

(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
2017年8月3日 (03.08.2017)



(10) 国际公布号
WO 2017/128973 A1

- (51) 国际专利分类号:
G06F 17/30 (2006.01)
- (21) 国际申请号: PCT/CN2017/071241
- (22) 国际申请日: 2017年1月16日 (16.01.2017)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201610048385.8 2016年1月25日 (25.01.2016) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。
- (72) 发明人; 及
- (71) 申请人 (仅对美国): 周俊 (ZHOU, Jun) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼 阿里巴巴集团法务部, Zhejiang 311121 (CN)。 崔卿 (CUI, Qing) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼 阿里巴巴集团法务部, Zhejiang 311121 (CN)。

- (74) 代理人: 北京国昊天诚知识产权代理有限公司 (CO-HORIZON INTELLECTUAL PROPERTY INC.); 中国北京市朝阳区小关北里甲 2 号渔阳置业大厦 B 座 605, Beijing 100029 (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, 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)。

[见续页]

(54) Title: METHOD AND APPARATUS FOR PROCESSING USER BEHAVIOUR DATA

(54) 发明名称: 用户行为数据的处理方法和装置

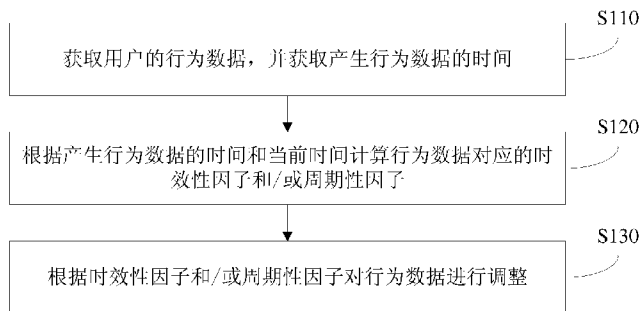


图 1

- S110 ACQUIRING USER BEHAVIOUR DATA, AND ACQUIRING A TIME WHEN THE BEHAVIOUR DATA IS GENERATED
- S120 ACCORDING TO THE TIME WHEN THE BEHAVIOUR DATA IS GENERATED AND A CURRENT TIME, CALCULATING A TIMELINESS FACTOR AND/OR A PERIODICITY FACTOR CORRESPONDING TO THE BEHAVIOUR DATA
- S130 ACCORDING TO THE TIMELINESS FACTOR AND/OR THE PERIODICITY FACTOR, ADJUSTING THE BEHAVIOUR DATA

(57) Abstract: A method and apparatus for processing user behaviour data. The processing method comprises the following steps: acquiring user behaviour data, and acquiring a time when the behaviour data is generated (S110); according to the time when the behaviour data is generated and a current time, calculating a timeliness factor and/or a periodicity factor corresponding to the behaviour data (S120); and according to the timeliness factor and/or the periodicity factor, adjusting the behaviour data (S130). The processing method in the embodiments can improve the timelessness of user behaviour data, and can more effectively acquire the preference and interest of a user, so that a search result more satisfying the user requirements is provided for the user, and the user experience is improved.

(57) 摘要: 一种用户行为数据的处理方法和装置, 其中, 处理方法包括以下步骤: 获取用户的行为数据, 并获取产生行为数据的时间(S110); 根据产生行为数据的时间和当前时间计算行为数据对应的时效性因子和/或周期性因子(S120); 根据时效性因子和/或周期性因子对行为数据进行调整(S130)。实施例的处理方法, 可以提升用户行为数据的时效性, 能够更加有效地获取到用户的偏好和兴趣, 为用户提供更加满足用户需求的搜索结果, 提升了用户的使用体验。

WO 2017/128973 A1

本国际公布:

- 包括国际检索报告(条约第 21 条(3))。

用户行为数据的处理方法和装置

技术领域

本申请涉及互联网技术领域，尤其涉及一种用户行为数据的处理方法和
5 装置。

背景技术

搜索引擎经常用机器学习方法通过对历史日志数据的学习来优化多个
搜索结果的排序结果，为了能够取得更理想的搜索结果，往往会使用过去
10 两周、一个月甚至更长时间的历史数据来预测用户对搜索结果的偏好。

然而，目前存在的问题是，搜索引擎在通过历史日志数据优化搜索结果
的排序结果时通常是平等对待所有的历史数据，例如，一个 10 天前的
行为日志和一个 1 天前的历史数据，搜索引擎在根据行为日志预测用户偏
好时并不区分历史数据的重要程度，将所有的历史数据按照同样的重要程
15 度用机器学习方法预测用户的偏好，导致搜索引擎无法有效地捕捉用户的
兴趣变化，不能满足用户的需求。

发明内容

本申请的目的旨在至少在一定程度上解决上述的技术问题之一。

20 为此，本申请的第一个目的在于提出一种用户行为数据的处理方法，
该处理方法可以提升用户行为数据的时效性，能够更加有效地获取到用户
的偏好和兴趣，为用户提供更加满足用户需求的搜索结果，提升了用户的
使用体验。

