

# (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



(10) 国际公布号

WO 2019/095572 A1

(43) 国际公布日

2019 年 5 月 23 日 (23.05.2019)

(51) 国际专利分类号:

G06Q 10/06 (2012.01)

(21) 国际申请号:

PCT/CN2018/076169

(22) 国际申请日:

2018 年 2 月 10 日 (10.02.2018)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

201711141730.3 2017 年 11 月 17 日 (17.11.2017) 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八卦岭工业区平安大厦六楼, Guangdong 518000 (CN)。

(72) 发明人: 汪伟 (WANG, Wei); 中国广东省深圳市

福田区八卦岭工业区平安大厦六楼, Guangdong

518000 (CN)。 罗傲雪 (LUO, Aoxue); 中国广东省深圳市福田区八卦岭工业区平安大厦六楼, Guangdong 518000 (CN)。 肖京 (XIAO, Jing); 中国广东省深圳市福田区八卦岭工业区平安大厦六楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市沃德知识产权代理事务所 (普通合伙) (SHENZHEN WORLD INTELLECTUAL PROPERTY AGENCY (GENERAL PARTNERSHIP)); 中国广东省深圳市福田区园岭街道八卦四路 10 号中浩大厦 1528-1530 室于志光, Guangdong 518000 (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,

(54) Title: ENTERPRISE INVESTMENT RISK ASSESSMENT METHOD, DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 企业投资风险评估方法、装置及存储介质

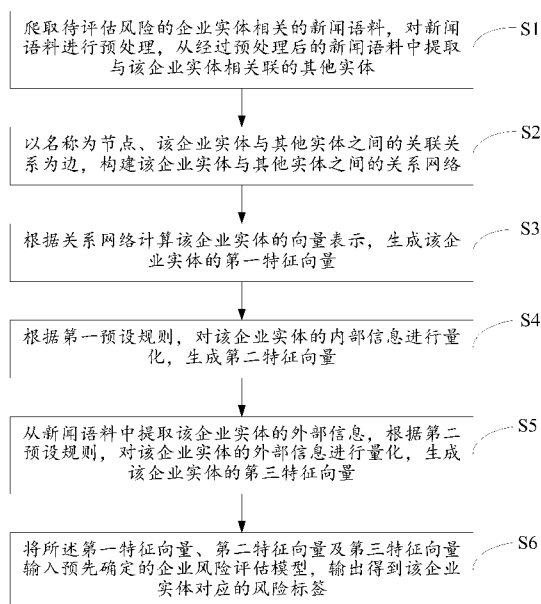


图 5

S1 Crawl news corpora associated with an enterprise entity whose risks are to be assessed, preprocess the news corpora, and extract other entities associated with the enterprise entity from the preprocessed news corpora  
S2 Build a relationship network by using names as nodes and association relationships between the enterprise entity and other entities as edges  
S3 Compute the vector representation of the enterprise entity according to the relationship network, and generate a first eigenvector of the enterprise entity  
S4 Quantize internal information of the enterprise entity according to a first preset rule, and generate a second eigenvector  
S5 Extract external information of the enterprise entity from the news corpora, quantize the external information of the enterprise entity according to a second preset rule, and generate a third eigenvector of the enterprise entity  
S6 Input the first eigenvector, the second eigenvector and the third eigenvector into a preset enterprise risk assessment model, and obtain and output a risk label corresponding to the enterprise entity

(57) Abstract: The present application provides an enterprise investment risk assessment method. The method comprises: crawling news corpora associated with an investment target enterprise entity, and extracting other entities associated with the enterprise entity; building a relationship network by using names as nodes and association relationships between the enterprise entity and other entities as edges; computing the vector representation of the enterprise entity, and generating a first eigenvector of the enterprise entity; quantizing internal information of the enterprise entity according to a first preset rule, and generating a second eigenvector; quantizing

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。

(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))。

external information of the enterprise entity according to a second preset rule, and generating a third eigenvector; and inputting the first eigenvector, the second eigenvector and the third eigenvector into an enterprise risk assessment model, and obtaining and outputting a risk label corresponding to the enterprise entity. The present application also provides an electronic device and a computer readable storage medium. By means of the present application, by analyzing information revealed in news corpora, risks of investment in target enterprises can be assessed.

(57) 摘要: 本申请提出一种企业投资风险评估方法, 包括: 爬取投资目标企业实体相关的新闻语料, 提取与该企业实体相关联的其他实体; 以名称为节点、该企业实体与其他实体之间的关联关系为边, 构建关系网络; 计算该企业实体的向量表示, 生成该企业实体的第一特征向量; 根据第一预设规则, 对该企业实体的内部信息进行量化, 生成第二特征向量; 根据第二预设规则, 对该企业实体的外部信息进行量化, 生成第三特征向量; 及, 将所述第一特征向量、第二特征向量及第三特征向量输入企业风险评估模型, 输出得到该企业实体对应的风险标签。本申请还提出一种电子装置及计算机可读存储介质。利用本申请, 对新闻语料中透露的信息进行分析, 可评估投资目标企业的风险。

## 企业投资风险评估方法、装置及存储介质

### 优先权申明

本申请基于巴黎公约申明享有 2017 年 11 月 17 日递交的申请号为 CN 201711141730.3、名称为“企业投资风险评估方法、装置及存储介质”的中国专利申请的优先权，该中国专利申请的整体内容以参考的方式结合在本申请中。

### 技术领域

本申请涉及计算机技术领域，尤其涉及一种企业投资风险评估方法、电子装置及计算机可读存储介质。

### 背景技术

目前，在观察投资标的角度上，市面上的相关工具都相对简单，大部分停留在传统的财务分析、报表分析层面，缺乏对上下游、关联方以及市场热点、政策线索的关联量化的考虑。

随着网络的普及，每个新闻网站每天有成千上万条新闻，并且新闻会实时更新。如果能从海量的新闻语料中，提取投资目标企业相关联的大数据，例如企业内部状况：经营、财务、高管、招聘、网站更新频率等，企业外部状况，例如关联公司状况如上下游、客户等，评级机构对该企业的评级，新闻媒体相关报道等信息，将这些信息形成关系网络，分析、评估投资目标企业的风险系数，从而投资方可根据风险系数考虑能否接受该风险并决定是否投资该企业。因此，如何从新闻语料中提取投资目标企业相关联的信息，并利用该信息进行风险评估是急需解决的问题。

