

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 3 月 12 日 (12.03.2020)



(10) 国际公布号
WO 2020/048059 A1

(51) 国际专利分类号:
G06F 16/9535 (2019.01) **G06Q 40/06** (2012.01)

(21) 国际申请号: PCT/CN2018/124252

(22) 国际申请日: 2018 年 12 月 27 日 (27.12.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201811022472.1 2018 年 9 月 3 日 (03.09.2018) CN

(71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518033 (CN)。

(72) 发明人: 陈泽晖(CHEN, ZeHui); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518033 (CN)。
刘琼(LIU, Qiong); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融

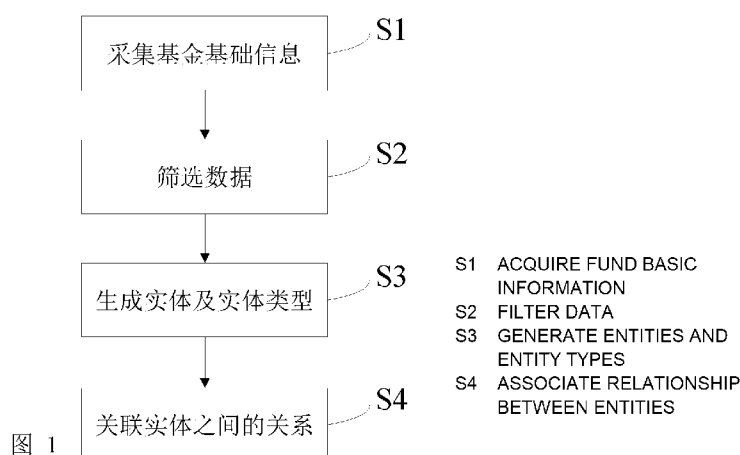
中心 23 楼, Guangdong 518033 (CN)。 蒋逸文(JIANG, YiWen); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518033 (CN)。

(74) 代理人: 北京市京大律师事务所(BEIJING JINGDA LAW FIRM); 中国北京市海淀区海淀中街 16 号 7 层 2 单元 703, Beijing 100080 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(54) Title: FUND DATA ASSOCIATION METHOD AND SYSTEM, COMPUTER DEVICE AND STORAGE MEDIUM

(54) 发明名称: 基金数据关联方法、系统、计算机设备和存储介质



(57) Abstract: A fund data association method and system, a computer device and a storage medium. Said method comprises: periodically acquiring fund basic information from a plurality of fund websites, filtering the fund basic information by means of a preset data screening rule to obtain a plurality of fund data sets, generating a plurality of entities for the same fund data set, each entity correspondingly generating a plurality of entity types, performing matching on every two entities, and when said entities are of the same entity types, configuring the entity types as a relationship type, associating the two entities by means of the relationship type, and using the two entities and the entity types to generate relational data having an association. By processing the fund basic information into relational data having an association for the user to subsequently find fund data having a tight association therewith, the present invention achieves effective extraction and association of mass public data.

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

(57) 摘要: 一种基金数据关联方法、系统、计算机设备和存储介质。该方法包括: 定时从多个基金网站采集基金基础信息, 通过预设的数据筛选规则, 对基金基础信息进行筛选, 得到多项基金数据集, 对同一项基金数据集生成多个实体, 每个实体对应生成多个实体类型, 对每两个实体进行匹配, 当具有相同的实体类型时, 将实体类型设置为关系类型, 以关系类型将两个实体进行关联, 将两个实体和实体类型生成有关联关系的关系数据。通过对基金基础信息处理成具有关联关系的关系数据, 供用户后续找到与其关联关系紧密的基金数据, 实现了对海量公开数据的有效提取和关联。

基金数据关联方法、系统、计算机设备和存储介质

本申请要求于 2018 年 09 月 03 日提交中国专利局、申请号为 201811022472.1、发明名称为“基金数据关联方法、系统、计算机设备和存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及金融技术领域，尤其涉及一种基金数据关联方法、系统、计算机设备和存储介质。

背景技术

基金市场高风险和高收益并存，因此对于基金数据的知识发现的研究一直受到人们的关注。近年来，随着计算机技术的飞速发展和存储能力的大大提高，使得这一方面的研究有了很大的发展。

目前市面上的财经类应用 APP 或者网站，每天有大量的基金信息数据产生，这些数据无疑有益于股民了解股市的走势，做出正确的投资决策。但是由于数据量大，股民无法从海量数据中提取或识别有效的、新颖的、有助于投资的数据信息。目前市面上的财经类应用 APP 或网站，仅能提供简单的基金的列表信息显示，简单的文字列表信息形式，客户无法获得各个基金之间相关联关系的直观感受，缺少定期的和实时的数据更新。

发明内容

有鉴于此，有必要针对目前市面上的基金信息数据，仅能提供简单的基金的列表信息显示的问题，提供一种基金数据关联方法、系统、计算机设备和存储介质。

一种基金数据关联方法，包括：

定时从多个基金网站采集基金基础信息，并将所述基金基础信息存储在缓存器中；

调用存储在所述缓存器中的所述基金基础信息，通过预设的数据筛选规则，对所述基金基础信息进行筛选，得到多项基金数据集，每项所述基金数据集内的基金数据具有关联关系，将多项所述基金数据集存储在所述缓冲器中；

依次调用存储在所述缓存器中的每项所述基金数据集，对同一项所述基金数据集生成多个实体，每个实体对应生成多个实体类型，所述实体包括基金、基金经理或基金公司，所述实体类型包括实体关联的其他所述基金数据；

对每两个所述实体进行匹配，当具有相同的所述实体类型时，将所述实体类型设置为关系类型，以关系类型将两个所述实体进行关联，将两个所述实体和所述实体类型生成有关联关系的关系数据，将所述关系数据存储在数据库中。

一种基金数据关联系统，包括：

采集单元，设置为定时从多个基金网站采集基金基础信息，并将所述基金基础信息存储在缓存器中；

筛选单元，设置为调用存储在所述缓存器中的所述基金基础信息，通过预设的数据筛选规则，对所述基金基础信息进行筛选，得到多项基金数据集，每项所述基金数据集内的基金数据具有关联关系，将多项所述基金数据集存储在所述缓冲器中；

生成实体及实体类型单元，设置为依次调用存储在所述缓存器中的每项所述基金数据集，对同一项所述基金数据集生成多个实体，每个实体对应生成多个实体类型，所述实体包括基金、基金经理或基金公司，所述实体类型包括实体关联的其他所述基金数据；

关联单元，设置为对每两个所述实体进行匹配，当具有相同的所述实体类型时，将所述实体类型设置为关系类型，以关系类型将两个所述实体进行关联，将两个所述实体和所述实体类型生成有关联关系的关系数据，将所述关系数据存储在数据库中。