本申请的第二个目的在于提出了一种用户行为数据的处理装置。

25 为达上述目的，根据本申请第一方面实施例提出的一种用户行为数据
的处理方法，包括以下步骤：获取用户的行为数据，并获取产生所述行为
数据的时间；根据所述产生所述行为数据的时间和当前时间计算所述行为

数据对应的时效性因子和/或周期性因子；根据所述时效性因子和/或周期性因子对所述行为数据进行调整。

本申请实施例的处理方法，结合时间衰减和周期性相似性为用户的行为数据赋予不同的重要程度，利用不同重要程度的行为数据通过机器学习方法优化多个搜索结果的排序结果，从而提升了用户行为数据的时效性，能够更加有效地获取到用户的偏好和兴趣，为用户提供更加满足用户需求的搜索结果，提升了用户的使用体验。

为达上述目的，根据本申请的第二方面实施例提出的一种用户行为数据的处理装置，包括：获取模块，用于获取用户的行为数据，并获取产生所述行为数据的时间；计算模块，用于根据所述产生所述行为数据的时间和当前时间计算所述行为数据对应的时效性因子和/或周期性因子；处理模块，用于根据所述时效性因子和/或周期性因子对所述行为数据进行调整。

本申请实施例的处理装置，结合时间衰减和周期性相似性为用户的行为数据赋予不同的重要程度，利用不同重要程度的行为数据通过机器学习方法优化多个搜索结果的排序结果，从而提升了用户行为数据的时效性，能够更加有效地获取到用户的偏好和兴趣，为用户提供更加满足用户需求的搜索结果，提升了用户的使用体验。

本申请附加的方面和优点将在下面的描述中部分给出，部分将从下面的描述中变得明显，或通过本申请的实践了解到。

20

附图说明

本申请的上述和/或附加的方面和优点从结合下面附图对实施例的描述中将变得明显和容易理解，其中：

图 1 是根据本申请一个实施例的用户行为数据的处理方法的流程图；

图 2 是本发明一个具体实施例的用户行为数据的处理方法的流程图；

图 3 是本发明另一个具体实施例的用户行为数据的处理方法的流程图；
以及

图 4 是根据本申请一个实施例的用户行为数据的处理装置的结构示意图。

具体实施方式

5 下面详细描述本申请的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的，旨在用于解释本申请，而不能理解为对本申请的限制。

下面参考附图描述本申请实施例的用户行为数据的处理方法和装置。

10 图 1 是根据本申请一个实施例的用户行为数据的处理方法的流程图。

如图 1 所示，该处理方法包括：

S110，获取用户的行为数据，并获取产生行为数据的时间。

15 例如，行为数据的格式一般如下表所示，其中，features 为行为数据中的特征，label 为用户的点击行为（用^分为两部分，第一部分为 1 表示用户点击，第二部分为 0 表示用户没有点击），date 为产生点击行为的时间。

features	label	date
city_1,age_2,age_city_4	0^1	2015-11-24
city_6,age_20,age_city_3	1^0	2015-11-20
city_199,age_10,age_city_1	0^1	2015-11-02

20 S120，根据产生行为数据的时间和当前时间计算行为数据对应的时效性因子和/或周期性因子。

例如，用户 A 在过去一个月里多次访问手机类的页面，然而，在过去的一周内用户浏览手机类页面和在一个月前用户浏览手机类页面的影响

是不同的，显然离现在越近时间点的用户行为的影响更大，更能符合用户的搜索需求。因此，在使用历史数据学习时，可以按照时间远近给历史数据赋予不同的重要程度。

再例如，有些用户行为是有周期性的，比如上班族用户 B 经常在周末的时候在网上购买食材，因为用户 B 周末的时候才有时间做饭。那么虽然周一离上个周末的时间上很接近，但是用户 B 的行为差异仍然很大。因此，在使用历史数据学习时，需要考虑周期性的影响，根据历史数据在某个维度上与当前周期性的相近程度来给历史数据赋予不同的重要程度。

具体地，计算产生行为数据的时间与当前时间的的时间差，根据特定的函数计算时效性因子 Z。根据产生行为数据的时间与当前时间，通过特定的函数计算出周期性因子 C。

在本发明的一个实施例中，在计算时效性因子 Z 和周期性因子 C 之前，还可以先将行为数据输入至数据切分装置进行切分，将行为数据切分为多个小的行为数据块。其中，数据切分装置根据行为数据的大小，以及处理行为数据的每台计算机器可以计算的数据量判断将行为数据切分成多少份。由于数据切分装置对行为数据的切分本身是在分布式环境下执行的，因此可以充分利用多个计算机器的资源来加速切分动作。数据切分装置将行为数据切分好后，将每个小的行为数据块分发至每台计算机器上，多台计算机器可以同时并行处理行为数据块，其中任意两台计算机器获取到的行为数据块不会重叠。