### 发明内容

本申请提供一种企业投资风险评估方法、电子装置及计算机可读存储介质，其主要目的在于通过对新闻语料中透露的信息进行分析，评估投资目标企业的风险。

为实现上述目的，本申请提供一种电子装置，该电子装置包括：存储器、处理器，所述存储器存储有可在所述处理器上运行的企业投资风险评估程序，

该程序被所述处理器执行时实现如下步骤：

A1、爬取待评估风险的企业实体相关的新闻语料，对新闻语料进行预处理，从经过预处理后的新闻语料中提取与该企业实体相关联的其他实体；

A2、以名称为节点、该企业实体与其他实体之间的关联关系为边，构建该企业实体与其他实体之间的关系网络；

A3、根据关系网络计算该企业实体的向量表示，生成该企业实体的第一特征向量；

A4、根据第一预设规则，对该企业实体的内部信息进行量化，生成第二特征向量；

A5、从新闻语料中提取该企业实体的外部信息，根据第二预设规则，对该企业实体的外部信息进行量化，生成该企业实体的第三特征向量；及

A6、将所述第一特征向量、第二特征向量及第三特征向量输入预先确定的企业风险评估模型，输出得到该企业实体对应的风险标签。

此外，为实现上述目的，本申请还提供一种企业投资风险评估方法，该方法包括：

S1、爬取待评估风险的企业实体相关的新闻语料，对新闻语料进行预处理，从经过预处理后的新闻语料中提取与该企业实体相关联的其他实体；

S2、以名称为节点、该企业实体与其他实体之间的关联关系为边，构建该企业实体与其他实体之间的关系网络；

S3、根据关系网络计算该企业实体的向量表示，生成该企业实体的第一特征向量；

S4、根据第一预设规则，对该企业实体的内部信息进行量化，生成第二特征向量；

S5、从新闻语料中提取该企业实体的外部信息，根据第二预设规则，对该企业实体的外部信息进行量化，生成该企业实体的第三特征向量；及

S6、将所述第一特征向量、第二特征向量及第三特征向量输入预先确定的企业风险评估模型，输出得到该企业实体对应的风险标签。

此外，为实现上述目的，本申请还提供一种计算机可读存储介质，所述计算机可读存储介质存储有企业投资风险评估程序，该程序被处理器执行时实现如上所述的企业投资风险评估方法的任意步骤。

此外，为实现上述目的，本申请还提供一种企业投资风险评估，该程序包括：提取模块、构建模块、第一计算模块、第二计算模块、第三计算模块以及评估模块，该程序被处理器执行时实现如上所述的企业投资风险评估方法的任意步骤。

本申请提出的企业投资风险评估方法、电子装置及计算机可读存储介质，通过从新闻语料中了解企业实体与其关联实体之间的关系、企业实体的内部信息及外部信息，分别得到该企业实体的第一特征向量、第二特征向量及第三特征向量，利用风险评估模型及所述第一特征向量、第二特征向量及第三特征向量，对投资该企业实体进行风险评估，便于投资方捕捉市场投资机会，提前预测投资风险。

## 附图说明

图1为本申请企业投资风险评估方法较佳实施例的应用环境示意图；

图2为企业实体A与相关联的其他实体之间的关系网络图；

图3为企业实体A的向量表示；

图4为图1中企业投资风险评估程序的模块示意图；

图5为本申请企业投资风险评估方法较佳实施例的流程图。

本申请目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

## 具体实施方式

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。

本申请提供一种企业投资风险评估方法，该方法应用于一种电子装置1。参照图1所示，为本申请企业投资风险评估方法较佳实施例的应用环境示意图。在本实施例中，所述电子装置1可以是PC（Personal Computer，个人电脑），也可以是智能手机、平板电脑、电子书阅读器、便携计算机等终端设备。该电子装置1包括存储器11、处理器12，通信总线13，以及网络接口14。

其中，存储器11至少包括一种类型的可读存储介质，所述可读存储介质包括闪存、硬盘、多媒体卡、卡型存储器（例如，SD或DX存储器等）、磁

性存储器、磁盘、光盘等。存储器 11 在一些实施例中可以是所述电子装置 1 的内部存储单元，例如该电子装置 1 的硬盘。存储器 11 在另一些实施例中也可以是所述电子装置 1 的外部存储设备，例如该电子装置 1 上配备的插接式硬盘，智能存储卡（Smart Media Card, SMC），安全数字（Secure Digital, SD）卡，闪存卡（Flash Card）等。进一步地，存储器 11 还可以既包括该电子装置 1 的内部存储单元也包括外部存储设备。存储器 11 不仅可以用于存储安装于该电子装置 1 的应用软件及各类数据，例如企业投资风险评估程序 10 等，还可以用于暂时地存储已经输出或者将要输出的数据。

处理器 12 在一些实施例中可以是一中央处理器（Central Processing Unit, CPU）、控制器、微控制器、微处理器或其他数据处理芯片，用于运行存储器 11 中存储的程序代码或处理数据，例如企业投资风险评估程序 10 等。

通信总线 13 用于实现这些组件之间的连接通信。

网络接口 14 可选的可以包括标准的有线接口、无线接口（如 WI-FI 接口），通常用于在该装置与其他电子设备之间建立通信连接。

图 1 仅示出了具有组件 11-14 的电子装置 1，但是应理解的是，并不要求实施所有示出的组件，可以替代的实施更多或者更少的组件。

可选地，该电子装置 1 还可以包括用户接口，用户接口可以包括显示器（Display）、输入单元比如键盘（Keyboard），可选的用户接口还可以包括标准的有线接口、无线接口。

可选地，在一些实施例中，显示器可以是 LED 显示器、液晶显示器、触控式液晶显示器以及 OLED（Organic Light-Emitting Diode，有机发光二极管）触摸器等。其中，显示器也可以称为显示屏或显示单元，用于显示在电子装置 1 中处理的信息以及用于显示可视化的用户界面。

在图 1 所示的装置实施例中，存储器 11 中存储有企业投资风险评估程序 10；处理器 12 执行存储器 11 中存储的企业投资风险评估程序 10 时实现如下步骤：

A1、爬取待评估风险的企业实体相关的新闻语料，对新闻语料进行预处理，从经过预处理后的新闻语料中提取与该企业实体相关联的其他实体；

A2、以名称为节点、该企业实体与其他实体之间的关联关系为边，构建该企业实体与其他实体之间的关系网络；

A3、根据关系网络计算该企业实体的向量表示，生成该企业实体的第一特征向量；

A4、根据第一预设规则，对该企业实体的内部信息进行量化，生成第二特征向量；

A5、从新闻语料中提取该企业实体的外部信息，根据第二预设规则，对该企业实体的外部信息进行量化，生成该企业实体的第三特征向量；及

A6、将所述第一特征向量、第二特征向量及第三特征向量输入预先确定的企业风险评估模型，输出得到该企业实体对应的风险标签。

语料涉及多个不同的领域，本实施例以新闻语料为例，对本申请的具体方案进行说明，但不仅仅限于新闻领域。当投资方需要了解时下的新闻，以获取投资目标企业相关联的内部数据及外部数据时，利用网络爬虫从互联网中爬取网络新闻，例如，通过爬虫爬取新浪、百度、腾讯等的网络新闻。可以理解的是，每个企业在不同的时间段内的经营情况都不一样，因此，为了使投资者更准确的了解投资目标企业的信息，在时间维度上对爬取的网络新闻进行过滤，设置预设时间区间，只爬取该时间段的网络新闻，例如，只爬取近半年的网络新闻。由于新闻语料的来源具有多样性，因此，语料中格式类型比较多，为便于对语料进行后续处理，需对新闻语料进行预处理，得到新闻语料文本数据，形成新闻语料文本集。

在其他实施例中，所述预处理可以将新闻语料的格式统一为文本格式，从新闻语料中去除广告噪声并过滤脏词和敏感词中的一种或多种。在将新闻语料的格式统一为文本格式时，可以将当前技术暂不能转换为文本格式的内容过滤掉。