一种计算机设备，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述处理器执行以下步骤：

定时从多个基金网站采集基金基础信息，并将所述基金基础信息存储在缓存器中；

调用存储在所述缓存器中的所述基金基础信息，通过预设的数据筛选规则，对所述基金基础信息进行筛选，得到多项基金数据集，每项所述基金数据集内的基金数据具有关联关系，将多项所述基金数据集存储在所述缓冲器中；

依次调用存储在所述缓存器中的每项所述基金数据集，对同一项所述基金数据集生成多个实体，每个实体对应生成多个实体类型，所述实体包括基金、

基金经理或基金公司，所述实体类型包括实体关联的其他所述基金数据；

对每两个所述实体进行匹配，当具有相同的所述实体类型时，将所述实体类型设置为关系类型，以关系类型将两个所述实体进行关联，将两个所述实体和所述实体类型生成有关联关系的关系数据，将所述关系数据存储在数据库中。

5 一种存储有计算机可读指令的存储介质，所述计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行以下步骤：

定时从多个基金网站采集基金基础信息，并将所述基金基础信息存储在缓存器中；

调用存储在所述缓存器中的所述基金基础信息，通过预设的数据筛选规则，
10 对所述基金基础信息进行筛选，得到多项基金数据集，每项所述基金数据集内的基金数据具有关联关系，将多项所述基金数据集存储在所述缓冲器中；

依次调用存储在所述缓存器中的每项所述基金数据集，对同一项所述基金数据集生成多个实体，每个实体对应生成多个实体类型，所述实体包括基金、基金经理或基金公司，所述实体类型包括实体关联的其他所述基金数据；

15 对每两个所述实体进行匹配，当具有相同的所述实体类型时，将所述实体类型设置为关系类型，以关系类型将两个所述实体进行关联，将两个所述实体和所述实体类型生成有关联关系的关系数据，将所述关系数据存储在数据库中。

上述基金数据关联方法、装置、计算机设备和存储介质，包括定时从多个基金网站采集基金基础信息，并将基金基础信息存储在缓存器中；调用存储在
20 缓存器中的基金基础信息，通过预设的数据筛选规则，对基金基础信息进行筛选，得到多项基金数据集，每项基金数据集内的基金数据具有关联关系，将多项基金数据集存储在缓冲器中；依次调用存储在缓存器中的每项基金数据集，对同一项基金数据集生成多个实体，每个实体对应生成多个实体类型，实体包括基金、基金经理或基金公司，实体类型包括实体关联的其他基金数据；对每
25 两个实体进行匹配，当具有相同的实体类型时，将实体类型设置为关系类型，以关系类型将两个实体进行关联，将两个实体和实体类型生成有关联关系的关系数据，将关系数据存储在数据库中。本申请通过从各个来源定时获取基金基础信息，进行深度挖掘，通过构建各个基金数据之间的关系生成有关联关系的关系数据，供用户后续通过某一基金数据，可以查找到与其关联关系紧密的其他基金数据，实现了对海量公开数据的有效提取和关联。
30

附图说明

通过阅读下文优选实施方式的详细描述，各种其他的优点和益处对于本领域普通技术人员将变得清楚明了。附图仅用于示出优选实施方式的目的，而并不认为是对本申请的限制。

- 5 图 1 为本申请一个实施例中的基金数据关联方法的流程图；
图 2 为一个实施例中步骤 S3 的流程图；
图 3 为一个实施例中两个实体之间的关联关系图；
图 4 为本申请一个实施例中基金数据关联系统的结构图。

10 具体实施方式

为了使本申请的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本申请进行进一步详细说明。应当理解，此处所描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

- 15 本技术领域技术人员可以理解，除非特意声明，这里使用的单数形式“一”、“一个”、“所述”和“该”也可包括复数形式。应该进一步理解的是，本申请的说明书中使用的措辞“包括”是指存在所述特征、整数、步骤、操作、元件和/或组件，但是并不排除存在或添加一个或多个其他特征、整数、步骤、操作、元件、组件和/或它们的组。

- 20 图 1 为本申请一个实施例中的基金数据关联方法的流程图，如图 1 所示，一种基金数据关联方法，包括以下步骤：

步骤 S1，采集基金基础信息：定时从多个基金网站采集基金基础信息，并将基金基础信息存储在缓存器中。

- 25 为了丰富本申请的基金数据，本步骤从不同的基金网站上分别进行采集，得到海量的基金基础信息。基金基础信息的来源可以是外部如：公开的基金公司官方网站、监管机构披露数据、中国知网、国内外学术论文网站、新浪微博、各类投资推荐网站，还可以是天眼查等 APP 渠道。此时可以通过网络爬虫算法定期或适时进行爬取，获取大量的基金基础信息，并将基金基础信息存储在缓存器中。

在一个实施例中，通过网络爬虫算法获取基金基础信息包括：

- 30 预设网址列表，网址列表中包含多个基金网站的网址；调用浏览器内核依次对网址列表中的网址发出网页访问请求，并等待接收网页访问请求的网站发

出的反馈信息，反馈信息包括接收访问的反馈信息和拒绝接收访问的反馈信息；当接收到接收访问的反馈信息时，调用预设 in 数据库中的网络爬虫算法，采集与基金相关的基金基础信息，然后继续调用浏览器内核访问网址列表中的其他网址，直到遍历网址列表中的所有网址；当接收到拒绝接收访问的反馈信息后，
5 继续调用浏览器内核访问网址列表中的其他网址，直到遍历网址列表中的所有网址；汇总网络爬虫算法采集到的基金基础信息。

本实施例利用网络爬虫算法爬取数据，可以实现自动操作，不需要人工筛选信息，方便且不会遗漏关键信息，能容易得到海量基金基础信息。

本步骤中基金网站的数据库还可以是本地数据库，基金基础信息的来源
10 如平安欧拉图谱数据服务网站、平安保险网站、陆金所网站，这些网站的数据库是本地数据库，可以从本地数据库中获取海量基金基础信息。

在一个实施例中，定时从数据库中采集基金基础信息包括：

预设数据库的目标路径列表，目标路径列表中含有提供基金基础信息的多个路径，每个路径对应数据库中的至少一个文件；调用定时任务，依次读取目
15 标路径列表中的路径，根据路径在数据库中查找对应的文件，采集文件内的基金基础信息。

本实施例利用本地数据库丰富的数据，通过定时任务，定时采集基金基础信息，为后续基金数据关联提供更可靠的数据。

步骤 S2，筛选数据：调用存储在缓存器中的基金基础信息，通过预设的数
20 据筛选规则，对基金基础信息进行筛选，得到多项基金数据集，每项基金数据集内的基金数据具有关联关系，将多项基金数据集存储在缓冲器中。

本步骤中的数据筛选规则是从基金基础信息中，将包括基金经理履历信息及对应的基金公司、管理过的基金、管理过的基金公司、管理的基金的持仓信息、社会关系和新闻媒体信息在内的关键字进行提取的规则。