S130，根据时效性因子和/或周期性因子对行为数据进行调整。

具体地，根据时效性因子 Z 与用户的点击行为 label 计算得到新的点击行为 $label_1$ ，即 $label_1 = Z * label$ 。进而，根据周期性因子 C 与新的点击行为 $label_1$ 计算得到新的点击行为 $label_2$ 。也就是说，首先对用户的行为数据施加时效性和周期性的影响，然后利用得到的调整后的行为数据通过机器学习方法优化多个搜索结果的排序结果。

应当理解的是，在对用户的行为数据进行调整时，可以仅根据时效性因子 Z 对用户的行为数据进行调整，或者仅根据周期性因子 C 对用户的行

为数据进行调整, 或者还可以同时根据时效性因子 Z 和周期性因子 C 对用户的行为数据进行调整, 即先根据时效性因子 Z 对用户的行为数据进行调整, 再根据周期性因子 C 对用户的行为数据进行调整。

本申请实施例中的用户行为数据的处理方法, 结合时间衰减和周期性相似性为用户的行为数据赋予不同的重要程度, 利用不同重要程度的行为数据通过机器学习方法优化多个搜索结果的排序结果, 从而提升了用户行为数据的时效性, 能够更加有效地获取到用户的偏好和兴趣, 为用户提供更加满足用户需求的搜索结果, 提升了用户的使用体验。

为了使得本申请中对用户行为数据的处理方法的描述更加的具体化, 使得本领域的技术人员更加清楚的了解本申请, 下面以同时根据时效性因子 Z 和周期性因子 C 对用户的行为数据进行调整的具体示例进行说明。

图 2 是本发明一个具体实施例的用户行为数据的处理方法的流程图, 图 3 是本发明另一个具体实施例的用户行为数据的处理方法的流程图, 如图 2 和 3 所示, 用户行为数据的处理方法包括:

S210, 计算产生行为数据的日期与当前日期之间的时间差。其中, 产生行为数据的时间和当前时间为日期。

S220, 根据时间差和时间衰减函数计算时效性因子。其中, 时间衰减函数为指数函数或幂函数。

S230, 将时效性因子与行为数据的乘积作为第一行为数据。

S240, 计算产生行为数据的第 i 日和当前第 j 日的日期差。其中, 产生行为数据的时间和当前时间属于预设时间周期中的第 i 日和第 j 日, 其中预设时间周期具有 N 日, i 和 j 为小于或等于 N 的正整数。

S250, 根据日期差生成周期性因子。其中, 通过以下公式生成周期性因子:

$$1/\sqrt{\min(|d_j - d_i|, N - |d_j - d_i|)},$$

其中, d_i 为产生行为数据的第 i 日, d_j 为当前第 j 日。

S260, 根据周期性因子、行为数据与第一行为数据生成第二行为数据。

在本发明的一个实施例中，如图3所示，根据周期性因子、行为数据与第一行为数据生成第二行为数据具体包括：

S261，根据周期性因子和行为数据生成第三行为数据。

S262，根据第一行为数据和第三行为数据生成第二行为数据。其中，
5 将第一行为数据和第三行为数据相加或相乘以生成第二行为数据。

具体而言，首先，输入用户行为数据中的用户的点击行为 label 和产生点击行为的时间 date，并按照时间衰减规律来给用户的行为数据赋予权重。应当理解的时，在对用户的行为数据进行切分后，由于只需要输入用户的点击行为 label 和产生点击行为的时间 date，因此可以大大降低每台计算机
10 器的数据处理量。设 t_0 是当前日期， t 是产生行为数据的日期，因此， $t_0 - t$ 可以衡量用户的行为数据离当前日期的远近程度，也就是距离。如果时间衰减函数 $f(x)$ 是一个递减函数，即 $f(x)$ 在 $[0, +\infty)$ 上是一个递减函数，那么 $f(t_0 - t)$ 就构成一个按时间衰减给行为数据进行加权的方法。例如，一天前的行为数据权重是 $f(1)$ ，三天前的行为数据权重是 $f(3)$ ，以此类推。由于
15 $f(x)$ 为递减函数， $f(1) > f(3)$ ，因而离现在时间间隔越大的行为数据的权重越小。其中，时间衰减函数 $f(x)$ 可以是指数函数 $f(x) = a^x, 0 < a < 1$ ，还可以是幂函数 $f(x) = x^{-a}, a > 0$ 等。

应当理解的是，在按照时间衰减规律给用户的行为数据赋予权重时，当前时间和产生行为数据的时间可以以天为单位，即当前日期和产生行为
20 数据的日期，也可以以小时或者分钟等单位，本实施例中仅以以天为单位进行说明，以小时或分钟为单位计算行为数据的权重和以天为单位计算行为数据的权重的方式相同，为了避免冗余，此处不在赘。