接下来，利用上述分词的方法，根据预先确定的企业名称库，从经过预处理后的新闻语料中提取出所有企业名称，然后根据待评估风险的企业实体（即投资目标企业）的关联企业数据，筛选出与待评估风险的企业实体相关联的其他实体，并将企业实体和其他实体构建成关系网络。其中，所述关联企业数据可以通过第三方数据获得。可以理解的是，从新闻语料中提取与企业实体相关联的其他实体可能很多，若要将所有关联实体全部构建在关系网络不合理，因此，在构建关系网络前，对提取出的与企业实体相关联的其他实体进行过滤筛选，具体地，通过过滤筛选步骤后保留的与企业实体相关联

的其他实体包括：该企业实体的股东公司、与企业实体有发生金钱往来的其他实体、供应商、客户、信用评结构等。

在本实施例中，以企业实体 A 为例，对从新闻语料中提取的与企业实体 A 相关联的其他实体进行筛选后，假设保留的其他实体分别为 B1、B2、B3，其中，B1 为给企业实体 A 进行信用评级的评级机构，从历史评级数据中可以了解到 B1 给企业实体 A 的信用评级为 BBB，B2 为给企业实体 A 提供原材料或货品的供应商，企业实体 A 对 B2 的欠款金额为 30 万，B3 为企业实体 A 的客户，企业实体 A 曾对 B3 违约 2 次。以企业实体 A、B1、B2、B3 为节点，以 B1、B2、B3 与 A 的关联关系为边，构建如图 2 所示的企业实体与其他实体之间的关系网络图。

然后，根据上述关系网络图，计算企业实体 A 的向量表示，本实施例采用的是 Skip-Gram 方法，因为关系网络中企业实体 A 的向量表示与其相关联的实体 B1、B2、B3 的向量表示之间存在着管理关系。对于企业实体名称向量的训练，Skip-Gram 方法利用当前企业实体去预测周围实体，如图 3 所示。图 3 中的 An1，An2，An3，An4 是没有顺序的，均表示为企业实体 A 的相邻实体。与利用 Skip-Gram 训练词向量的方法类似，设置一个固定的预测长度 L，来预测企业实体 A 周围的 L 个相邻实体，若真实情况相邻实体不及 L 个，则输出为 NULL。通过该方法，可以得到企业实体 A 的向量表示 embedding(E1)，embedding(E2)，...，将该向量表示作为企业实体 A 的第一特征向量。

可以理解的是，要了解投资企业实体 A 的风险，必须了解企业实体 A 的财务、经营等方面的信息，因此，需要考虑到企业实体 A 的内部信息，其中，内部信息包括企业实体 A 的经营、财务、招聘、网站更新频率等信息，其中部分信息为数字信息，例如财务信息包括企业上一年度的净利润、股票收益等。按照规则将企业内部信息中每一个参考因素转换为数字进行量化，例如，财务信息里的数值可以转化为特征值，比如在本实施例中，净利润为 30 万元，取 30 为对应的特征值，网站更新频率、最近一年的招聘人数也是数值，也可以按预设的转换规则为相应的数值。在其他实施例中，也可以按预设的转换比例将 30 万元转换为其他数值。将企业实体 A 的内部信息中的每个参考因素量化后，生成企业实体 A 的第二特征向量。

需要说明的是，企业实体 A 的经营好坏，除了自身因素外，外界的因素也至关重要，因此还需要考虑企业实体 A 的外部信息，其中，外部信息包括企业实体 A 的上下游关系，例如供应商、客户，该企业是否对上下游关系的其他实体产生过违约、欠款，如有，违约次数、欠款周期分别为多少。此外，企业实体 A 的外部信息还包括评级机构对企业实体 A 的评级（评级级别 3A，2A 表示优良，A 表示良好，BBB 表示一般等）、新闻媒体对该企业实体 A 的正面/负面报道等。然后，按照规则将企业内部信息中每一个参考因素转换为数字进行量化，例如，在本实施例中，违约次数可以量化为 3 个数值，无违约-0，轻度违约-1，重度违约-2；欠款可以量化为 2 个数值，无欠款-0，有欠款-1；评级可以量化为多个数值，评级级别 3A-6，评级级别 2A-5，评级级别 A-4，评级级别 BBB-3，评级级别 BB-2，评级级别 B-1。根据企业实体 A 的具体情况，将其外部信息进行量化，违约次数-1，欠款-1，评级-3，依据量化后的信息生成企业实体 A 的第三特征向量。

至此，了解了与企业实体 A 相关联的其他实体、企业实体 A 的内部信息及外部信息后，接下来就可以对投资企业实体 A 进行风险评估。将企业实体 A 的名称及企业实体 A 的第一特征向量、第二特征向量及第三特征向量输入预先确定的风险评估模型中进行风险评估，并输出风险评估结果。其中，所述预先确定的风险评估模型的训练步骤包括：利用上述 A1-A5 步骤，获取大量企业实体的第一特征向量、第二特征向量及第三特征向量，其具体实施方式与上述步骤一致，这里不再赘述。然后为各个企业实体标注风险标签，对“无风险”的企业实体，标注风险标签为 0，对“高风险”的企业实体，标注风险标签为 1，然后将各个企业实体的第一特征向量、第二特征向量、第三特征向量及对应的风险标签作为样本数据。从样本数据中的随机抽取第一比例（例如 60%）的企业实体的第一特征向量、第二特征向量、第三特征向量及该第一比例（例如 60%）的企业实体对应的风险标签作为训练集，从剩下的样本集中的随机抽取第二比例（例如 50%）的企业实体的第一特征向量、第二特征向量、第三特征向量及该第二比例（例如 50%）的企业实体对应的风险标签作为验证集，也就是说，抽取样本数据的 20% 的样本数据作为验证集；利用所述 50% 的样本数据对支持向量机进行训练，确定风险评估模型的模型参数，确定出企业实体的关联实体、内部信息、外部信息与投资该企业实体的风险

之间的关系；利用 20%的样本数据对所述风险评估模型的准确性进行验证，若准确率大于或者等于预设准确率（例如 90%），则训练结束，或者，若准确率小于预设准确率（例如 90%），则增加样本数量并重新执行训练步骤。

将企业实体 A 的第一特征向量、第二特征向量及第三特征向量输入所述风险评估模型后，若模型输出结果为 0，则表示投资企业实体 A 基本无风险，若模型输出结果为 1，则表示投资企业实体 A 有较大风险。

上述实施例提出的电子装置 1，通过了解企业实体与其关联实体之间的关系、企业实体的内部信息及外部信息，分别得到该企业实体的第一特征向量、第二特征向量及第三特征向量，利用风险评估模型及所述第一特征向量、第二特征向量及第三特征向量，对投资该企业实体进行风险评估，便于投资方捕捉市场投资机会。

可选地，在其他的实施例中，企业投资风险评估程序 10 还可以被分割为一个或者多个模块，一个或者多个模块被存储于存储器 11 中，并由一个或者多个处理器（本实施例为处理器 12）所执行，以完成本申请，本申请所称的模块是指能够完成特定功能的一系列计算机程序指令段。例如，参照图 4 所示，为图 1 中企业投资风险评估程序 10 的模块示意图，在本实施例中，该程序可以被分割为提取模块 110、构建模块 120、第一计算模块 130、第二计算模块 140、第三计算模块 150 以及评估模块 160，所述模块 110-160 所实现的功能或操作步骤均与上文类似，此处不再详述，示例性地，例如其中：