其中，基金经理履历信息包括基金经理的姓名、婚姻状况、毕业院校、导师、同学等信息。社会关系包括父母、兄弟、姐妹、子女和配偶等亲属信息。
25 新闻媒体信息包括基金经理有关系的公开新闻、报道等信息。

在一个实施例中，对基金基础信息进行筛选，得到多项基金数据集，包括：

调用数据筛选规则采用自然语言处理或正则表达式对基金基础信息进行提
30 取，提取出多个基金数据，将每个基金经理及关联关系的基金数据汇总成一项基金数据集，多个基金经理生成多项基金数据集，关联关系包括基金经理、基

基金公司、管理过的基金、管理过的基金公司、毕业院校、导师、同学、配偶中的至少一项关系。

自然语言处理技术是通过软件来对文本语言以人类的逻辑思维进行处理，根据数据筛选规则，对基金基础信息提取，得到多个基金数据。

- 5 正则表达式是一种文本模式，是对字符串和特殊字符操作的一种逻辑公式，通常用事先定义好的一些特定字符、及这些特定字符的组合，组成一个规则字符串，规则字符串用来表达对字符串的一种过滤逻辑。当基金基础信息是结构化数据时，通过正则表达式对基金基础信息进行提取，提取出多个基金数据。

本实施例中的多项基金数据集可以采用如下表 1 中所示：

关联关系 基金数据集	基金公司	基金	毕业院校	导师	同学	配偶
基金经理 1	嘉实基金 管理有限公司	嘉实基 本面 50	上海 XX 大学	导师 A	同学 B	配偶 C
基金经理 2	嘉实基金 管理有限公司	嘉实基 本面 50	北京 XX 大学	导师 D	同学 E	配偶 F

表 1

本实施例通过不同手段对基金基础信息进行处理，获取到各个基金相关的核心数据，这些核心数据用以后续关联使用，能将各个基金经理之间的各种关系串联起来，挖掘基金的关联关系，发现关联最紧密的基金，为用户提供可靠的投资数据。

- 15 步骤 S3，生成实体及实体类型：依次调用存储在缓存器中的每项基金数据集，对同一项基金数据集生成多个实体，每个实体对应生成多个实体类型，实体包括基金、基金经理或基金公司，实体类型包括实体关联的其他基金数据。

在一个实施例中，如图 2 所示，在对同一项基金数据集生成多个实体，每个实体对应生成多个实体类型，包括：

- 20 步骤 S301，标记唯一代码：对基金数据集中的基金经理、基金公司和基金分别标记唯一代码，并将唯一代码及对应的基金数据保存在目标关键词文本中。

由于基金经理、基金公司和基金在金融领域中，对于基金投资数据举足轻重，而且相关的基金经理之间往往存在抱团倾向，因此本步骤将基金数据集中

所有的基金经理、基金公司和基金分别标记唯一代码，用于后续实体的生成，以基金经理、基金公司和基金为主线，展现某一实体的关联关系。

目标关键词文本可以采用如表 2 所示的内容：

基金数据集	唯一代码	基金数据
基金	160716	嘉实基本面 50
基金经理	30284601	基金经理 1
	30414880	基金经理 2
基金公司	80000223	嘉实基金管理有限公司

表 2

- 5 步骤 S302，生成实体：获取目标关键词文本中的唯一代码及唯一代码对应的基金数据，将基金数据设置为名称，生成一个实体，实体包括唯一代码及名称。

本步骤中，将目标关键词文本中所有的基金经理、基金公司和基金都各自生成一个实体，实体包括唯一代码和名称，具体的，如下表 3、表 4 和表 5 所示，
10 分别是基金、基金经理和基金公司三种实体。

基金
-唯一代码：160716 -名称：嘉实基本面 50

表 3

如上表所示，生成一个实体，实体包括唯一代码为 160716、名称为嘉实基本面 50 的基金实体。

基金经理
-唯一代码：30284601 -名称：基金经理 1

表 4

- 15 如上表所示，生成一个实体，实体包括唯一代码为 30284601、名称为基金经理 1 的基金经理实体。

基金公司
-唯一代码：80000223 -名称：嘉实基金管理有限公司

表 5

如上表所示，生成一个实体，实体包括唯一代码为 80000223、名称为嘉实基金管理有限公司的基金公司实体。

步骤 S303，生成实体类型：获取基金数据集中实体关联的其他基金数据，将实体与其他基金数据关联关系设置为节点类型，将其他基金数据设置为实体名称，生成一个实体类型，实体类型包括节点类型及实体名称，实体还包括节点类型，以节点类型将实体和实体类型进行关联。

由于同一个基金数据集中包括了基金经理，及基金经理具有对应关系的其他基金数据，因此当在同一个基金数据集中生成若干实体后，每个实体与其他基金数据的关系从实体类型进行体现和关联。比如，步骤 S302 生成了一个如表 4 所示的唯一代码为 30284601、名称为基金经理 1 的基金经理实体。则此基金经理 1 在基金数据集中相关联的其他基金数据如表 1 所示，节点类型包括基金公司、基金、毕业院校、导师、同学及配偶，则此基金经理实体关联了六个实体类型。以基金公司为例，将关联关系基金公司设置为节点类型，将嘉实基金管理有限公司设置为实体名称，生成节类型为基金公司、实体名称为嘉实基金管理有限公司的实体类型。基金经理实体与此实体类型通过基金公司进行关联。

实体类型还包括权重系数，生成一个实体类型时，还包括：调用预设的权重列表，权重列表中包含有节点类型及对应的权重系数，根据节点类型，得到权重系数，将节点类型、实体名称及权重系数生成实体类型。由于据有关联关系的实体之间往往具有抱团倾向，所以本申请采用有权重的关系，即实体类型是有关重的，以便于能后续关联实体时，确定两个实体之间的重要程度。具体的，节点类型为基金公司，则权重 $W=1$ ，节点类型为基金，则权重 $W=2$ ，节点类型为毕业院校，则权重 $W=1$ ，节点类型为导师，则权重 $W=2$ ，节点类型为同学，则权重 $W=1$ ，节点类型为配偶，则权重 $W=2$ 。

具体的，表 1 中的基金经理 1 通过步骤 S302 和步骤 S303 生成的实体及其多项实体类型如表 6 所示。表 1 中的基金经理 2 通过步骤 S302 和步骤 S303 生成的实体及其中多项实体类型如表 7 所示。

实体	实体类型	实体类型
-唯一代码：30284601 -名称：基金经理 1 -节点类型：基金公司 -节点类型：毕业院校	-节点类型：基金公司 -实体名称：嘉实基金管理有 限公司 -权重：1	-节点类型：毕业院校 -实体名称：上海 XX 大学 -权重：1

表 6

实体	实体类型	实体类型
-唯一代码: 30414880 -名称: 基金经理 2 -节点类型: 基金公司 -节点类型: 同学	-节点类型: 基金公司 -实体名称: 嘉实基金管理有 限公司 -权重: 1	-节点类型: 同学 -实体名称: 同学 E -权重: 1