进而，计算产生行为数据的日期与当前日期之间的差值得到时间差，即将当前日期减去行为数据中 date 的日期以得到时间差，然后选择指数函数
25 或者幂函数计算得到时效性因子 Z，例如时效性因子 Z 为指数函数 $f(x) = 0.98^x$ ，然后计算时效性处理后的用户的点击行为 $label_1$ ， $label_1$ 等于用户的点击行为 label 乘以时效性因子 Z，即 $label_1 = Z * label$ 。

然后，输入对用户的时效性处理后的用户的点击行为 $label_1$ 、用户的点

击行为 label 以及产生点击行为的时间 date，并根据时间差通过公式
 $1/\sqrt{\min(|d_j - d_i|, N - |d_j - d_i|)}$ 得到周期性因子 C。例如，假设产生行为
 数据的是一周中的第 d_i 天，以一周中的七天为例，以数字 1 到 7 来记录周
 一到周日，因此当前第 d_j 天时行为数据的权重就应该是
 5 $1/\sqrt{\min(|d_j - d_i|, N - |d_j - d_i|)}$ ，其中，N 等于 7。

应当理解的是，这里仅以一周中的第几天来描述周期性信息，任何其他
 周期性的单位都可以用类似方法衡量，例如，一天中的 24 小时，节气的
 变更，温度的升降等周期性信息。

进而，根据选择的叠加关系计算得到 $label_2$ 。其中，叠加关系可以为加
 10 或者乘等，如果叠加关系是加，则：

$$label_2 = label_1 + label * 1/\sqrt{\min(|d_j - d_i|, N - |d_j - d_i|)}$$

如果叠加关系是乘，则：

$$label_2 = label_1 * label * 1/\sqrt{\min(|d_j - d_i|, N - |d_j - d_i|)}$$

然后，然后利用得到的调整后的行为数据 $label_2$ 通过机器学习方法优化
 15 多个搜索结果的排序结果。

本申请实施例中的用户行为数据的处理方法，先对行为数据进行时效
 性的处理，再对时效性处理后的行为数据进行周期性处理，从而根据周期
 性处理后的行为数据通过机器学习方法优化多个搜索结果的排序结果以
 捕捉用户的偏好，从而提升了用户行为数据的时效性，能够更加有效地获
 20 取到用户的偏好和兴趣，为用户提供更加满足用户需求的搜索结果，提升
 了用户的使用体验。

为了实现上述实施例，本申请还提出一种用户行为数据的处理装置。

图 4 是根据本申请一个实施例的用户行为数据的处理装置的结构示意
 图。

25 如图 4 所示，该处理装置包括：获取模块 100、计算模块 200 和处理
 模块 300。

其中，获取模块 100 用于获取用户的行为数据，并获取产生行为数据

的时间。

计算模块 200 用于根据产生行为数据的时间和当前时间计算行为数据对应的时效性因子和/或周期性因子。

5 处理模块 300 用于根据时效性因子和/或周期性因子对行为数据进行调整。

需要说明的是，前述对用户行为数据的处理方法的解释说明也适用于该实施例的用户行为数据的处理装置，其实现原理类似，此处不再赘述。

10 本申请实施例中的用户行为数据的处理装置，结合时间衰减和周期性相似性为用户的行为数据赋予不同的重要程度，利用不同重要程度的行为数据通过机器学习方法优化多个搜索结果的排序结果，从而提升了用户行为数据的时效性，能够更加有效地获取到用户的偏好和兴趣，为用户提供更加满足用户需求的搜索结果，提升了用户的使用体验。

15 为了使得本申请中对用户行为数据的处理装置的描述更加的具体化，使得本领域的技术人员更加清楚的了解本申请，下面以同时根据时效性因子 Z 和周期性因子 C 对用户的行为数据进行调整的具体示例进行说明。

在本申请的一个实施例中，产生行为数据的时间和当前时间为日期，其中，计算模块 200 用于计算产生行为数据的日期与当前日期之间的时间差，并根据时间差和时间衰减函数计算时效性因子。其中，时间衰减函数为指数函数或幂函数。

20 在本申请的另一个实施例中，产生行为数据的时间和当前时间属于预设时间周期中的第 i 日和第 j 日，其中，预设时间周期具有 N 日，i 和 j 为小于或等于 N 的正整数，其中，计算模块 200 还用于计算产生行为数据的第 i 日和当前第 j 日的日期差，并根据日期差生成周期性因子。其中，通过以下公式生成周期性因子： $1/\sqrt{\min(|d_j - d_i|, N - |d_j - d_i|)}$ ，其中， d_i 为产生行为数据的第 i 日， d_j 为当前第 j 日。

25 在本申请的另一个实施例中，处理模块 300 用于根据时效性因子和周期性因子对行为数据进行调整，并将时效性因子与行为数据的乘积作为第一行为数据，以及根据周期性因子、行为数据与第一行为数据生成第二行

