

(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
2016 年 9 月 15 日 (15.09.2016)



(10) 国际公布号
WO 2016/141821 A1

- (51) 国际专利分类号:
G06F 17/30 (2006.01)
- (21) 国际申请号: PCT/CN2016/075379
- (22) 国际申请日: 2016 年 3 月 2 日 (02.03.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510108576.4 2015 年 3 月 12 日 (12.03.2015) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 英属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): 富国萍 (FU, Guoping) [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)。王瑶 (WANG, Yao) [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)。郑伟 (ZHENG, Wei) [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)。王全剑 (WANG, Quanjian) [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, 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, KN, KP, KR, 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,

[见续页]

(54) Title: PRODUCT INFORMATION PROCESSING METHOD AND SERVER

(54) 发明名称: 一种产品信息处理的方法及服务器

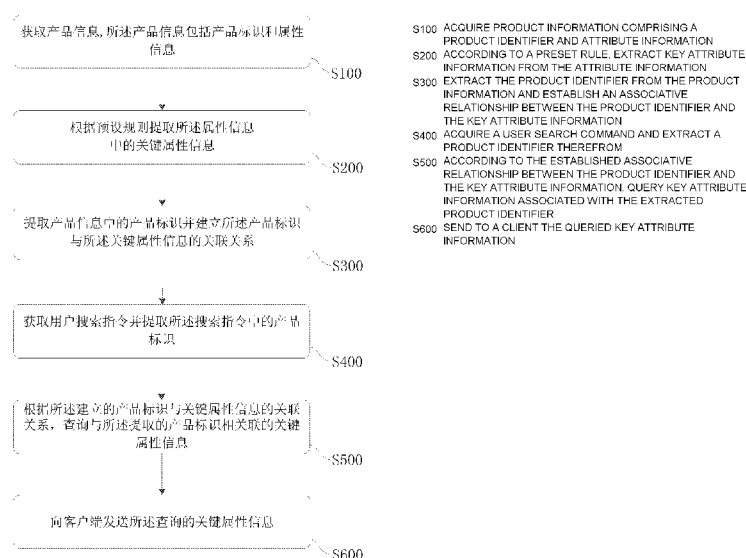


图 1

(57) Abstract: A product information processing method and server. The method comprises: acquiring product information comprising a product identifier and attribute information (S100); according to a preset rule, extracting key attribute information from the attribute information (S200); extracting the product identifier from the product information and establishing an associative relationship between the product identifier and the key attribute information (S300); acquiring a user search command and extracting a product identifier therefrom (S400); according to the established associative relationship between the product identifier and the key attribute information, querying key attribute information associated with the extracted product identifier (S500); sending to a client the queried key attribute information (S600). The key attribute information of a product related to a user search command can be sent to a client and viewed by a user, thus providing the user product information which the user is truly interested in.

(57) 摘要:

[见续页]



WO 2016/141821 A1

RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, **本国际公布:**
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, — 包括国际检索报告(条约第 21 条(3))。
TG)。

一种产品信息处理的方法及服务器。所述方法包括：获取产品信息，所述产品信息包括产品标识和属性信息（S100）；根据预设规则提取所述属性信息中的关键属性信息（S200）；提取所述产品信息中的产品标识并建立所述产品标识与所述关键属性信息的关联关系（S300）；获取用户搜索指令并提取所述搜索指令中的产品标识（S400）；根据所述建立的产品标识与关键属性信息的关联关系，查询与所述提取的产品标识相关联的关键属性信息（S500）；向客户端发送所述查询的关键属性信息（S600）。可以将与用户的搜索指令相关的产品的关键属性信息发送给客户端供用户浏览，能够给用户提用户真正关注的产品信息。

一种产品信息处理的方法及服务器

技术领域

本申请涉及互联网技术领域，特别涉及一种产品信息处理的方法及服务器。

背景技术

随着互联网技术的应用和发展，人们越来越习惯于在互联网上进行产品信息的筛选。在进行产品信息筛选时，用户往往会在提供搜索服务的网站的搜索引擎中输入自己感兴趣的产品关键词，那么网站便可以展示与该产品相关的信息供用户浏览。网站在向用户展示信息时，往往会在产品词条的下方列举出该产品的属性信息，所述属性信息例如可以是产品的产地，制作材料，颜色，最低成交价等。

