电路模块，或者将它们中的多个模块或步骤制作成单个集成电路模块来实现。这样，本申请不限制于任何特定的硬件和软件结合。

[0100] 以上所述仅为本申请的优选实施例而已，并不用于限制本申请，对于本领域的技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

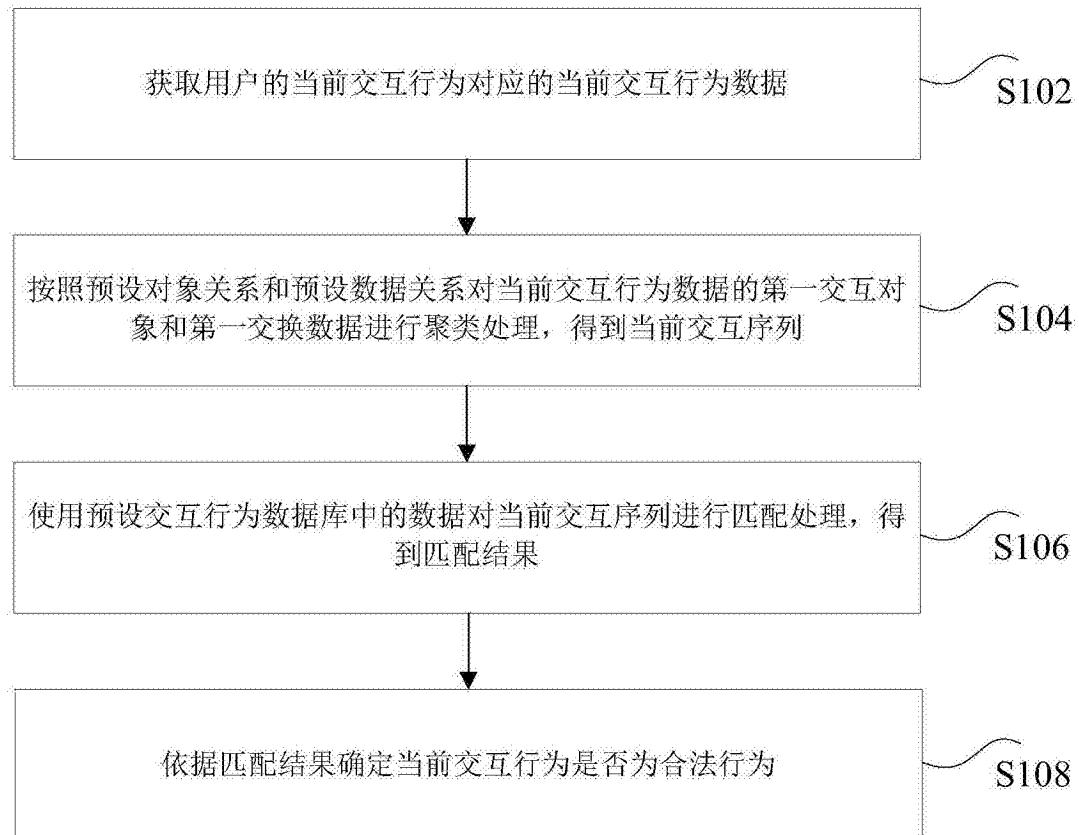


图1

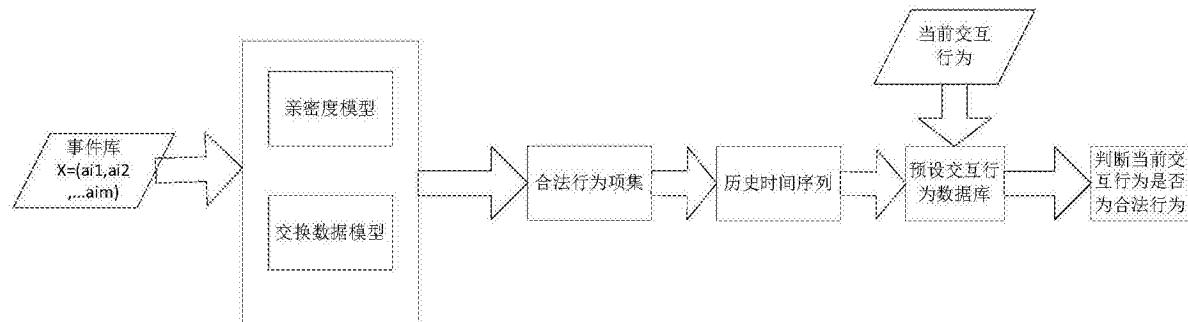


图2



图3