表 7

本实施例，通过标记唯一代码区别基金数据集中众多的基金数据，使基金数据集中最重要的基金经理、基金公司和基金具有唯一性，为后续生成实体和实体类型提供数据支持。在生成实体和实体类型时，对同一基金数据集中分别生成多个实体，每个实体生成多个与之关联的实体类型，尽可能的将基金数据集中的含有关系的联系在一起，为后续生成关系数据提供链接关系。

步骤 S4，关联实体之间的关系：对每两个实体进行匹配，当具有相同的实体类型时，将实体类型设置为关系类型，以关系类型将两个实体进行关联，将两个实体和实体类型生成有关联关系的关系数据，将关系数据存储在数据库中。

在一个实施例中，对每两个实体进行匹配，当具有相同的实体类型时，将实体类型设置为关系类型，并生成关系链接，关系链接包括关系类型中的节点类型、链接来源和链接目标，将其中一个实体的唯一代码设置为链接来源，另一个实体的唯一代码设置为链接目标；将关系链接与链接来源进行关联，将链接来源、关系链接和链接目标生成有关联关系的关系数据。

具体的，步骤 S3 的表 6 和表 7 生成了两个实体，对此两个实体进行匹配，发现具有相同的实体类型，将实体类型设置为关系类型，还生成了一个关系链接，关系链接中的节点类型为基金公司，链接来源为 30284601，链接目标为 30414880。如图 3 所示，唯一代码为 30284601 的基金经理实体通过链接来源与关系链接关联，通过关系链接与唯一代码为 30414880 的链接目标生成关联关系，他们的关联关系是关系类型中的关系。且唯一代码为 30284601 的基金经理实体和唯一代码为 30414880 的基金经理实体分别各自对应有实体类型。

本实施例通过增设关系链接，通过链接来源和链接目标将两个有关系的实体进行连接，实现关联，生成关系数据。这种方法，可以将某一实体的其他实体类型相区别，没有关系的实体类型分别与各种的实体相关联，整个关联关系简洁清楚，关系数据易于保存和理解。

本申请的基金数据关联方法，通过定时从外部多个渠道通过网络爬虫算法或查找数据库的方式，采集大量的基金基础信息，丰富了基金数据，减少遗漏基金相关信息的可能，且由于本申请是定时采集，因此得到的基金基础信息可做到实时更新。在得到大量基金基础信息后，通过预设的数据筛选规则进行筛选，得到与基金密切相关的基金数据，如基金经理履历信息及对应的基金公司、管理过的基金、管理过的基金公司、管理的基金的持仓信息、社会关系和新闻媒体信息等。由于基金数据量大，因此通过基金经理为标识，将与某一基金经理相关联的基金数据整合在一起生成基金数据集。并依次对每项基金数据集中的基金数据生成实体和实体类型，并以基金经理、基金公司和基金为主线，将各个实体间的连接关系通过关系链接进行关联，得到多项关系数据，实现对海量基金相关信息的梳理、整合和关联目的，用户可以从纷繁复杂的海量基金相关信息中得到关联关系最紧密的基金相关信息。

在一个实施例中，提出了一种基金数据关联系统，如图 4 所示，包括如下单元：

采集单元，设置为定时从多个基金网站采集基金基础信息，并将基金基础信息存储在缓存器中；

筛选单元，设置为调用存储在缓存器中的基金基础信息，通过预设的数据筛选规则，对基金基础信息进行筛选，得到多项基金数据集，每项基金数据集内的基金数据具有关联关系，将多项基金数据集存储在缓冲器中；

生成实体及实体类型单元，设置为依次调用存储在缓存器中的每项基金数据集，对同一项基金数据集生成多个实体，每个实体对应生成多个实体类型，实体包括基金、基金经理或基金公司，实体类型包括实体关联的其他基金数据；

关联单元，设置为对每两个实体进行匹配，当具有相同的实体类型时，将实体类型设置为关系类型，以关系类型将两个实体进行关联，将两个实体和实体类型生成有关联关系的关系数据，将关系数据存储在数据库中。

在一个实施例中，采集单元，包括：预设模块，设置为预设网址列表，网址列表中包含多个基金网站的网址；请求模块，设置为调用浏览器内核依次对网址列表中的网址发出网页访问请求，并等待接收网页访问请求的网站发出的反馈信息，反馈信息包括接收访问的反馈信息和拒绝接收访问的反馈信息；采集模块，设置为当接收到接收访问的反馈信息时，调用预设存储在数据库中的网络爬虫算法，采集与基金相关的基金基础信息，然后继续调用浏览器内核访问网

址列表中的其他网址，直到遍历网址列表中的所有网址；遍历模块，设置为当接收到拒绝接收访问的反馈信息后，继续调用浏览器内核访问网址列表中的其他网址，直到遍历网址列表中的所有网址；汇总模块，设置为汇总网络爬虫算法采集到的基金基础信息。

5 在一个实施例中，采集单元，包括：预设列表模块，设置为预设数据库的目标路径列表，目标路径列表中含有提供基金基础信息的多个路径，每个路径对应数据库中的至少一个文件；定时采集模块，设置为调用定时任务，依次读取目标路径列表中的路径，根据路径在数据库中查找对应的文件，采集文件内的基金基础信息。

10 在一个实施例中，筛选单元，还设置为调用数据筛选规则采用自然语言处理或正则表达式对基金基础信息进行提取，提取出多个基金数据，将每个基金经理及关联关系的基金数据汇总成一项基金数据集，多个基金经理生成多项基金数据集，关联关系包括基金经理、基金公司、管理过的基金、管理过的基金公司、毕业院校、导师、同学、配偶中的至少一项关系；数据筛选规则是从基金基础信息中，将包括基金经理履历信息及对应的基金公司、管理过的基金、管理过的基金公司、管理的基金的持仓信息、社会关系和新闻媒体信息在内的关键字进行提取的规则。

15 在一个实施例中，生成实体及实体类型单元，包括：标记模块，设置为对基金数据集中的基金经理、基金公司和基金分别标记唯一代码，并将唯一代码及对应的基金数据保存在目标关键词文本中；生成实体模块，设置为获取目标关键词文本中的唯一代码及唯一代码对应的基金数据，将基金数据设置为名称，生成一个实体，实体包括唯一代码及名称；生成实体类型模块，设置为获取基金数据集中实体关联的其他基金数据，将实体与其他基金数据关联关系设置为节点类型，将其他基金数据设置为实体名称，生成一个实体类型，实体类型包括节点类型及实体名称，实体还包括节点类型，以节点类型将实体和实体类型进行关联。

20 在一个实施例中，关联模块，还设置为调用预设的权重列表，权重列表中包含有节点类型及对应的权重系数，根据节点类型，得到权重系数，将节点类型、实体名称及权重系数生成实体类型。