提取模块 110，用于爬取待评估风险的企业实体相关的新闻语料，对新闻语料进行预处理，从经过预处理后的新闻语料中提取与该企业实体相关联的其他实体；

构建模块 120，用于以名称为节点、该企业实体与其他实体之间的关联关系为边，构建该企业实体与其他实体之间的关系网络；

第一计算模块 130，用于根据关系网络计算该企业实体的向量表示，生成该企业实体的第一特征向量；

第二计算模块 140，用于根据第一预设规则，对该企业实体的内部信息进行量化，生成第二特征向量；

第三计算模块 150，用于从新闻语料中提取该企业实体的外部信息，根据

第二预设规则，对该企业实体的外部信息进行量化，生成该企业实体的第三特征向量；及

评估模块 160，用于将所述第一特征向量、第二特征向量及第三特征向量输入预先确定的企业风险评估模型，输出得到该企业实体对应的风险标签。

此外，本申请还提供一种企业投资风险评估方法。参照图 5 所示，为本申请企业投资风险评估方法较佳实施例的流程图。该方法可以由一个装置执行，该装置可以由软件和/或硬件实现。

在本实施例中，企业投资风险评估方法包括：

S1、爬取待评估风险的企业实体相关的新闻语料，对新闻语料进行预处理，从经过预处理后的新闻语料中提取与该企业实体相关联的其他实体；

S2、以名称为节点、该企业实体与其他实体之间的关联关系为边，构建该企业实体与其他实体之间的关系网络；

S3、根据关系网络计算该企业实体的向量表示，生成该企业实体的第一特征向量；

S4、根据第一预设规则，对该企业实体的内部信息进行量化，生成第二特征向量；

S5、从新闻语料中提取该企业实体的外部信息，根据第二预设规则，对该企业实体的外部信息进行量化，生成该企业实体的第三特征向量；及

S6、将所述第一特征向量、第二特征向量及第三特征向量输入预先确定的企业风险评估模型，输出得到该企业实体对应的风险标签。

语料涉及多个不同的领域，本实施例以新闻语料为例，对本申请的具体方案进行说明，但不仅仅限于新闻领域。当投资方需要了解时下的新闻，以获取投资目标企业相关联的内部数据及外部数据时，利用网络爬虫从互联网中爬取网络新闻，例如，通过爬虫爬取新浪、百度、腾讯等的网络新闻。可以理解的是，每个企业在不同的时间段内的经营情况都不一样，因此，为了使投资者更准确的了解投资目标企业的信息，在时间维度上对爬取的网络新闻进行过滤，设置预设时间区间，只爬取该时间段的网络新闻，例如，只爬取近半年的网络新闻。由于新闻语料的来源具有多样性，因此，语料中格式类型比较多，为便于对语料进行后续处理，需对新闻语料进行预处理，得到

新闻语料文本数据，形成新闻语料文本集。

在其他实施例中，所述预处理可以将新闻语料的格式统一为文本格式，从新闻语料中去除广告噪声并过滤脏词和敏感词中的一种或多种。在将新闻语料的格式统一为文本格式时，可以将当前技术暂不能转换为文本格式的内容过滤掉。

接下来，利用上述分词的方法，根据预先确定的企业名称库，从经过预处理后的新闻语料中提取出所有企业名称，然后根据待评估风险的企业实体（即投资目标企业）的关联企业数据，筛选出与待评估风险的企业实体相关联的其他实体，并将企业实体和其他实体构建成关系网络。其中，所述关联企业数据可以通过第三方数据获得。可以理解的是，从新闻语料中提取与企业实体相关联的其他实体可能很多，若要将所有关联实体全部构建在关系网络不合理，因此，在构建关系网络前，对提取出的与企业实体相关联的其他实体进行过滤筛选，具体地，通过过滤筛选步骤后保留的与企业实体相关联的其他实体包括：该企业实体的股东公司、与企业实体有发生金钱往来的其他实体、供应商、客户、信用评结构等。

在本实施例中，以企业实体 A 为例，对从新闻语料中提取的与企业实体 A 相关联的其他实体进行筛选后，假设保留的其他实体分别为 B1、B2、B3，其中，B1 为给企业实体 A 进行信用评级的评级机构，从历史评级数据中可以了解到 B1 给企业实体 A 的信用评级为 BBB，B2 为给企业实体 A 提供原材料或货品的供应商，企业实体 A 对 B2 的欠款金额为 30 万，B3 为企业实体 A 的客户，企业实体 A 曾对 B3 违约 2 次。以企业实体 A、B1、B2、B3 为节点，以 B1、B2、B3 与 A 的关联关系为边，构建如图 2 所示的企业实体与其他实体之间的关系网络图。

然后，根据上述关系网络图，计算企业实体 A 的向量表示，本实施例采用的是 Skip-Gram 方法，因为关系网络中企业实体 A 的向量表示与其相关联的实体 B1、B2、B3 的向量表示之间存在着管理关系。对于企业实体名称向量的训练，Skip-Gram 方法利用当前企业实体去预测周围实体，如图 3 所示。图 3 中的 An1，An2，An3，An4 是没有顺序的，均表示为企业实体 A 的相邻实体。与利用 Skip-Gram 训练词向量的方法类似，设置一个固定的预测长度 L，来预测企业实体 A 周围的 L 个相邻实体，若真实情况相邻实体不及 L 个，则

输出为 NULL。通过该方法,可以得到企业实体 A 的向量表示 embedding(E1), embedding(E2), ..., 将该向量表示作为企业实体 A 的第一特征向量。

可以理解的是,要了解投资企业实体 A 的风险,必须了解企业实体 A 的财务、经营等方面的信息,因此,需要考虑到企业实体 A 的内部信息,其中,内部信息包括企业实体 A 的经营、财务、招聘、网站更新频率等信息,其中部分信息为数字信息,例如财务信息包括企业上一年度的净利润、股票收益等。按照规则将企业内部信息中每一个参考因素转换为数字进行量化,例如,财务信息里的数值可以转化为特征值,比如在本实施例中,净利润为 30 万元,取 30 为对应的特征值,网站更新频率、最近一年的招聘人数也是数值,也可以按预设的转换规则为相应的数值。在其他实施例中,也可以按预设的转换比例将 30 万元转换为其他数值。将企业实体 A 的内部信息中的每个参考因素量化后,生成企业实体 A 的第二特征向量。

需要说明的是,企业实体 A 的经营好坏,除了自身因素外,外界的因素也至关重要,因此还需要考虑企业实体 A 的外部信息,其中,外部信息包括企业实体 A 的上下游关系,例如供应商、客户,该企业是否对上下游关系的其他实体产生过违约、欠款,如有,违约次数、欠款周期分别为多少。此外,企业实体 A 的外部信息还包括评级机构对企业实体 A 的评级(评级级别 3A, 2A 表示优良, A 表示良好, BBB 表示一般等)、新闻媒体对该企业实体 A 的正面/负面报道等。然后,按照规则将企业内部信息中每一个参考因素转换为数字进行量化,例如,在本实施例中,违约次数可以量化为 3 个数值,无违约-0,轻度违约-1,重度违约-2;欠款可以量化为 2 个数值,无欠款-0,有欠款-1;评级可以量化为多个数值,评级级别 3A-6,评级级别 2A-5,评级级别 A-4,评级级别 BBB-3,评级级别 BB-2,评级级别 B-1。根据企业实体 A 的具体情况,将其外部信息进行量化,违约次数-1,欠款-1,评级-3,依据量化后的信息生成企业实体 A 的第三特征向量。