现有技术中的网站往往按照下述的方式向用户展示产品信息：

网站首先可以预先设置固定的产品属性信息，该固定的产品属性信息例如可以是产地，颜色，最低成交价。无论用户输入的是什么产品关键词，网站均可以向用户提供搜索到的产品的这些固定属性信息。例如，网站预先设置的固定属性信息为原产地，颜色，最低成交价，那么当用户在网站的搜索引擎中输入“iPhone6”时，网站会向该用户展示与“iPhone6”有关的产品信息，并且在这些产品词条的下方列举出该产品的原产地，颜色以及最低成交价。当用户在该网站的搜索引擎中输入另一个产品关键词时，该网站同样会展示与该产品有关的信息，并且在这些产品词条的下方列举出原产地，颜色以及最低成交价这些预设的产品属性信息。

在实施本申请的过程中，发明人发现现有技术至少存在如下问题：

对于不同的产品，用户关注的属性信息往往是不同的。例如，对于“葡萄酒”而言，用户比较关注该产品的原产地。但对于“U盘”而言，用户却不是很关注该产品的原产地，而往往比较关注该产品的容量。如果对所有产品均设置固定的属性信息，那么网站往往会展示用户不关注的部分属性信息，而用户真正关注的属性信息却没能直接展示。

发明内容

本申请实施例的目的在于提供一种产品信息处理的方法及服务器，能够提供用户关注的属性信息。

本申请实施例提供的一种产品信息处理的方法及服务器是这样实现的：

一种产品信息处理的方法，包括：

获取产品信息，所述产品信息包括产品标识和属性信息；

根据预设规则提取所述属性信息中的关键属性信息；

提取所述产品信息中的产品标识，并建立所述产品标识与所述关键属性信息的关联关

系；

获取用户搜索指令并提取所述搜索指令中的产品标识；

根据所述建立的产品标识与所述关键属性信息的关联关系，查询与所述提取的产品标识相关联的关键属性信息；

向客户端发送所述查询的关键属性信息。

一种产品信息处理的服务器，包括：产品信息获取单元，关键属性信息提取单元，关联关系建立单元，产品标识获取单元，关键属性信息查询单元，发送单元，其中：

所述产品信息获取单元，用来获取产品信息，所述产品信息包括产品标识和属性信息；

所述关键属性信息提取单元，用来根据预设规则提取所述属性信息中的关键属性信息；

所述关联关系建立单元，用来提取所述产品信息中的产品标识，并建立所述产品标识与所述关键属性信息的关联关系；

所述产品标识获取单元，用来获取用户搜索指令并提取所述搜索指令中的产品标识；

所述关键属性信息查询单元，用来根据所述建立的产品标识与关键属性信息的关联关系，查询与所述提取的产品标识相关联的关键属性信息；

所述发送单元，用来向客户端发送所述查询的关键属性信息。

本申请实施例提供的一种产品信息处理的方法及服务器，通过对产品属性信息进行分析，针对不同的产品，提取出该产品的关键属性信息，并根据用户的搜索指令，将与用户的搜索指令相关的产品的关键属性信息发送给客户端供用户浏览，能够提供用户真正关注的属性信息。

附图说明

图 1 为本申请一实施例提供的一种产品信息处理的方法流程图；

图 2 为本申请一个例子中以线性分类为例解释通过支持向量机算法进行二维分类的基本原理示意图；

图 3 为本申请一个例子中以线性分类为例解释通过支持向量机算法进行三维分类的基本原理示意图；

图 4 为本申请一实施例提供的一种产品信息处理的服务器功能模块图；

图 5 为本申请另一实施例提供的一种产品信息处理的服务器功能模块图。

具体实施方式

本申请实施例提供一种产品信息处理的方法及服务器。为了使本技术领域的人员更好地

理解本申请中的技术方案，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它实施例，都应当属于本申请保护的范围。

本申请提供一种产品信息处理的方法及服务器涉及服务器和客户端。其中，客户端可以接收用户的搜索指令，并将该搜索指令发送至服务器。服务器获取用户的搜索指令后，可以对该指令进行一系列的处理，并最终向客户端返回用户感兴趣的产品信息。具体的过程如下所述：