25 在一个实施例中，关联单元，包括：匹配模块，设置为对每两个实体进行匹配，当具有相同的实体类型时，将实体类型设置为关系类型，并生成关系链

接，关系链接包括关系类型中的节点类型、链接来源和链接目标，将其中一个实体的唯一代码设置为链接来源，另一个实体的唯一代码设置为链接目标；生成关系数据模块，设置为将关系链接与链接来源进行关联，将链接来源、关系链接和链接目标生成有关联关系的关系数据。

- 5 在一个实施例中，提出了一种计算机设备，包括存储器和处理器，存储器中存储有计算机可读指令，计算机可读指令被处理器执行时，使得处理器执行计算机可读指令时实现上述各实施例中基金数据关联方法中的步骤。

10 在一个实施例中，提出了一种存储有计算机可读指令的存储介质，计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行上述各实施例中基金数据关联方法中的步骤。其中，存储介质可以为非易失性存储介质。

本领域普通技术人员可以理解上述实施例的各种方法中的全部或部分步骤是可以通过程序来指令相关的硬件来完成，该程序可以存储于一计算机可读存储介质中，存储介质可以包括：只读存储器（ROM，Read Only Memory）、随机存取存储器（RAM，Random Access Memory）、磁盘或光盘等。

- 15 以上所述实施例的各技术特征可以进行任意的组合，为使描述简洁，未对上述实施例中的各个技术特征所有可能的组合都进行描述，然而，只要这些技术特征的组合不存在矛盾，都应当认为是本说明书记载的范围。

20 以上所述实施例仅表达了本申请一些示例性实施例，其描述较为具体和详细，但并不能因此而理解为对本申请专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本申请构思的前提下，还可以做出若干变形和改进，这些都属于本申请的保护范围。因此，本申请专利的保护范围应以所附权利要求为准。

权 利 要 求 书

1、一种基金数据关联方法，包括：

定时从多个基金网站采集基金基础信息，并将所述基金基础信息存储在缓存器中；

调用存储在所述缓存器中的所述基金基础信息，通过预设的数据筛选规则，
5 对所述基金基础信息进行筛选，得到多项基金数据集，每项所述基金数据集内的基金数据具有关联关系，将多项所述基金数据集存储在所述缓冲器中；

依次调用存储在所述缓存器中的每项所述基金数据集，对同一项所述基金数据集生成多个实体，每个实体对应生成多个实体类型，所述实体包括基金、基金经理或基金公司，所述实体类型包括实体关联的其他所述基金数据；

10 对每两个所述实体进行匹配，当具有相同的所述实体类型时，将所述实体类型设置为关系类型，以关系类型将两个所述实体进行关联，将两个所述实体和所述实体类型生成有关联关系的关系数据，将所述关系数据存储在数据库中。

2、根据权利要求 1 所述的基金数据关联方法，其中，所述定时从多个基金网站采集基金基础信息，并将所述基金基础信息存储在缓存器中，包括：

15 预设网址列表，所述网址列表中包含多个所述基金网站的网址；

调用浏览器内核依次对所述网址列表中的网址发出所述网页访问请求，并等待接收所述网页访问请求的网站发出的反馈信息，所述反馈信息包括接收访问的反馈信息和拒绝接收访问的反馈信息；

当接收到所述接收访问的反馈信息时，调用预设在所述数据库中的网络爬虫算法，采集与基金相关的所述基金基础信息，然后继续调用所述浏览器内核访问所述网址列表中的其他网址，直到遍历所述网址列表中的所有网址；

当接收到所述拒绝接收访问的反馈信息后，继续调用所述浏览器内核访问所述网址列表中的其他网址，直到遍历所述网址列表中的所有网址；

汇总所述网络爬虫算法采集到的所述基金基础信息。

25 3、根据权利要求 1 所述的基金数据关联方法，其中，所述基金网站的数据库是本地数据库时，还包括定时从所述数据库中采集基金基础信息；

预设所述数据库的目标路径列表，所述目标路径列表中含有提供所述基金基础信息的多个路径，每个所述路径对应所述数据库中的至少一个文件；

调用定时任务，依次读取所述目标路径列表中的路径，根据路径在数据库
30 中查找对应的所述文件，采集所述文件内的基金基础信息。

4、根据权利要求 1 所述的基金数据关联方法，其中，所述调用存储在所述

缓存器中的所述基金基础信息，通过预设的数据筛选规则，对所述基金基础信息进行筛选，得到多项基金数据集，包括：

所述数据筛选规则是从所述基金基础信息中，将包括基金经理履历信息及对应的基金公司、管理过的基金、管理过的基金公司、管理的基金的持仓信息、
5 社会关系和新闻媒体信息在内的关键字进行提取的规则；

调用所述数据筛选规则采用自然语言处理或正则表达式对所述基金基础信息进行提取，提取出多个基金数据，将每个基金经理及关联关系的基金数据汇总成一项所述基金数据集，多个所述基金经理生成多项所述基金数据集，所述关联关系包括基金经理、基金公司、管理过的基金、管理过的基金公司、毕业
10 院校、导师、同学、配偶中的至少一项关系。

5、根据权利要求 1 所述的基金数据关联方法，其中，所述对同一项所述基金数据集生成多个实体，每个实体对应生成多个实体类型，包括：

对所述基金数据集中的基金经理、基金公司和基金分别标记唯一代码，并将所述唯一代码及对应的基金数据保存在目标关键词文本中；

15 获取所述目标关键词文本中的所述唯一代码及所述唯一代码对应的基金数据，将所述基金数据设置为名称，生成一个所述实体，所述实体包括唯一代码及名称；

获取所述基金数据集中所述实体关联的其他所述基金数据，将所述实体与其他所述基金数据关联关系设置为节点类型，将其他所述基金数据设置为实体
20 名称，生成一个所述实体类型，所述实体类型包括节点类型及实体名称，所述实体还包括所述节点类型，以所述节点类型将所述实体和所述实体类型进行关联。

6、根据权利要求 5 所述的基金数据关联方法，其中，所述实体类型还包括权重系数，所述生成一个所述实体类型时，还包括：

25 调用预设的权重列表，所述权重列表中包含有节点类型及对应的权重系数，根据所述节点类型，得到所述权重系数，将所述节点类型、所述实体名称及所述权重系数生成所述实体类型。

7、根据权利要求 5 所述的基金数据关联方法，其中，所述对每两个所述实体进行匹配，当具有相同的所述实体类型时，将所述实体类型设置为关系类型，
30 以关系类型将两个所述实体进行关联，将两个所述实体和所述实体类型生成有
关联关系的关系数据，包括：

对每两个所述实体进行匹配，当具有相同的所述实体类型时，将所述实体类型设置为所述关系类型，并生成关系链接，所述关系链接包括所述关系类型中的节点类型、链接来源和链接目标，将其中一个所述实体的唯一代码设置为链接来源，另一个所述实体的唯一代码设置为链接目标；