至此,了解了与企业实体 A 相关联的其他实体、企业实体 A 的内部信息及外部信息后,接下来就可以对投资企业实体 A 进行风险评估。将企业实体 A 的名称及企业实体 A 的第一特征向量、第二特征向量及第三特征向量输入预先确定的风险评估模型中进行风险评估,并输出风险评估结果。其中,所述预先确定的风险评估模型的训练步骤包括:利用上述 S1-S5 步骤,获取大

量企业实体的第一特征向量、第二特征向量及第三特征向量，其具体实施方式与上述步骤一致，这里不再赘述。然后为各个企业实体标注风险标签，对“无风险”的企业实体，标注风险标签为 0，对“高风险”的企业实体，标注风险标签为 1，然后将各个企业实体的第一特征向量、第二特征向量、第三特征向量及对应的风险标签作为样本数据。从样本数据中的随机抽取第一比例（例如 60%）的企业实体的第一特征向量、第二特征向量、第三特征向量及该第一比例（例如 60%）的企业实体对应的风险标签作为训练集，从剩下的样本集中的随机抽取第二比例（例如 50%）的企业实体的第一特征向量、第二特征向量、第三特征向量及该第二比例（例如 50%）的企业实体对应的风险标签作为验证集，也就是说，抽取样本数据的 20% 的样本数据作为验证集；利用所述 50% 的样本数据对支持向量机进行训练，确定风险评估模型的模型参数，确定出企业实体的关联实体、内部信息、外部信息与投资该企业实体的风险之间的关系；利用 20% 的样本数据对所述风险评估模型的准确性进行验证，若准确率大于或者等于预设准确率（例如 90%），则训练结束，或者，若准确率小于预设准确率（例如 90%），则增加样本数量并重新执行训练步骤。

将企业实体 A 的第一特征向量、第二特征向量及第三特征向量输入所述风险评估模型后，若模型输出结果为 0，则表示投资企业实体 A 基本无风险，若模型输出结果为 1，则表示投资企业实体 A 有较大风险。

上述实施例提出的企业投资风险评估方法，通过了解企业实体与其关联实体之间的关系、企业实体的内部信息及外部信息，分别得到该企业实体的第一特征向量、第二特征向量及第三特征向量，利用风险评估模型及所述第一特征向量、第二特征向量及第三特征向量，对投资该企业实体进行风险评估，便于投资方捕捉市场投资机会。

此外，本申请实施例还提出一种计算机可读存储介质，所述计算机可读存储介质存储有企业投资风险评估程序，所述企业投资风险评估程序被处理器执行时实现如下操作：

A1、爬取待评估风险的企业实体相关的新闻语料，对新闻语料进行预处理，从经过预处理后的新闻语料中提取与该企业实体相关联的其他实体；

A2、以名称为节点、该企业实体与其他实体之间的关联关系为边，构建

该企业实体与其他实体之间的关系网络；

A3、根据关系网络计算该企业实体的向量表示，生成该企业实体的第一特征向量；

A4、根据第一预设规则，对该企业实体的内部信息进行量化，生成第二特征向量；

A5、从新闻语料中提取该企业实体的外部信息，根据第二预设规则，对该企业实体的外部信息进行量化，生成该企业实体的第三特征向量；及

A6、将所述第一特征向量、第二特征向量及第三特征向量输入预先确定的企业风险评估模型，输出得到该企业实体对应的风险标签。

本申请计算机可读存储介质具体实施方式与上述企业投资风险评估方法和电子装置各实施例基本相同，在此不作累述。

需要说明的是，上述本申请实施例序号仅仅为了描述，不代表实施例的优劣。并且本文中的术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、装置、物品或者方法不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、装置、物品或者方法所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括该要素的过程、装置、物品或者方法中还存在另外的相同要素。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在如上所述的一个存储介质(如 ROM/RAM、磁碟、光盘)中，包括若干指令用以使得一台终端设备(可以是手机，计算机，服务器，或者网络设备等)执行本申请各个实施例所述的方法。

以上仅为本申请的优选实施例，并非因此限制本申请的专利范围，凡是利用本申请说明书及附图内容所作的等效结构或等效流程变换，或直接或间接运用在其他相关的技术领域，均同理包括在本申请的专利保护范围内。

## 权利要求书

1、一种企业投资风险评估方法，应用于电子装置，其特征在于，该方法包括：

5 S1、爬取待评估风险的企业实体相关的新闻语料，对新闻语料进行预处理，从经过预处理后的新闻语料中提取与该企业实体相关联的其他实体；

S2、以名称为节点、该企业实体与其他实体之间的关联关系为边，构建该企业实体与其他实体之间的关系网络；

10 S3、根据关系网络计算该企业实体的向量表示，生成该企业实体的第一特征向量；

S4、根据第一预设规则，对该企业实体的内部信息进行量化，生成第二特征向量；

S5、从新闻语料中提取该企业实体的外部信息，根据第二预设规则，对该企业实体的外部信息进行量化，生成该企业实体的第三特征向量；及

15 S6、将所述第一特征向量、第二特征向量及第三特征向量输入预先确定的企业风险评估模型，输出得到该企业实体对应的风险标签。

2、如权利要求1所述的企业投资风险评估方法，其特征在于，所述“根据关系网络计算该企业实体的向量表示”的步骤采用的是 Skip-Gram 方法。

20 3、如权利要求1所述的企业投资风险评估方法，其特征在于，所述第一预设规则为：将所述企业实体的内部信息中每一个参考因素转换为数字量化的规则。

4、如权利要求1所述的企业投资风险评估方法，其特征在于，所述第二预设规则为：将所述企业实体的外部信息中每一个参考因素转换为数字量化的规则。

25 5、如权利要求1所述的企业投资风险评估方法，其特征在于，所述步骤S1中预处理包括：将新闻语料的格式统一为文本格式，从新闻语料中去除广告噪声。

6、如权利要求1所述的企业投资风险评估方法，其特征在于，所述预先确定的企业风险评估模型的训练步骤包括：

30 爬取多个企业实体相关的新闻语料，从新闻语料提取与该多个企业实体

相关联的其他实体，以名称为节点、实体之间的关联关系为边，分别构建该多个企业实体与其他实体之间的关系网络；

根据关系网络分别计算所述多个企业实体的向量表示，生成所述多个企业实体的第一特征向量；

- 5        根据第一预设规则，对所述多个企业实体的内部信息进行量化，生成第二特征向量；

从新闻语料中提取该企业实体的外部信息，根据第二预设规则，对该企业实体的外部信息进行量化，生成该企业实体的第三特征向量；

- 10       根据历史风险评估记录，分别给所述多个企业实体标注风险标签，将多个企业实体的第一特征向量、第二特征向量、第三特征向量及风险标签作为样本数据；

抽取第一比例的样本数据作为训练集，抽取第二比例的样本数据作为验证集；

利用所述训练集对支持向量机进行训练，得到所述风险评估模型；及

- 15       利用所述验证集对所述风险评估模型的准确性进行验证，若准确率大于或者等于预设准确率，则训练结束，或者，若准确率小于预设准确率，则增加样本数量并重新执行训练步骤。