图1为本申请一实施例提供的一种产品信息处理的方法流程图。如图1所示，所述方法包括：

S100：获取产品信息，所述产品信息包括产品标识和属性信息。

所述产品信息可以是产品的所有者或者网站服务器的管理者在发布该产品时预先登记在服务器内的。所述产品信息可以包括产品的价格，原产地，颜色，制作材料，容量等属性信息中的至少一个属性信息。产品的所有者或者网站的管理者在发布每个产品时，可以给每个产品分配唯一的编号。通过该编号可以查询对应产品的信息。另外，每个产品还可以对应所属的产品类目，所述产品类目可以根据涵盖范围的不同分为多个。例如，对于某品牌特定型号的鼠标而言，该特定型号的鼠标可以拥有一个唯一的编号。通过该编号可以查询该特定型号的鼠标的产地，制作材料，颜色，重量，接口，价格等一系列的产品信息。另外，该鼠标还可以拥有多个产品类目，例如鼠标，电脑配件，电子产品。所述的产品编号以及产品类目中的至少一个可以作为该产品的标识包含在该产品的产品信息中。服务器获取的产品信息可以存储在数据库中，该数据库可以被其他实体或者逻辑体访问。在实际应用中，本申请实施例可以采用MySQL+iBATIS技术进行产品信息的存储和读取。本申请实施例在后续对存储的产品信息进行处理时，可以利用iBATIS的ORM（Object Relational Mapping，对象关系映射）技术对存储的产品信息进行读取。

S200：根据预设规则提取所述属性信息中的关键属性信息。

本申请实施例中的服务器可以根据产品类别的不同，从众多的属性信息中提取关键的属性信息，并向用户展示所述提取的关键属性信息。例如，对于鼠标这类产品，服务器可以从众多的属性信息中提取接口，重量，颜色，有线/无线这些关键属性信息；对于葡萄酒这类产品，服务器可以提取原产地，年份这些关键属性信息；对于U盘这类产品，服务器可以提取容量，接口，制作材料这些关键属性信息。具体地，本申请实施例可以通过机器学习的方法提取所述产品信息中的关键属性信息。下面首先介绍利用机器学习方法提取产品信息中的

关键属性信息的第一实现方式:

用户在搜索产品信息时,一般可以直接输入产品的类目进行搜索,还可以输入产品的类目加上自己比较关注的属性信息进行搜索。例如,用户 A 在搜索产品 MP3 的信息时,可以仅仅输入 MP3 进行搜索,也可以输入 MP3 8GB 进行搜索。对于后者,8GB 代表了用户对 MP3 的容量比较关注。本申请实施例可以针对每类产品,获取预设数量的携带属性信息的产品搜索指令,然后可以统计所述获取的产品搜索指令中各个属性信息所占的比重值,并且通过设置第一阈值,将所述比重值大于或者等于所述第一阈值的属性信息确定为所述产品的关键属性信息。例如,本申请可以获取 10000 条有关 MP3 的携带属性信息的搜索指令,并且统计出这些搜索指令中包括:携带 MP3 容量的搜索指令有 7000 条,携带 MP3 颜色的搜索指令有 1000 条,携带 MP3 接口的搜索指令有 5000 条。那么,本申请实施例可以计算出 MP3 容量这个属性信息的比重值为 $7000/10000=0.7$,同样地,可以计算出 MP3 颜色这个属性信息的比重值为 0.1,可以计算出 MP3 接口这个属性信息的比重值为 0.5。接着,本申请实施例可以将 0.3 设置为第一阈值,并将比重值大于或者等于 0.3 的属性信息提取出来,例如 MP3 的容量和接口,作为 MP3 这个产品的关键属性信息。

在本申请另一实施例中,还可以将上述获取的关键属性信息作为参考的样本,通过计算所述产品的各个属性信息与上述获取的关键属性信息的相似度,设置第二阈值,并将所述相似度大于或者等于该第二阈值的属性信息确定为所述产品的关键属性信息。

举个例子来详细说明如何计算两个属性信息的相关度:

用户在利用产品的属性信息进行搜索的时候,可以根据搜索结果对该属性信息进行评估。假设 A 为根据上述步骤初步提取的关键属性信息,B 为需要计算相关度的属性信息。对于同时关注该类产品的属性信息 A 和属性信息 B 的用户群 C,可以获取用户群 C 中每个用户对该类产品中这两个属性信息的评分情况,现在便可以通过 Pearson 公式计算属性信息 A 与属性信息 B 的相关度,具体如式 1 所示:

$$\text{sim}(a,b) = \frac{\sum_{u \in C} (R_{ua} - \bar{R}_a)(R_{ub} - \bar{R}_b)}{\sqrt{\sum_{u \in C} (R_{ua} - \bar{R}_a)^2} \sqrt{\sum_{u \in C} (R_{ub} - \bar{R}_b)^2}} \quad \text{式 1}$$