为数据。

在本申请的另一个实施例中，处理模块 300 还用于根据周期性因子和
行为数据生成第三行为数据，并根据第一行为数据和第三行为数据生成第
二行为数据。其中，将第一行为数据和第三行为数据相加或相乘以生成第
5 二行为数据。

需要说明的是，前述对用户行为数据的处理方法的解释说明也适用于
该实施例的用户行为数据的处理装置，其实现原理类似，此处不再赘述。

本申请实施例中的用户行为数据的处理装置，先对行为数据进行时效
性的处理，再对时效性处理后的行为数据进行周期性处理，从而根据周期
10 性处理后的行为数据通过机器学习方法优化多个搜索结果的排序结果以
捕捉用户的偏好，从而提升了用户行为数据的时效性，能够更加有效地获
取到用户的偏好和兴趣，为用户提供更加满足用户需求的搜索结果，提升
了用户的使用体验。

在本申请的描述中，需要理解的是，术语“第一”、“第二”仅用于描述
15 目的，而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特
征的数量。由此，限定有“第一”、“第二”的特征可以明示或者隐含地包括
至少一个该特征。在本申请的描述中，“多个”的含义是至少两个，例如两
个，三个等，除非另有明确具体的限定。

在本说明书的描述中，参考术语“一个实施例”、“一些实施例”、“示例”、
20 “具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体
特征、结构、材料或者特点包含于本申请的至少一个实施例或示例中。在
本说明书中，对上述术语的示意性表述不必须针对的是相同的实施例或示
例。而且，描述的具体特征、结构、材料或者特点可以在任一个或多个实
施例或示例中以合适的方式结合。此外，在不相互矛盾的情况下，本领域
25 的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例
或示例的特征进行结合和组合。

流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为，
表示包括一个或更多个用于实现特定逻辑功能或过程的步骤的可执行指
令的代码的模块、片段或部分，并且本申请的优选实施方式的范围包括另
30 外的实现，其中可以不按所示出或讨论的顺序，包括根据所涉及的功能按

基本同时的方式或按相反的顺序，来执行功能，这应被本申请的实施例所属技术领域的技术人员所理解。

应当理解，本申请的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中，多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。例如，如果用硬件来实现，
5 和在另一实施方式中一样，可用本领域公知的下列技术中的任一项或他们的组合来实现：具有用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路，具有合适的组合逻辑门电路的专用集成电路，可编程门阵列（PGA），现场可编程门阵列（FPGA）等。

10 本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通程序来指令相关的硬件完成，所述的程序可以存储于一种计算机可读存储介质中，该程序在执行时，包括方法实施例的步骤之一或其组合。

此外，在本申请各个实施例中的各功能单元可以集成在一个处理模块
15 中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用，也可以存储在一个计算机可读取存储介质中。

20 上述提到的存储介质可以是只读存储器，磁盘或光盘等。尽管上面已经示出和描述了本申请的实施例，可以理解的是，上述实施例是示例性的，不能理解为对本申请的限制，本领域的普通技术人员在本申请的范围内可以对上述实施例进行变化、修改、替换和变型。

25

权 利 要 求 书

1、一种用户行为数据的处理方法，其特征在于，包括以下步骤：

获取用户的行为数据，并获取产生所述行为数据的时间；

5 根据所述产生所述行为数据的时间和当前时间计算所述行为数据对应的时效性因子和/或周期性因子；

根据所述时效性因子和/或周期性因子对所述行为数据进行调整。

2、如权利要求 1 所述的用户行为数据的处理方法，其特征在于，所述根据所述时效性因子和所述周期性因子对所述行为数据进行调整；

10 将所述时效性因子与所述行为数据的乘积作为第一行为数据；

根据所述周期性因子、所述行为数据与所述第一行为数据生成第二行为数据。

3、如权利要求 1 所述的用户行为数据的处理方法，其特征在于，所述产生所述行为数据的时间和所述当前时间为日期，其中，所述根据所述产生所述行为数据的时间和当前时间计算所述行为数据对应的时效性因子
15 具体包括：

计算所述产生所述行为数据的日期与所述当前日期之间的时间差；

根据所述时间差和时间衰减函数计算所述时效性因子。

4、如权利要求 3 所述的用户行为数据的处理方法，其特征在于，所述
20 时间衰减函数为指数函数或幂函数。

5、如权利要求 1 所述的用户行为数据的处理方法，其特征在于，所述产生所述行为数据的时间和所述当前时间属于预设时间周期中的第 i 日和第 j 日，其中，所述预设时间周期具有 N 日， i 和 j 为小于或等于 N 的正整数，其中，所述根据所述产生所述行为数据的时间和当前时间计算所述
25 行为数据对应的周期性因子具体包括：