- 20       7、一种电子装置，其特征在于，该装置包括：存储器、处理器，所述存储器存储有可在所述处理器上运行的企业投资风险评估程序，该程序被所述处理器执行时实现如下步骤：

A1、爬取待评估风险的企业实体相关的新闻语料，对新闻语料进行预处理，从经过预处理后的新闻语料中提取与该企业实体相关联的其他实体；

A2、以名称为节点、该企业实体与其他实体之间的关联关系为边，构建该企业实体与其他实体之间的关系网络；

- 25       A3、根据关系网络计算该企业实体的向量表示，生成该企业实体的第一特征向量；

A4、根据第一预设规则，对该企业实体的内部信息进行量化，生成第二特征向量；

- 30       A5、从新闻语料中提取该企业实体的外部信息，根据第二预设规则，对该企业实体的外部信息进行量化，生成该企业实体的第三特征向量；及

A6、将所述第一特征向量、第二特征向量及第三特征向量输入预先确定的企业风险评估模型，输出得到该企业实体对应的风险标签。

8、根据权利要求7所述的电子装置，其特征在于，所述“根据关系网络计算该企业实体的向量表示”的步骤采用的是 Skip-Gram 方法。

5 9、根据权利要求7所述的电子装置，其特征在于，所述第一预设规则为：将所述企业实体的内部信息中每一个参考因素转换为数字量化的规则。

10、根据权利要求7所述的电子装置，其特征在于，所述第二预设规则为：将所述企业实体的外部信息中每一个参考因素转换为数字量化的规则。

10 11、根据权利要求7所述的电子装置，其特征在于，所述步骤A1中预处理包括：将新闻语料的格式统一为文本格式，从新闻语料中去除广告噪声。

12、根据权利要求7所述的电子装置，其特征在于，所述预先确定的企业风险评估模型的训练步骤包括：

15 爬取多个企业实体相关的新闻语料，从新闻语料提取与该多个企业实体相关联的其他实体，以名称为节点、实体之间的关联关系为边，分别构建该多个企业实体与其他实体之间的关系网络；

根据关系网络分别计算所述多个企业实体的向量表示，生成所述多个企业实体的第一特征向量；

根据第一预设规则，对所述多个企业实体的内部信息进行量化，生成第二特征向量；

20 从新闻语料中提取该企业实体的外部信息，根据第二预设规则，对该企业实体的外部信息进行量化，生成该企业实体的第三特征向量；

根据历史风险评估记录，分别给所述多个企业实体标注风险标签，将多个企业实体的第一特征向量、第二特征向量、第三特征向量及风险标签作为样本数据；

25 抽取第一比例的样本数据作为训练集，抽取第二比例的样本数据作为验证集；

利用所述训练集对支持向量机进行训练，得到所述风险评估模型；及

30 利用所述验证集对所述风险评估模型的准确性进行验证，若准确率大于或者等于预设准确率，则训练结束，或者，若准确率小于预设准确率，则增加样本数量并重新执行训练步骤。

13、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质上存储有企业投资风险评估程序，该程序被处理器执行时实现如下步骤：

A1、爬取待评估风险的企业实体相关的新闻语料，对新闻语料进行预处理，从经过预处理后的新闻语料中提取与该企业实体相关联的其他实体；

5 A2、以名称为节点、该企业实体与其他实体之间的关联关系为边，构建该企业实体与其他实体之间的关系网络；

A3、根据关系网络计算该企业实体的向量表示，生成该企业实体的第一特征向量；

10 A4、根据第一预设规则，对该企业实体的内部信息进行量化，生成第二特征向量；

A5、从新闻语料中提取该企业实体的外部信息，根据第二预设规则，对该企业实体的外部信息进行量化，生成该企业实体的第三特征向量；及

A6、将所述第一特征向量、第二特征向量及第三特征向量输入预先确定的企业风险评估模型，输出得到该企业实体对应的风险标签。

15 14、根据权利要求 13 所述的计算机可读存储介质，其特征在于，所述“根据关系网络计算该企业实体的向量表示”的步骤采用的是 Skip-Gram 方法。

15、根据权利要求 13 所述的计算机可读存储介质，其特征在于，所述第一预设规则为：将所述企业实体的内部信息中每一个参考因素转换为数字量化的规则。

20 16、根据权利要求 13 所述的计算机可读存储介质，其特征在于，所述第二预设规则为：将所述企业实体的外部信息中每一个参考因素转换为数字量化的规则。

25 17、根据权利要求 13 所述的计算机可读存储介质，其特征在于，所述步骤 A1 中预处理包括：将新闻语料的格式统一为文本格式，从新闻语料中去除广告噪声。

18、根据权利要求 13 所述的计算机可读存储介质，其特征在于，所述预先确定的企业风险评估模型的训练步骤包括：

30 爬取多个企业实体相关的新闻语料，从新闻语料提取与该多个企业实体相关联的其他实体，以名称为节点、实体之间的关联关系为边，分别构建该多个企业实体与其他实体之间的关系网络；

根据关系网络分别计算所述多个企业实体的向量表示，生成所述多个企业实体的第一特征向量；

根据第一预设规则，对所述多个企业实体的内部信息进行量化，生成第二特征向量；

- 5       从新闻语料中提取该企业实体的外部信息，根据第二预设规则，对该企业实体的外部信息进行量化，生成该企业实体的第三特征向量；

根据历史风险评估记录，分别给所述多个企业实体标注风险标签，将多个企业实体的第一特征向量、第二特征向量、第三特征向量及风险标签作为样本数据；

- 10       抽取第一比例的样本数据作为训练集，抽取第二比例的样本数据作为验证集；

利用所述训练集对支持向量机进行训练，得到所述风险评估模型；及

- 15       利用所述验证集对所述风险评估模型的准确性进行验证，若准确率大于或者等于预设准确率，则训练结束，或者，若准确率小于预设准确率，则增加样本数量并重新执行训练步骤。

19、一种企业投资风险评估程序，其特征在于，该程序包括：提取模块、构建模块、第一计算模块、第二计算模块、第三计算模块以及评估模块。

- 20       20、如权利要求 19 所述的企业投资风险评估程序，其特征在于，改程序被处理器执行时，可实现如权利要求 1 至 6 中任意一项企业投资风险评估方法的步骤。