其中, $\text{sim}(a,b)$ 代表属性信息 A 与属性信息 B 的相关度, R_{ua} 代表用户群 C 中的用户 u 对属性信息 A 的评分, R_{ub} 代表用户群 C 中的用户 u 对属性信息 B 的评分, $\bar{R}_a$ 代表用户群 C 中所有用户对属性信息 A 的平均评分, $\bar{R}_b$ 代表用户群 C 中所有用户对属性信息 B 的平均评分。

根据上述方法,可以计算得到该产品的属性信息与部分关键属性信息参考样本的相关度,然后可以设置第二阈值,将相关度大于或者等于第二阈值的属性信息确定为该产品的关键属性信息。

本申请另一实施例还可以通过机器学习方法中的支持向量机 (SVM, Support Vector

Machine) 来提取产品信息中的关键属性信息。下面介绍本申请实施例利用机器学习方法提取产品信息中的关键属性信息的第二实现方式:

本申请实施例可以将用户群对某类产品中部分属性信息的评分作为特征向量组,建立训练样本集。然后对该训练样本集进行机器学习,计算得到划分该训练样本集中特征向量组的归类条件。该归类条件能够区分该训练样本集中用户关注的属性信息对应的特征向量组 and 用户不关注的属性信息对应的特征向量组。得到所述归类条件后,便可以利用该归类条件对该类产品中的属性信息进行归类,最终可以得到该类产品的产品信息中用户关注的关键属性信息。

这里以采用支持向量机为例介绍计算所述训练样本中特征向量组的归类条件的实现方式。支持向量机是 Cortes 和 Vapnik 于 1995 年首先提出的,它在解决小样本、非线性及高维模式识别中表现出许多特有的优势,并能够推广应用到函数拟合等其他机器学习问题中。整体来说,支持向量机可以解决复杂事务的分类及分类标准的问题。

利用图 2 显示的线性分类的例子解释通过支持向量机进行分类的基本原理。如图 2 所示,左侧坐标图中的点表示输入的训练样本,右侧坐标图中的叉代表的点表示计算得到的 C1 类训练样本,圆圈代表的点表示计算得到的 C2 类训练样本。如图 2 所示,将训练样本通过支持向量机算法计算后,可以获得分类后的 C1 和 C2 两类训练样本,并且可以得到划分 C1 和 C2 两类的归类条件。

对于图 2 的线性分类来说,所述归类条件(图中的 vv' 线,也称为超平面)可以用一个线性函数来表示,例如表示为:

$$f(x)=wx+b \quad \text{式 2}$$

其中, w 和 b 为支持向量机对特征值向量集合进行计算(支持向量机中称为“训练”)后得到的参数, x 代表特征值向量。

$f(x)$ 表示支持向量机中的映射关系。对于 $f(x)=0$ 的情况,此时的特征值向量 x 即位于所述超平面上。对于 $f(x)$ 大于 0 的情况,对应图 2 右侧坐标图中超平面右上侧的特征值向量;对于 $f(x)$ 小于 0 的情况,对应图 2 右侧坐标图中超平面左下侧的特征值向量。

输入的特征值向量例如均为二维向量,即对应图 2 中坐标上的每个点。支持向量机算法,即不断搜索输入的特征值向量范围内的直线,通过尝试计算每一个搜索到的这种直线与每一特征值向量(图中的点)的距离,得到一个这样的直线:该直线距离两侧最近特征值向量的距离最大且相等。如图 2 中右侧坐标图所示,计算得到的直线 vv' 即超平面。从图 2 中右侧坐标图可以看出,二维情况下超平面 vv' 为一直线,该直线距离两侧最近特征值向量的距离最大且相等,该距离均为 L 。

当输入的特征值向量为三维向量时，处理过程与特征值向量为二维向量的情况类似。图 3 为本申请一个例子中以线性分类为例解释通过支持向量机算法进行三维分类的基本原理示意图。如图 3 所示，左侧坐标图中的点表示输入的训练样本，右侧坐标图中的叉代表的点表示计算得到的 C1 类训练样本，圆圈代表的点表示计算得到的 C2 类训练样本。如图 3 所示，将训练样本通过支持向量机算法计算后，可以获得分类后的 C1 和 C2 两类训练样本，并且可以得到划分 C1 和 C2 两类的归类条件。