计算产生所述行为数据的所述第 i 日和当前第 j 日的日期差；

根据所述日期差生成所述周期性因子。

6、如权利要求 5 所述的用户行为数据的处理方法，其特征在于，通过以下公式生成所述周期性因子：

$1/\sqrt{\min(|d_j - d_i|, N - |d_j - d_i|)}$ ，其中， d_i 为产生所述行为数据的所述第 i 日， d_j 为当前第 j 日。

5 7、如权利要求 2 所述的用户行为数据的处理方法，其特征在于，所述根据所述周期性因子、所述行为数据与所述第一行为数据生成第二行为数据具体包括：

根据所述周期性因子和所述行为数据生成第三行为数据；

根据所述第一行为数据和所述第三行为数据生成所述第二行为数据。

10 8、如权利要求 7 所述的用户行为数据的处理方法，其特征在于，将所述第一行为数据和所述第三行为数据相加或相乘以生成所述第二行为数据。

9、一种用户行为数据的处理装置，其特征在于，包括：

15 获取模块，用于获取用户的行为数据，并获取产生所述行为数据的时间；

计算模块，用于根据所述产生所述行为数据的时间和当前时间计算所述行为数据对应的时效性因子和/或周期性因子；

处理模块，用于根据所述时效性因子和/或周期性因子对所述行为数据进行调整。

20 10、如权利要求 9 所述的用户行为数据的处理装置，其特征在于，所述处理模块用于：

根据所述时效性因子和所述周期性因子对所述行为数据进行调整，并将所述时效性因子与所述行为数据的乘积作为第一行为数据，以及根据所述周期性因子、所述行为数据与所述第一行为数据生成第二行为数据。

25 11、如权利要求 9 所述的用户行为数据的处理装置，其特征在于，所述产生所述行为数据的时间和所述当前时间为日期，其中，所述计算模块用于：

计算所述产生所述行为数据的日期与所述当前日期之间的时间差，并根据所述时间差和时间衰减函数计算所述时效性因子。

12、如权利要求 11 所述的用户行为数据的处理装置，其特征在于，所述时间衰减函数为指数函数或幂函数。

5 13、如权利要求 1 所述的用户行为数据的处理装置，其特征在于，所述产生所述行为数据的时间和所述当前时间属于预设时间周期中的第 i 日和第 j 日，其中，所述预设时间周期具有 N 日， i 和 j 为小于或等于 N 的正整数，其中，所述计算模块用于：

10 计算产生所述行为数据的所述第 i 日和当前第 j 日的日期差，并根据所述日期差生成所述周期性因子。

14、如权利要求 13 所述的用户行为数据的处理装置，其特征在于，通过以下公式生成所述周期性因子：

$1/\sqrt{\min(|d_j - d_i|, N - |d_j - d_i|)}$ ，其中， d_i 为产生所述行为数据的所述第 i 日， d_j 为当前第 j 日。