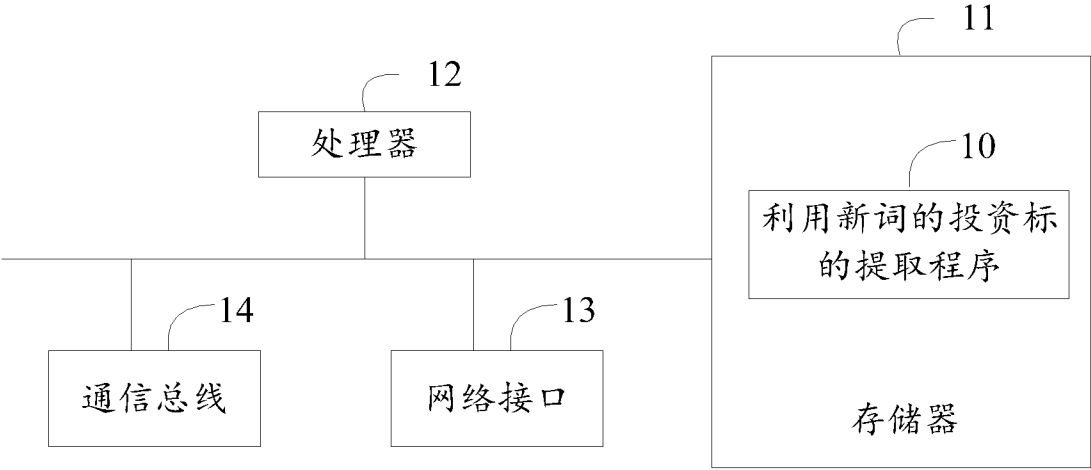


图 1

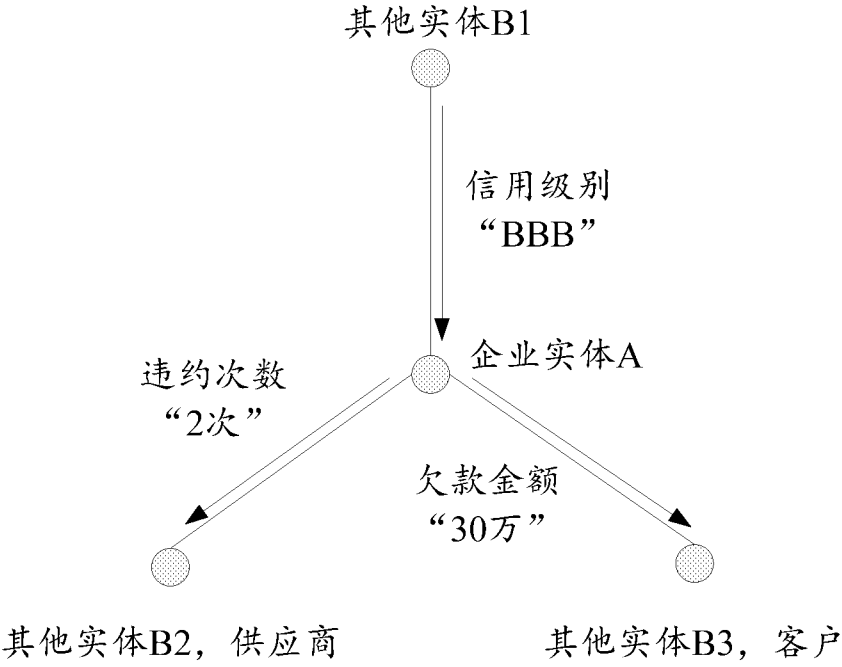


图 2

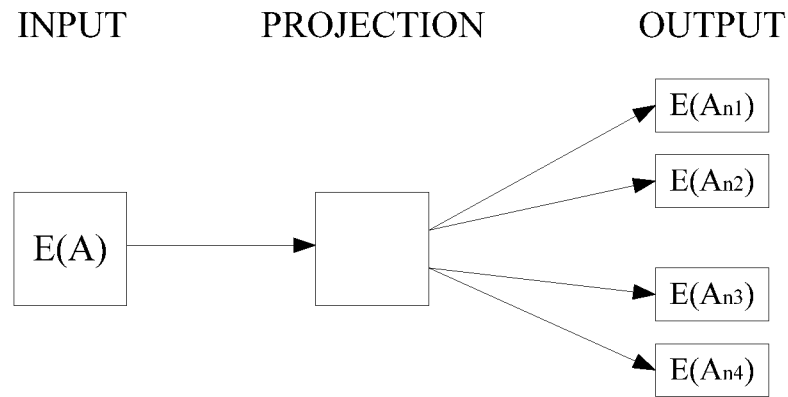


图 3



图 4

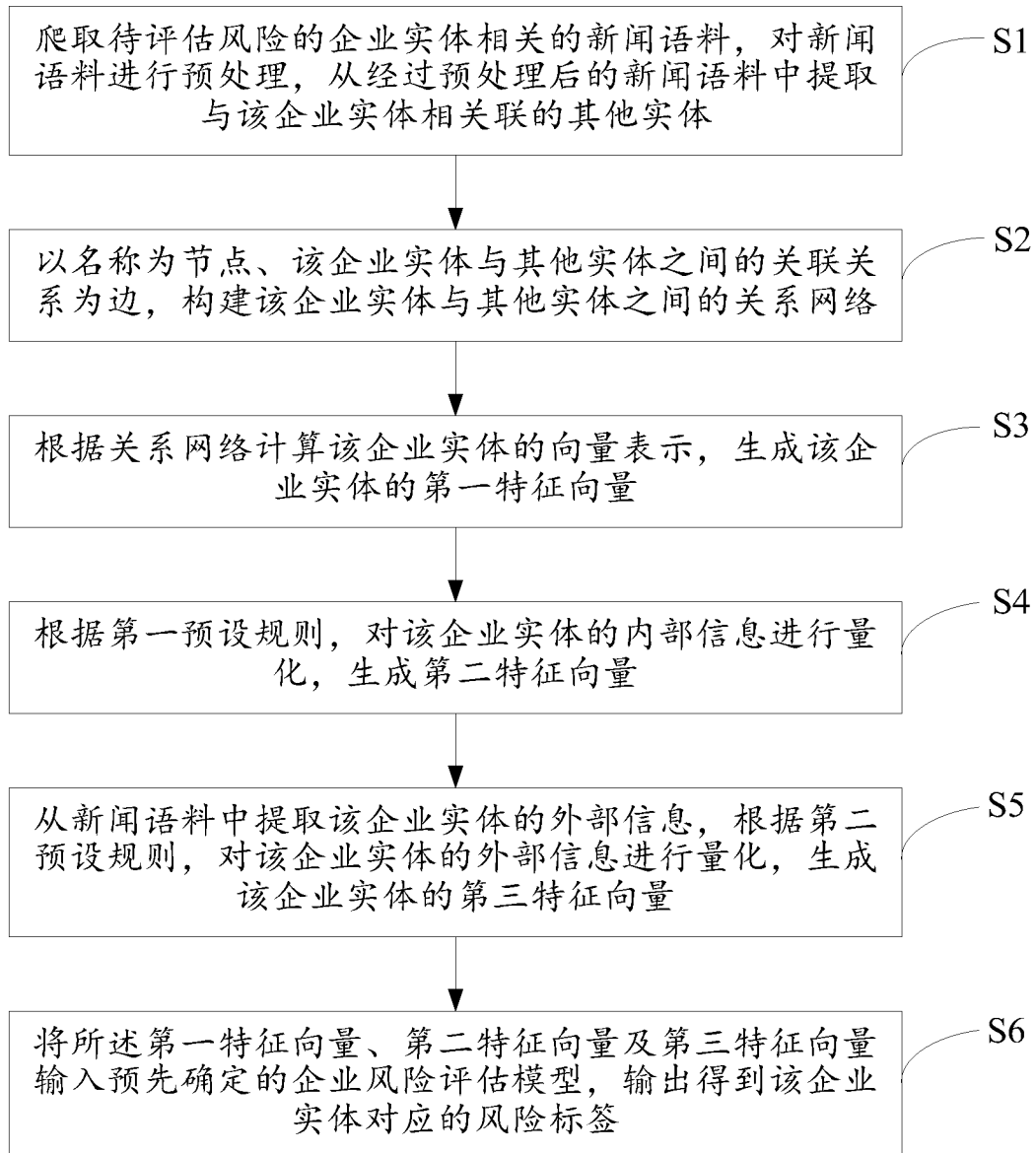


图 5

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/CN2018/076169

## A. CLASSIFICATION OF SUBJECT MATTER

G06Q 10/06 (2012.01) i

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)