对于图 3 的线性分类来说，所述归类条件（图中的 vv' 面，也称为超平面）可以用一个线性函数来表示，例如表示为：

$$f(x) = wx + b \quad \text{式 3}$$

其中， w 和 b 为支持向量机对特征值向量集合进行计算（支持向量机中称为“训练”）后得到的参数， x 代表特征值向量。

这里需要说明的是，由于这是针对三维向量的情况，因此式 3 中的参数 w 和 b 均为三维向量。

$f(x)$ 表示支持向量机中的映射关系。对于 $f(x) = 0$ 的情况，此时的特征值向量 x 即位于所述超平面上。对于 $f(x)$ 大于 0 的情况，对应图 3 右侧坐标图中超平面右上侧的特征值向量；对于 $f(x)$ 小于 0 的情况，对应图 3 右侧坐标图中超平面左下侧的特征值向量。

输入的特征值向量例如均为三维向量，即对应图 3 中坐标上的每个点，每个点均由 X , Y , Z 三个坐标确定。此时支持向量机算法可以不断搜索输入的特征值向量范围内的平面，通过尝试计算每一个搜索到的这种平面与每一特征值向量的距离，得到一个这样的平面：该平面距离两侧最近特征值向量的距离最大且相等。如图 3 右侧坐标图所示，计算得到的平面 vv' 即超平面。该平面距离两侧最近特征值向量的距离最大且相等。

对于输入的多个特征值向量，可以计算得到这些特征值向量的两种训练样本类型，以及这两种训练样本类型的归类条件。

在本申请实施例中，可以选取用户群中 N 个用户，并且将该 N 个用户对该类产品中各个属性信息的评分作为特征值向量。例如， N 个用户对 MP3 的产品信息中容量这个属性信息的评分依次为 4.0, 5.1, 4.8, ..., 5.1。那么可以将这 N 个用户的评分按顺序排列，形成特征值向量 (4.0, 5.1, 4.8, ..., 5.1)。同理，可以将这 N 个用户对 MP3 的产品信息中原产地，颜色，制作材料等属性信息的评分作为其他特征值向量，并将这多个特征值向量作为训练样本集。由于用户对于自身关注的属性信息往往会给予较高的评分，而对于自身不关注的属性信息往往会给予较低的评分，那么在训练样本集中的特征值向量之间会存在比较明显的差异。本申请实施例可以利用支持向量机对这种差异进行上述的计算过程，这样可以得到一个

归类条件。该归类条件可以区分该训练样本集中用户关注的属性信息对应的特征向量组 and 用户不关注的属性信息对应的特征向量组。利用该归类条件，便可以将该类产品的属性信息划分为用户关注和用户不关注的属性信息。本申请实施例可以提取用户关注的属性信息，作为该类产品的关键属性信息。

S300: 提取产品信息中的产品标识并建立所述产品标识与所述关键属性信息的关联关系。

本申请实施例中的服务器在提取产品信息的关键属性信息后，可以建立该产品与所述关键属性信息的关联关系。具体地，本申请实施例可以提取产品信息中的产品标识，并建立该标识与所述关键属性信息的关联关系。所述的产品标识可以包括产品编号和产品类目。其中，产品编号与产品一一对应，通过产品编号可以查询到对应的产品；产品类目可以有根据涵盖范围的不同分为多个，例如对于某品牌的特定型号的鼠标这个产品而言，鼠标、电脑配件以及电子产品均可以作为它的产品类目，这三个产品类目涵盖的范围逐级增大。从产品信息中提取出产品标识后，可以建立所述产品标识与所述关键属性信息的关联关系。通过产品标识可以查询与该产品标识相关联的关键属性信息。例如，对于某品牌特定型号的 MP3 这个产品而言，其产品编号例如可以为 201502031331，其产品类别例如可以为 MP3，随身播放器，电子产品，其关键属性信息例如可以为容量、大小、重量、颜色。本申请实施例可以按照该 MP3 产品标识涵盖范围从大到小的顺序，建立该 MP3 的产品标识和它的关键属性信息的关联关系，例如可以为：电子产品-随身播放器-MP3-201502031331-容量、大小、重量、颜色。这样，通过该 MP3 的产品标识，便可以查询与该产品标识相关联的关键属性信息。具体地，可以将产品标识作为 key，将该产品的关键属性信息作为 value 进行存储。通过产品的标识可以查询出与该标识相关联的关键属性信息，这样便可以建立所述产品标识与所述关键属性信息的关联关系。