15 15、如权利要求 10 所述的用户行为数据的处理装置，其特征在于，所述处理模块还用于：

根据所述周期性因子和所述行为数据生成第三行为数据，并根据所述第一行为数据和所述第三行为数据生成所述第二行为数据。

20 16、如权利要求 15 所述的用户行为数据的处理装置，其特征在于，将所述第一行为数据和所述第三行为数据相加或相乘以生成所述第二行为数据。

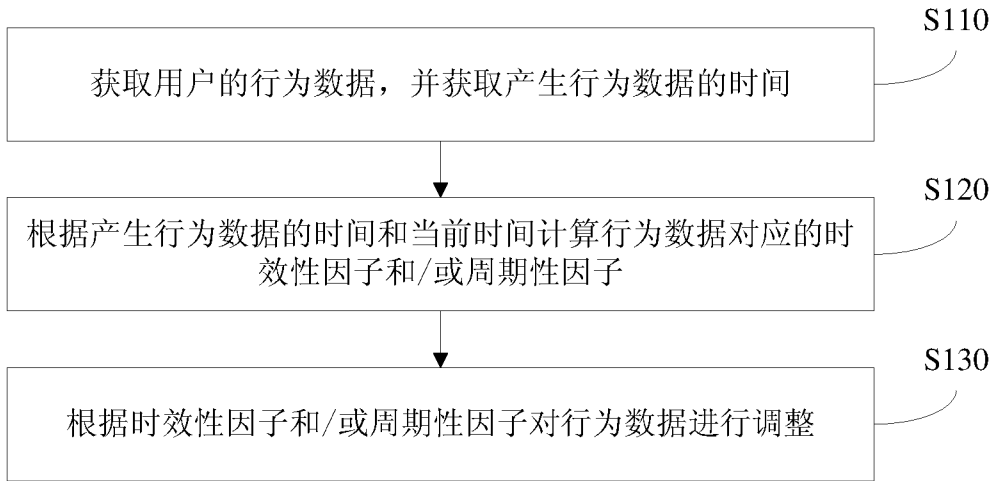


图 1

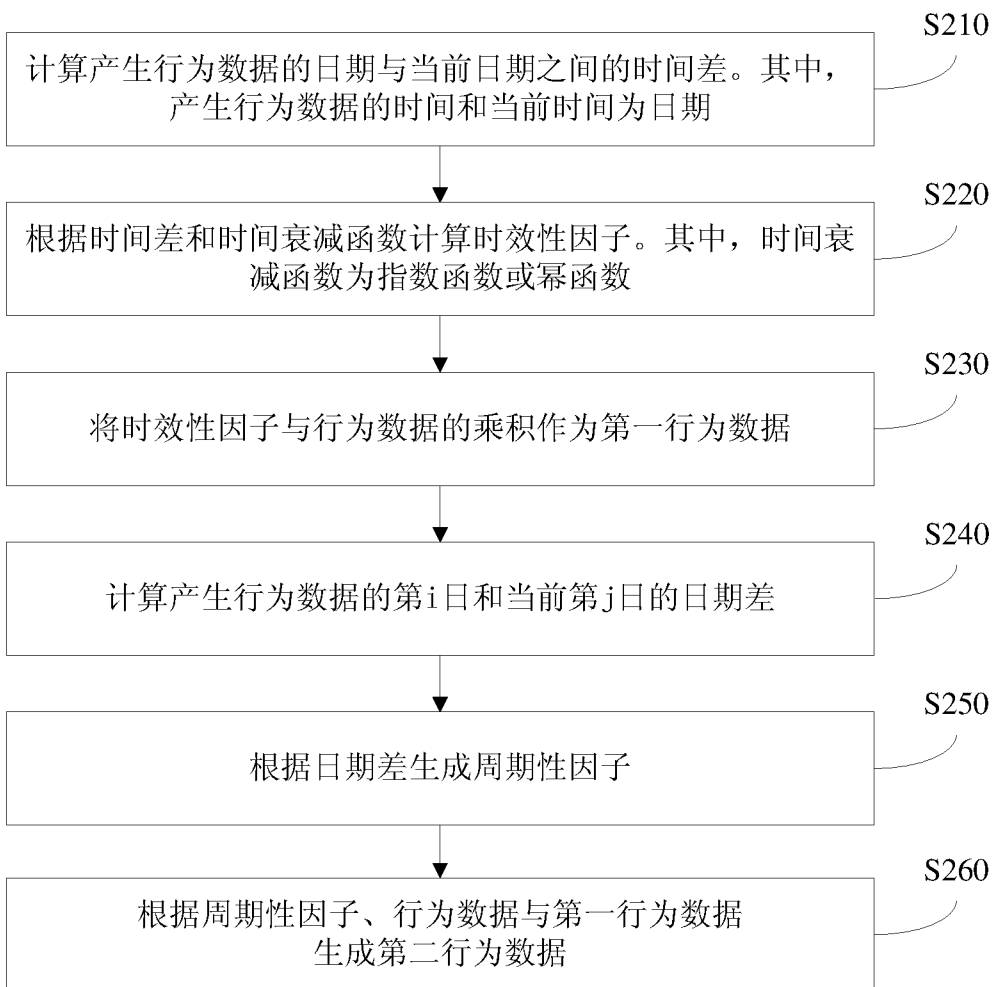


图 2

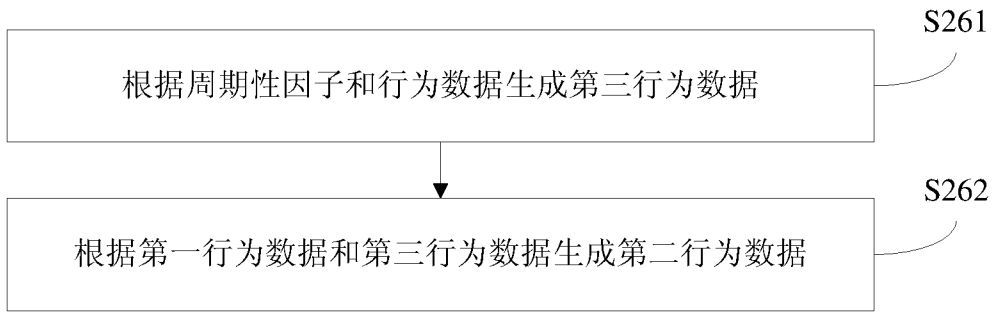


图 3



图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2017/071241

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/30 (2006.01) 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)

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; IEEE: aging, adjustment, factor, data, time, history, gene, user, behaviour

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 103164804 A (ALIBABA GROUP HOLDING LIMITED), 19 June 2013 (19.06.2013), description, paragraphs [0023]-[0072], and figures 1-6	1, 5, 9, 13
A	CN 102075352 A (BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS), 25 May 2011 (25.05.2011), the whole document	1-16
A	CN 104216884 A (KUSHENG (TIANJIN) SCIENCE AND TECHNOLOGY CO., LTD.), 17 December 2014 (17.12.2014), the whole document	1-16
A	US 2009089233 A1 (ROCKWELL AUTOMATION TECHNOLOGIES, INC.), 02 April 2009 (02.04.2009), the whole document	1-16