- 5 将所述关系链接与所述链接来源进行关联，将所述链接来源、所述关系链接和所述链接目标生成有关联关系的关系数据。

8、一种基金数据关联系统，包括：

采集单元，设置为定时从多个基金网站采集基金基础信息，并将所述基金基础信息存储在缓存器中；

- 10 筛选单元，设置为调用存储在所述缓存器中的所述基金基础信息，通过预设的数据筛选规则，对所述基金基础信息进行筛选，得到多项基金数据集，每项所述基金数据集内的基金数据具有关联关系，将多项所述基金数据集存储在所述缓冲器中；

- 15 生成实体及实体类型单元，设置为依次调用存储在所述缓存器中的每项所述基金数据集，对同一项所述基金数据集生成多个实体，每个实体对应生成多个实体类型，所述实体包括基金、基金经理或基金公司，所述实体类型包括实体关联的其他所述基金数据；

- 20 关联单元，设置为对每两个所述实体进行匹配，当具有相同的所述实体类型时，将所述实体类型设置为关系类型，以关系类型将两个所述实体进行关联，将两个所述实体和所述实体类型生成有关联关系的关系数据，将所述关系数据存储在数据库中。

9、根据权利要求8所述的基金数据关联系统，其中，所述采集单元，包括：

预设模块，设置为预设网址列表，所述网址列表中包含多个所述基金网站的网址；

- 25 请求模块，设置为调用浏览器内核依次对所述网址列表中的网址发出所述网页访问请求，并等待接收所述网页访问请求的网站发出的反馈信息，所述反馈信息包括接收访问的反馈信息和拒绝接收访问的反馈信息；

- 30 采集模块，设置为当接收到所述接收访问的反馈信息时，调用预设在该数据库中的网络爬虫算法，采集与基金相关的所述基金基础信息，然后继续调用所述浏览器内核访问所述网址列表中的其他网址，直到遍历所述网址列表中的所有网址；

遍历模块，设置为当接收到所述拒绝接收访问的反馈信息后，继续调用所述浏览器内核访问所述网址列表中的其他网址，直到遍历所述网址列表中的所有网址；

汇总模块，设置为汇总所述网络爬虫算法采集到的所述基金基础信息。

5 10、根据权利要求 8 所述的基金数据关联系统，其中，所述采集单元，包括：

预设列表模块，设置为预设所述数据库的目标路径列表，所述目标路径列表中含有提供所述基金基础信息的多个路径，每个所述路径对应所述数据库中的至少一个文件；

10 定时采集模块，设置为调用定时任务，依次读取所述目标路径列表中的路径，根据路径在数据库中查找对应的所述文件，采集所述文件内的基金基础信息。

11、根据权利要求 8 所述的基金数据关联系统，其中，所述筛选单元，还设置为调用所述数据筛选规则采用自然语言处理或正则表达式对所述基金基础
15 信息进行提取，提取出多个基金数据，将每个基金经理及关联关系的基金数据汇总成一项所述基金数据集，多个所述基金经理生成多项所述基金数据集，所述关联关系包括基金经理、基金公司、管理过的基金、管理过的基金公司、毕业院校、导师、同学、配偶中的至少一项关系；

所述数据筛选规则是从所述基金基础信息中，将包括基金经理履历信息及
20 对应的基金公司、管理过的基金、管理过的基金公司、管理的基金的持仓信息、社会关系和新闻媒体信息在内的关键字进行提取的规则。

12、根据权利要求 8 所述的基金数据关联系统，其中，所述生成实体及实体类型单元，包括：

标记模块，设置为对所述基金数据集中的基金经理、基金公司和基金分别
25 标记唯一代码，并将所述唯一代码及对应的基金数据保存在目标关键词文本中；

生成实体模块，设置为获取所述目标关键词文本中的所述唯一代码及所述唯一代码对应的基金数据，将所述基金数据设置为名称，生成一个所述实体，所述实体包括唯一代码及名称；

生成实体类型模块，设置为获取所述基金数据集中所述实体关联的其他所
30 述基金数据，将所述实体与其他所述基金数据关联关系设置为节点类型，将其他所述基金数据设置为实体名称，生成一个所述实体类型，所述实体类型包括

节点类型及实体名称，所述实体还包括所述节点类型，以所述节点类型将所述实体和所述实体类型进行关联。

13、根据权利要求 12 所述的基金数据关联系统，其中，所述关联模块，还设置为调用预设的权重列表，所述权重列表中包含有节点类型及对应的权重系数，根据所述节点类型，得到所述权重系数，将所述节点类型、所述实体名称及所述权重系数生成所述实体类型。

14、根据权利要求 12 所述的基金数据关联系统，其中，所述关联单元，包括：

匹配模块，设置为对每两个所述实体进行匹配，当具有相同的所述实体类型时，将所述实体类型设置为所述关系类型，并生成关系链接，所述关系链接包括所述关系类型中的节点类型、链接来源和链接目标，将其中一个所述实体的唯一代码设置为链接来源，另一个所述实体的唯一代码设置为链接目标；

生成关系数据模块，设置为将所述关系链接与所述链接来源进行关联，将所述链接来源、所述关系链接和所述链接目标生成有关联关系的关系数据。

15、一种计算机设备，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述处理器执行以下步骤：

定时从多个基金网站采集基金基础信息，并将所述基金基础信息存储在缓存器中；

调用存储在所述缓存器中的所述基金基础信息，通过预设的数据筛选规则，对所述基金基础信息进行筛选，得到多项基金数据集，每项所述基金数据集内的基金数据具有关联关系，将多项所述基金数据集存储在所述缓冲器中；

依次调用存储在所述缓存器中的每项所述基金数据集，对同一项所述基金数据集生成多个实体，每个实体对应生成多个实体类型，所述实体包括基金、基金经理或基金公司，所述实体类型包括实体关联的其他所述基金数据；

对每两个所述实体进行匹配，当具有相同的所述实体类型时，将所述实体类型设置为关系类型，以关系类型将两个所述实体进行关联，将两个所述实体和所述实体类型生成有关联关系的关系数据，将所述关系数据存储在数据库中。

16、根据权利要求 15 所述的计算机设备，其中，所述调用存储在所述缓存器中的所述基金基础信息，通过预设的数据筛选规则，对所述基金基础信息进行筛选，得到多项基金数据集时，使得所述处理器执行以下步骤：

所述数据筛选规则是从所述基金基础信息中，将包括基金经理履历信息及对应的基金公司、管理过的基金、管理过的基金公司、管理的基金的持仓信息、社会关系和新闻媒体信息在内的关键字进行提取的规则；

调用所述数据筛选规则采用自然语言处理或正则表达式对所述基金基础信息
5 息进行提取，提取出多个基金数据，将每个基金经理及关联关系的基金数据汇总成一项所述基金数据集，多个所述基金经理生成多项所述基金数据集，所述关联关系包括基金经理、基金公司、管理过的基金、管理过的基金公司、毕业院校、导师、同学、配偶中的至少一项关系。