S400: 获取用户搜索指令并提取所述搜索指令中的产品标识。

用户在搜索自身感兴趣的产品时，可以在网站的搜索框中输入搜索指令。该搜索指令可以为具体的产品编号（例如网站的运营人员为了测试网站性能，直接输入产品的编号进行检测），也可以为产品的类目（一般用户在进行搜索时往往会输入产品的类目）。例如，用户 A 想购买一款鼠标，他往往会在网站的搜索框中输入“鼠标”这个搜索指令。服务器获取到用户的搜索指令后，可以提取所述搜索指令中的产品标识。

S500: 根据所述建立的产品标识与关键属性信息的关联关系，查询与所述提取的产品标识相关联的关键属性信息；

S600: 向客户端发送所述查询的关键属性信息。

服务器提取出用户搜索指令中的产品标识后，本申请实施例可以根据所述建立的产品标识与关键属性信息的关联关系，查询所述提取的产品标识相对应的关键属性信息，并向客户端发送所述查询的关键属性信息。具体地，服务器可以预先设置网页模板，该网页模板可以是一个页面框架，页面中产品的关键属性信息均未填充。服务器提取出用户搜索指令中的产品标识后，可以通过预先建立的产品标识与关键属性信息的对应关系，查询出该产品标识对应的关键属性信息。例如，用户的搜索指令为“鼠标”，服务器从该搜索指令中提取出的产品标识为“鼠标”，那么便可以根据产品标识与关键属性信息的关联关系，查询到产品标识为鼠标的一系列产品的关键属性信息，这些产品例如可以是不同品牌、不同型号的有线鼠标、无线鼠标。每款产品均具有一系列的关键属性信息，例如接口，重量，颜色，制作材料等。

服务器查询到与提取的产品标识相关联的关键属性信息后，可以将这些关键属性信息填充到上述的网页模板中，这样便可以形成最终的搜索页面。服务器可以将该搜索页面发送至客户端，供用户浏览。

服务器根据用户搜索指令中的产品标识进行关键属性信息的查询时，有时会查询到处于不同产品类目中的关键属性信息。例如，用户提供的搜索指令为“苹果”，那么服务器从用户的搜索指令中提取的产品标识就可以为“苹果”。当服务器进行查询与“苹果”相关联的关键属性信息时，可以查询出水果行业和电子商品行业的关键属性信息。其中，水果行业的“苹果”就属于水果类目，而电子商品行业中的“苹果”手机则属于电子商品类目。此时，服务器可以根据用户的偏好，对查询出的关键属性信息进行筛选。鉴于此，在本申请一优选实施例中，可以在步骤 S500 与 S600 之间增加下述步骤：

S501：根据用户偏好，对所述查询的关键属性信息进行筛选。

具体地，本申请实施例的服务器可以分析预设时间段内用户的搜索记录，确定用户关注度较高的产品类目。例如，服务器可以分析最近一个月内用户的搜索记录，并且提取每条搜索记录中的产品标识。通过提取的产品标识，服务器可以统计出该用户经常搜索的产品类目，从而确定出该用户关注度较高的产品类目。当服务器根据用户搜索指令中的产品标识查询到多个产品类目中的关键属性信息时，可以将用户关注度较低的产品类目的关键属性信息去除，并将用户关注度较高的产品类目的关键属性信息发送给客户端。具体地，服务器同样可以预设网页模板，并将用户关注度较高的产品类目的关键属性信息填充至预设的网页模板中，并将填充关键属性信息的网页发送至客户端供用户浏览。

本申请实施例还提供一种产品信息处理的服务器。图 4 为本申请一实施例提供的一种产品信息处理的服务器功能模块图。如图 4 所示，所述服务器包括：

产品信息获取单元 100，用来获取产品信息，所述产品信息包括产品标识和属性信息；

关键属性信息提取单元 200, 用来根据预设规则提取所述属性信息中的关键属性信息;

关联关系建立单元 300, 用来提取所述产品信息中的产品标识, 并建立所述产品标识与所述关键属性信息的关联关系;

产品标识获取单元 400, 用来获取用户搜索指令并提取所述搜索指令中的产品标识;

关键属性信息查询单元 500, 用来根据所述建立的产品标识与关键属性信息的关联关系, 查询与所述提取的产品标识相关联的关键属性信息;