G06Q; G06F

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: 企业, 公司, 投资, 融资, 贷款, 爬取, 爬虫, 关联, 关系, 新闻, 内部, 外部, 经营, 财务, 招聘, 违约, 欠款, 评级, company, loan, invest, financing, associate, relate, news, management, invite, breach, arrearage, evaluate, grade

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                       | Relevant to claim No. |
|-----------|--------------------------------------------------------------------------------------------------------------------------|-----------------------|
| X         | CN 106203808 A (CHINA MINSHENG BANKING CORP., LTD.) 07 December 2016 (07.12.2016), description, paragraphs [0031]-[0096] | 19                    |
| Y         | CN 106203808 A (CHINA MINSHENG BANKING CORP., LTD.) 07 December 2016 (07.12.2016), description, paragraphs [0031]-[0096] | 1-18, 20              |
| Y         | CN 105740335 A (SHANDONG HEETIAN CORPORATION) 06 July 2016 (06.07.2016), description, paragraphs [0003]-[0036]           | 1-18, 20              |
| A         | CN 107229756 A (SHANDONG HEETIAN CORPORATION) 03 October 2017 (03.10.2017), entire document                              | 1-20                  |
| A         | US 2015149247 A1 (THE DUN & BRADSTREET CORPORATION) 28 May 2015 (28.05.2015), entire document                            | 1-20                  |

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

|                                                                                                                                                                         |                                                                                                                                                                                                                                                  |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| * Special categories of cited documents:                                                                                                                                | "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                                              |
| "A" document defining the general state of the art which is not considered to be of particular relevance                                                                | "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                                                                     |
| "E" earlier application or patent but published on or after the international filing date                                                                               | "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 |
| "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) | "&" document member of the same patent family                                                                                                                                                                                                    |
| "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                                                                  |                                                                                                                                                                                                                                                  |

|                                                                           |                                                                      |
|---------------------------------------------------------------------------|----------------------------------------------------------------------|
| Date of the actual completion of the international search<br>27 June 2018 | Date of mailing of the international search report<br>08 August 2018 |
|---------------------------------------------------------------------------|----------------------------------------------------------------------|

|                                                                                                                                                                                                               |                                                                       |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------|
| Name and mailing address of the ISA<br>State Intellectual Property Office of the P. R. China<br>No. 6, Xitucheng Road, Jimenqiao<br>Haidian District, Beijing 100088, China<br>Facsimile No. (86-10) 62019451 | Authorized officer<br>YANG, Weiming<br>Telephone No. (86-10) 53961571 |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------|

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
PCT/CN2018/076169

| Patent Documents referred<br>in the Report | Publication Date | Patent Family    | Publication Date  |
|--------------------------------------------|------------------|------------------|-------------------|
| CN 106203808 A                             | 07 December 2016 | None             |                   |
| CN 105740335 A                             | 06 July 2016     | None             |                   |
| CN 107229756 A                             | 03 October 2017  | None             |                   |
| US 2015149247 A1                           | 28 May 2015      | AU 2014202660 B2 | 24 September 2015 |
|                                            |                  | HK 1206465 A1    | 08 January 2016   |
|                                            |                  | AU 2014202660 A1 | 20 November 2014  |
|                                            |                  | SG 11201402420 A | 27 February 2015  |
|                                            |                  | AU 2014202660 C1 | 09 June 2016      |
|                                            |                  | CN 104321794 B   | 02 February 2018  |
|                                            |                  | WO 2014179552 A1 | 06 November 2014  |
|                                            |                  | CN 104321794 A   | 28 January 2015   |
|                                            |                  | CA 2851464 A1    | 02 November 2014  |

## 国际检索报告

国际申请号

PCT/CN2018/076169

## A. 主题的分类

G06Q 10/06 (2012.01) i

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)

G06Q; G06F

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, EPODOC, CNPAT, CNKI: 企业, 公司, 投资, 融资, 贷款, 爬取, 爬虫, 关联, 关系, 新闻, 内部, 外部, 经营, 财务, 招聘, 违约, 欠款, 评级, company, loan, invest, financing, associate, relate, news, management, invite, breach, arrearage, evaluate, grade

## 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 106203808 A (中国民生银行股份有限公司) 2016年 12月 7日 (2016 - 12 - 07)<br>说明书第[0031]-[0096]段       | 19       |
| Y    | CN 106203808 A (中国民生银行股份有限公司) 2016年 12月 7日 (2016 - 12 - 07)<br>说明书第[0031]-[0096]段       | 1-18, 20 |
| Y    | CN 105740335 A (山东合天智汇信息技术有限公司) 2016年 7月 6日 (2016 - 07 - 06)<br>说明书第[0003]-[0036]段      | 1-18, 20 |
| A    | CN 107229756 A (山东合天智汇信息技术有限公司) 2017年 10月 3日 (2017 - 10 - 03)<br>全文                     | 1-20     |
| A    | US 2015149247 A1 (THE DUN & BRADSTREET CORPORATION) 2015年 5月 28日 (2015 - 05 - 28)<br>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

“&amp;” 同族专利的文件

国际检索实际完成的日期

2018年 6月 27日

国际检索报告邮寄日期

2018年 8月 8日

ISA/CN的名称和邮寄地址

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

传真号 (86-10) 62019451

受权官员

杨威明

电话号码 86-10-53961571

国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2018/076169

| 检索报告引用的专利文件 |            |    | 公布日<br>(年/月/日) | 同族专利 |              |    | 公布日<br>(年/月/日) |
|-------------|------------|----|----------------|------|--------------|----|----------------|
| CN          | 106203808  | A  | 2016年 12月 7日   | 无    |              |    |                |
| CN          | 105740335  | A  | 2016年 7月 6日    | 无    |              |    |                |
| CN          | 107229756  | A  | 2017年 10月 3日   | 无    |              |    |                |
| US          | 2015149247 | A1 | 2015年 5月 28日   | AU   | 2014202660   | B2 | 2015年 9月 24日   |
|             |            |    |                | HK   | 1206465      | A1 | 2016年 1月 8日    |
|             |            |    |                | AU   | 2014202660   | A1 | 2014年 11月 20日  |
|             |            |    |                | SG   | 11201402420V | A  | 2015年 2月 27日   |
|             |            |    |                | AU   | 2014202660   | C1 | 2016年 6月 9日    |
|             |            |    |                | CN   | 104321794    | B  | 2018年 2月 2日    |
|             |            |    |                | WO   | 2014179552   | A1 | 2014年 11月 6日   |
|             |            |    |                | CN   | 104321794    | A  | 2015年 1月 28日   |
|             |            |    |                | CA   | 2851464      | A1 | 2014年 11月 2日   |

表 PCT/ISA/210 (同族专利附件) (2015年1月)