17、根据权利要求 15 所述的计算机设备，其中，所述对同一项所述基金数
10 据集生成多个实体，每个实体对应生成多个实体类型时，使得所述处理器执行以下步骤：

对所述基金数据集中的基金经理、基金公司和基金分别标记唯一代码，并将所述唯一代码及对应的基金数据保存在目标关键词文本中；

获取所述目标关键词文本中的所述唯一代码及所述唯一代码对应的基金数
15 据，将所述基金数据设置为名称，生成一个所述实体，所述实体包括唯一代码及名称；

获取所述基金数据集中所述实体关联的其他所述基金数据，将所述实体与其他所述基金数据关联关系设置为节点类型，将其他所述基金数据设置为实体名称，生成一个所述实体类型，所述实体类型包括节点类型及实体名称，所述
20 实体还包括所述节点类型，以所述节点类型将所述实体和所述实体类型进行关联。

18、一种存储有计算机可读指令的存储介质，所述计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行以下步骤：

定时从多个基金网站采集基金基础信息，并将所述基金基础信息存储在缓
25 存器中；

调用存储在所述缓存器中的所述基金基础信息，通过预设的数据筛选规则，对所述基金基础信息进行筛选，得到多项基金数据集，每项所述基金数据集内的基金数据具有关联关系，将多项所述基金数据集存储在所述缓冲器中；

依次调用存储在所述缓存器中的每项所述基金数据集，对同一项所述基金
30 数据集生成多个实体，每个实体对应生成多个实体类型，所述实体包括基金、基金经理或基金公司，所述实体类型包括实体关联的其他所述基金数据；

对每两个所述实体进行匹配，当具有相同的所述实体类型时，将所述实体类型设置为关系类型，以关系类型将两个所述实体进行关联，将两个所述实体和所述实体类型生成有关联关系的关系数据，将所述关系数据存储在数据库中。

19、根据权利要求 18 所述的存储介质，其中，所述调用存储在所述缓存器中的所述基金基础信息，通过预设的数据筛选规则，对所述基金基础信息进行筛选，得到多项基金数据集时，使得一个或多个所述处理器执行以下步骤：

所述数据筛选规则是从所述基金基础信息中，将包括基金经理履历信息及对应的基金公司、管理过的基金、管理过的基金公司、管理的基金的持仓信息、社会关系和新闻媒体信息在内的关键字进行提取的规则；

10 调用所述数据筛选规则采用自然语言处理或正则表达式对所述基金基础信息进行提取，提取出多个基金数据，将每个基金经理及关联关系的基金数据汇总成一项所述基金数据集，多个所述基金经理生成多项所述基金数据集，所述关联关系包括基金经理、基金公司、管理过的基金、管理过的基金公司、毕业院校、导师、同学、配偶中的至少一项关系。

15 20、根据权利要求 18 所述的存储介质，其中，所述对同一项所述基金数据集生成多个实体，每个实体对应生成多个实体类型时，使得一个或多个所述处理器执行以下步骤：

对所述基金数据集中的基金经理、基金公司和基金分别标记唯一代码，并将所述唯一代码及对应的基金数据保存在目标关键词文本中；

20 获取所述目标关键词文本中的所述唯一代码及所述唯一代码对应的基金数据，将所述基金数据设置为名称，生成一个所述实体，所述实体包括唯一代码及名称；

25 获取所述基金数据集中所述实体关联的其他所述基金数据，将所述实体与其他所述基金数据关联关系设置为节点类型，将其他所述基金数据设置为实体名称，生成一个所述实体类型，所述实体类型包括节点类型及实体名称，所述实体还包括所述节点类型，以所述节点类型将所述实体和所述实体类型进行关联。