发送单元 600, 用来向客户端发送所述查询的关键属性信息。

在本申请一优选实施例中, 所述关键属性信息提取单元 200 具体可以包括:

搜索指令获取模块 201, 用来获取预设数量的携带属性信息的产品搜索指令;

比重值计算模块 202, 用来计算所述获取的产品搜索指令中各个属性信息所占的比重值;

第一阈值设置模块 203, 用来设置第一阈值;

第一关键属性信息确定模块 204, 用来将所述比重值大于或者等于所述第一阈值的属性信息确定为所述产品的关键属性信息。

在本申请另一优选实施例中, 所述关键属性信息提取单元 200 具体还可以包括:

参考样本设置模块 205, 用来将所述确定的关键属性信息作为参考样本;

相似度计算模块 206, 用来计算所述产品的各个属性信息与所述参考样本中关键属性信息的相似度;

第二阈值设置模块 207, 用来设置第二阈值;

第二关键属性信息确定模块 208, 用来将所述相似度大于或者等于第二阈值的属性信息确定为所述产品的关键属性信息。

在本申请另一优选实施例中, 所述关键属性信息提取单元 200 具体可以包括:

训练样本集建立模块 210, 用来建立训练样本集, 所述训练样本集包括关键属性信息与非关键属性信息的特征值向量;

归类条件计算模块 220, 用来计算所述包括关键属性信息与非关键属性信息的特征值向量的训练样本集的归类条件;

关键属性信息获取模块 230, 用来利用所述归类条件, 对所述产品信息中的属性信息进行归类, 得到所述属性信息中的关键属性信息。

在本申请另一优选实施例中, 所述发送单元 600 具体包括:

网页模板设置模块 601, 用来预先设置网页模板;

填充模块 602, 用来将所述查询的关键属性信息填充至所述网页模板中;

发送模块 603, 用来向客户端发送所述填充关键属性信息后的网页。

图 5 为本申请另一实施例提供的一种产品信息处理的服务器功能模块图。如图 5 所示，所述服务器还包括：

筛选单元 501，用来根据用户偏好，对所述关键属性信息查询单元发来的关键属性信息进行筛选，并将筛选后的关键属性信息发送给所述发送单元。

本申请实施例提供的一种产品信息处理的方法及服务器，通过对产品属性信息进行分析，针对不同的产品，提取出该产品的关键属性信息，并根据用户的搜索指令，将与用户的搜索指令相关的产品的关键属性信息发送给客户端供用户浏览，能够给用户真正关注的产品信息。

在20世纪90年代，对于一个技术的改进可以很明显地区分是硬件上的改进（例如，对二极管、晶体管、开关等电路结构的改进）还是软件上的改进（对于方法流程的改进）。然而，随着技术的发展，当今的很多方法流程的改进已经可以视为硬件电路结构的直接改进。设计人员几乎都通过将改进的方法流程编程到硬件电路中来得到相应的硬件电路结构。因此，不能说一个方法流程的改进就不能用硬件实体模块来实现。例如，可编程逻辑器件

（Programmable Logic Device, PLD）（例如现场可编程门阵列（Field Programmable Gate Array, FPGA））就是这样一种集成电路，其逻辑功能由用户对器件编程来确定。由设计人员自行编程来把一个数字系统“集成”在一片PLD上，而不需要请芯片制造厂商来设计和制作专用的集成电路芯片²。而且，如今，取代手工地制作集成电路芯片，这种编程也多半改用“逻辑编译器（logic compiler）”软件来实现，它与程序开发撰写时所用的软件编译器相类似，而要编译之前的原始代码也得用特定的编程语言来撰写，此称之为硬件描述语言（Hardware Description Language, HDL），而HDL也并非仅有一种，而是有许多种，如ABEL（Advanced Boolean Expression Language）、AHDL（Altera Hardware Description Language）、Confluence、CUPL（Cornell University Programming Language）、HDCal、JHDL（Java Hardware Description Language）、Lava、Lola、MyHDL、PALASM、RHDH（Ruby Hardware Description Language）等，目前最普遍使用的是VHDL（Very-High-Speed Integrated Circuit Hardware Description Language）与Verilog²。本领域技术人员也应该清楚，只需要将方法流程用上述几种硬件描述语言稍作逻辑编程并编程到集成电路中，就可以很容易得到实现该逻辑方法流程的硬件电路。