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p>
---	---

Date of the actual completion of the international search

21 February 2017 (21.02.2017)

Date of mailing of the international search report

17 April 2017 (17.04.2017)

Name and mailing address of the ISA/CN:
 State Intellectual Property Office of the P. R. China
 No. 6, Xitucheng Road, Jimenqiao
 Haidian District, Beijing 100088, China
 Facsimile No.: (86-10) 62019451

Authorized officer

LI, Meng

Telephone No.: (86-10) **62413008**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/071241

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 103164804 A	19 June 2013	EP 2791832 A1 WO 2013090686 A1 TW 201327233 A JP 2014533418 A US 2013166488 A1	22 October 2014 20 June 2013 01 July 2013 11 December 2014 27 June 2013
CN 102075352 A	25 May 2011	None	
CN 104216884 A	17 December 2014	None	
US 2009089233 A1	02 April 2009	CN 101408767 A EP 2042959 A1	15 April 2009 01 April 2009

<p>A. 主题的分类</p> <p>G06F 17/30 (2006.01) i</p> <p>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</p>																	
<p>B. 检索领域</p> <p>检索的最低限度文献 (标明分类系统和分类号)</p> <p>G06F</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>WPI; EPODOC; CNPAT; CNKI; IEEE: 用户, 行为, 历史, 时效, 时间, 调整, 数据, 因子, data, time, history, gene, user, behavior</p>																	
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 103164804 A (阿里巴巴集团控股有限公司) 2013年 6月 19日 (2013 - 06 - 19) 说明书第[0023]-[0072]段, 附图1-6</td> <td>1, 5, 9, 13</td> </tr> <tr> <td>A</td> <td>CN 102075352 A (北京邮电大学) 2011年 5月 25日 (2011 - 05 - 25) 全文</td> <td>1-16</td> </tr> <tr> <td>A</td> <td>CN 104216884 A (酷盛天津科技有限公司) 2014年 12月 17日 (2014 - 12 - 17) 全文</td> <td>1-16</td> </tr> <tr> <td>A</td> <td>US 2009089233 A1 (ROCKWELL AUTOMATION TECHNOLOGIES, INC.) 2009年 4月 2日 (2009 - 04 - 02) 全文</td> <td>1-16</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 103164804 A (阿里巴巴集团控股有限公司) 2013年 6月 19日 (2013 - 06 - 19) 说明书第[0023]-[0072]段, 附图1-6	1, 5, 9, 13	A	CN 102075352 A (北京邮电大学) 2011年 5月 25日 (2011 - 05 - 25) 全文	1-16	A	CN 104216884 A (酷盛天津科技有限公司) 2014年 12月 17日 (2014 - 12 - 17) 全文	1-16	A	US 2009089233 A1 (ROCKWELL AUTOMATION TECHNOLOGIES, INC.) 2009年 4月 2日 (2009 - 04 - 02) 全文	1-16
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
X	CN 103164804 A (阿里巴巴集团控股有限公司) 2013年 6月 19日 (2013 - 06 - 19) 说明书第[0023]-[0072]段, 附图1-6	1, 5, 9, 13															
A	CN 102075352 A (北京邮电大学) 2011年 5月 25日 (2011 - 05 - 25) 全文	1-16															
A	CN 104216884 A (酷盛天津科技有限公司) 2014年 12月 17日 (2014 - 12 - 17) 全文	1-16															
A	US 2009089233 A1 (ROCKWELL AUTOMATION TECHNOLOGIES, INC.) 2009年 4月 2日 (2009 - 04 - 02) 全文	1-16															
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p>																	
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p>																	
<p>国际检索实际完成的日期</p> <p>2017年 2月 21日</p>		<p>国际检索报告邮寄日期</p> <p>2017年 4月 17日</p>															
<p>ISA/CN的名称和邮寄地址</p> <p>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</p> <p>传真号 (86-10) 62019451</p>		<p>授权官员</p> <p>李萌</p> <p>电话号码 (86-10) 62413008</p>															

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/071241

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	103164804	A	2013年 6月 19日	EP	2791832	A1	2014年 10月 22日
				WO	2013090686	A1	2013年 6月 20日
				TW	201327233	A	2013年 7月 1日
				JP	2014533418	A	2014年 12月 11日
				US	2013166488	A1	2013年 6月 27日
CN	102075352	A	2011年 5月 25日	无			
CN	104216884	A	2014年 12月 17日	无			
US	2009089233	A1	2009年 4月 2日	CN	101408767	A	2009年 4月 15日
				EP	2042959	A1	2009年 4月 1日

表 PCT/ISA/210 (同族专利附件) (2009年7月)