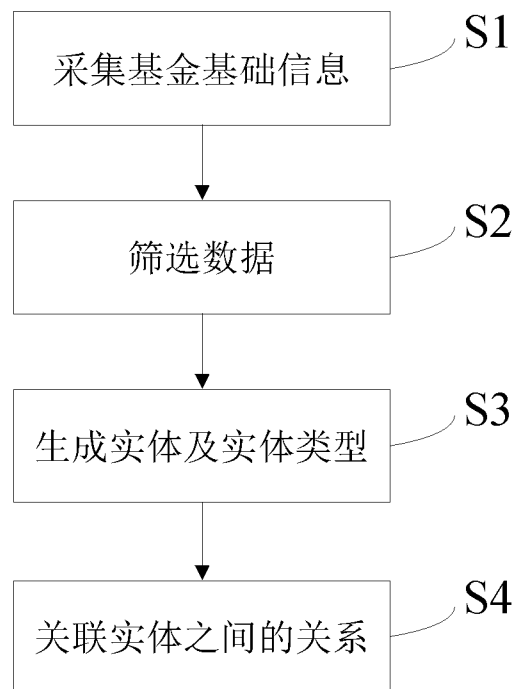


图 1

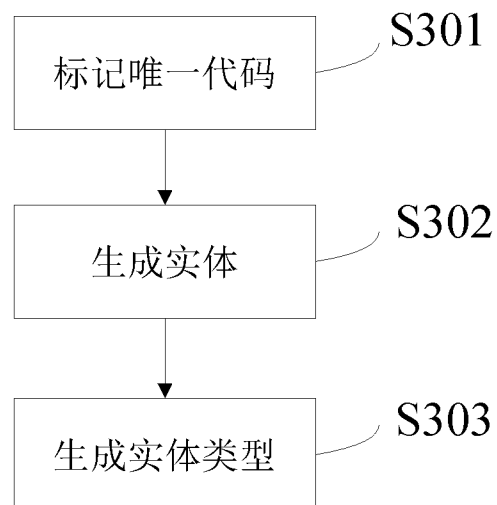


图 2

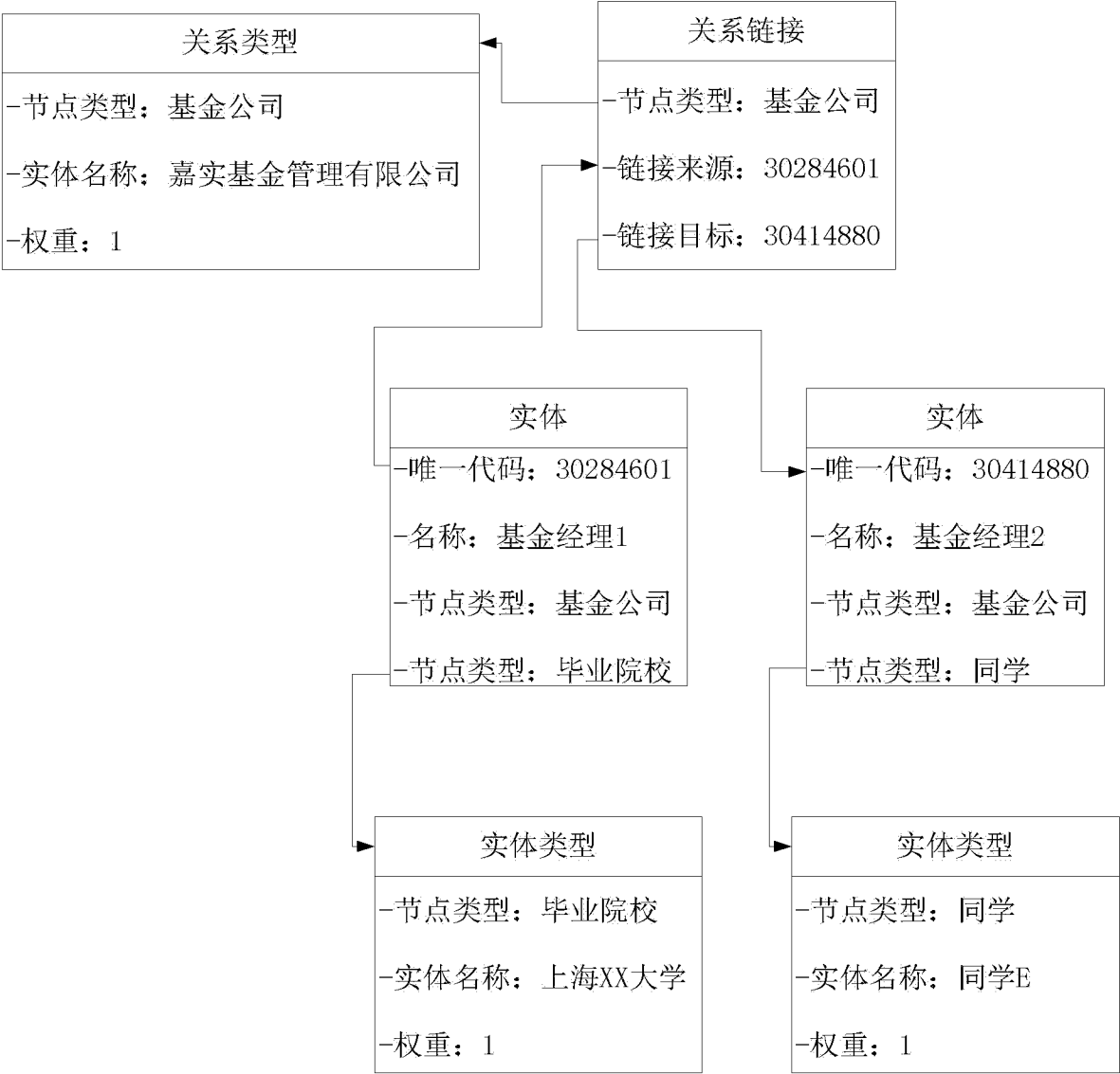


图 3

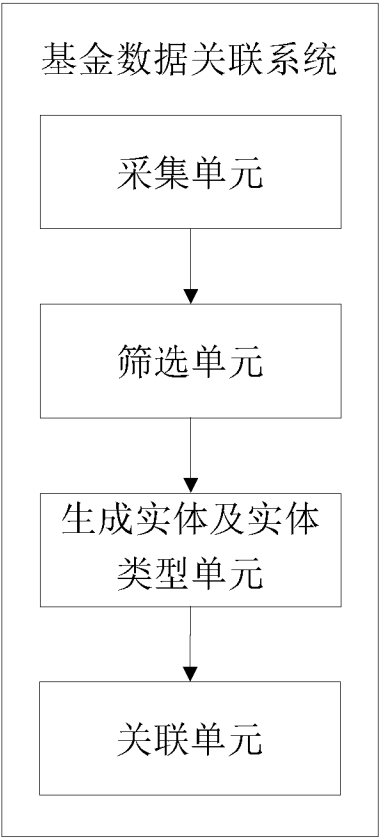


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/124252

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/9535(2019.01)i; G06Q 40/06(2012.01)n

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F, G06Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNPAT, CNKI, IEEE, GOOGLE: 基金, 股票, 金融产品, 关联, 相关, 关系, 联系, 类型, 实体, 匹配, 采集, 抓取, 数据, fund, stock, financial, product, associat+, correlation, relation, connect+, type, entity, match+, collect+, gather+, grasp+, data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 106294588 A (GUANGDONG INTELLIGENT DATA TECHNOLOGY CO., LTD.) 04 January 2017 (2017-01-04) claims 1-8, description, paragraphs [0049]-[0059], and figures 1-8	1-20
A	CN 107369091 A (ALIBABA GROUP HOLDING LIMITED) 21 November 2017 (2017-11-21) entire document	1-20
A	CN 107194754 A (MERIT DATA CO., LTD.) 22 September 2017 (2017-09-22) entire document	1-20
A	CN 103064837 A (XI'AN UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) 24 April 2013 (2013-04-24) entire document	1-20

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

23 May 2019

Date of mailing of the international search report

12 June 2019

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/124252

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	106294588	A	04 January 2017	None	
CN	107369091	A	21 November 2017	None	
CN	107194754	A	22 September 2017	None	
CN	103064837	A	24 April 2013	None	

国际检索报告

国际申请号

PCT/CN2018/124252

A. 主题的分类

G06F 16/9535 (2019.01)i; G06Q 40/06 (2012.01)n

按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献 (标明分类系统和分类号)

G06F, G06Q

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用))

WPI, EPDOC, CNPAT, CNKI, IEEE, GOOGLE: 基金, 股票, 金融产品, 关联, 相关, 关系, 联系, 类型, 实体, 匹配, 采集, 抓取, 数据, fund, stock, financial, product, associat+, correlation, relation, connect+, type, entity, match+, collect+, gather+, grasp+, data

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	CN 106294588 A (广东中标数据科技股份有限公司) 2017年 1月 4日 (2017 - 01 - 04) 权利要求1-8, 说明书第[0049]-[0059]段, 附图1-8	1-20
A	CN 107369091 A (阿里巴巴集团控股有限公司) 2017年 11月 21日 (2017 - 11 - 21) 全文	1-20
A	CN 107194754 A (美林数据技术股份有限公司) 2017年 9月 22日 (2017 - 09 - 22) 全文	1-20
A	CN 103064837 A (西安邮电学院) 2013年 4月 24日 (2013 - 04 - 24) 全文	1-20

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2019年 5月 23日

国际检索报告邮寄日期

2019年 6月 12日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

受权官员

李敏

传真号 (86-10)62019451

电话号码 86-(10)-53961526

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/124252

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	106294588	A	2017年 1月 4日	无	
CN	107369091	A	2017年 11月 21日	无	
CN	107194754	A	2017年 9月 22日	无	
CN	103064837	A	2013年 4月 24日	无	