控制器可以按任何适当的方式实现，例如，控制器可以采取例如微处理器或处理器以及存储可由该（微）处理器执行的计算机可读程序代码（例如软件或固件）的计算机可读介质、逻辑门、开关、专用集成电路（Application Specific Integrated Circuit, ASIC）、可

编程逻辑控制器和嵌入微控制器的形式，控制器的例子包括但不限于以下微控制器：ARC 625D、Atmel AT91SAM、Microchip PIC18F26K20 以及Silicone Labs C8051F320，存储器控制器还可以被实现为存储器的控制逻辑的一部分。

本领域技术人员也知道，除了以纯计算机可读程序代码方式实现控制器以外，完全可以通过将方法步骤进行逻辑编程来使得控制器以逻辑门、开关、专用集成电路、可编程逻辑控制器和嵌入微控制器等的形式来实现相同功能。因此这种控制器可以被认为是一种硬件部件，而对其内包括的用于实现各种功能的装置也可以视为硬件部件内的结构。或者甚至，可以将用于实现各种功能的装置视为既可以是实现方法的软件模块又可以是硬件部件内的结构。

上述实施例阐明的系统、装置、模块或单元，具体可以由计算机芯片或实体实现，或者由具有某种功能的产品来实现。

为了描述的方便，描述以上装置时以功能分为各种单元分别描述。当然，在实施本申请时可以把各单元的功能在同一个或多个软件和/或硬件中实现。

通过以上的实施方式的描述可知，本领域的技术人员可以清楚地了解到本申请可借助软件加必需的通用硬件平台的方式来实现。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，该计算机软件产品可以存储在存储介质中，如ROM/RAM、磁碟、光盘等，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等等）执行本申请各个实施例或者实施例的某些部分所述的方法。

本说明书中的各个实施例均采用递进的方式描述，各个实施例之间相同相似的部分互相参见即可，每个实施例重点说明的都是与其他实施例的不同之处。尤其，对于系统实施例而言，由于其基本相似于方法实施例，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

本申请可用于众多通用或专用的计算机系统环境或配置中。例如：个人计算机、服务器计算机、手持设备或便携式设备、平板型设备、多处理器系统、基于微处理器的系统、置顶盒、可编程的消费电子设备、网络PC、小型计算机、大型计算机、包括以上任何系统或设备的分布式计算环境等等。

本申请可以在由计算机执行的计算机可执行指令的一般上下文中描述，例如程序模块。一般地，程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等等。也可以在分布式计算环境中实践本申请，在这些分布式计算环境中，由通过通信网络而被连接的远程处理设备来执行任务。在分布式计算环境中，程序模块可以位于包

括存储设备在内的本地和远程计算机存储介质中。

虽然通过实施例描绘了本申请，本领域普通技术人员知道，本申请有许多变形和变化而不脱离本申请的精神，希望所附的权利要求包括这些变形和变化而不脱离本申请的精神。

权利要求

- 1、一种产品信息处理的方法，其特征在于，包括：
获取产品信息，所述产品信息包括产品标识和属性信息；
根据预设规则提取所述属性信息中的关键属性信息；
提取所述产品信息中的产品标识，并建立所述产品标识与所述关键属性信息的关联关系；
获取用户搜索指令并提取所述搜索指令中的产品标识；
根据所述建立的产品标识与所述关键属性信息的关联关系，查询与所述提取的产品标识相关联的关键属性信息；
向客户端发送所述查询的关键属性信息。
- 2、如权利要求 1 所述的一种产品信息处理的方法，其特征在于，所述根据预设规则提取所述属性信息中的关键属性信息具体包括：
获取预设数量的携带属性信息的产品搜索指令；
计算所述获取的产品搜索指令中各个属性信息所占的比重值；
设置第一阈值；
将所述比重值大于或者等于所述第一阈值的属性信息确定为所述产品的关键属性信息。
- 3、如权利要求 2 所述的一种产品信息处理的方法，其特征在于，所述根据预设规则提取所述属性信息中的关键属性信息还包括：
将所述确定的关键属性信息作为参考样本；
计算所述产品的各个属性信息与所述参考样本中关键属性信息的相似度；
设置第二阈值；
将所述相似度大于或者等于第二阈值的属性信息确定为所述产品的关键属性信息。
- 4、如权利要求 1 所述的一种产品信息处理的方法，其特征在于，所述根据预设规则提取所述属性信息中的关键属性信息具体包括：
建立训练样本集，所述训练样本集包括关键属性信息与非关键属性信息的特征值向量；
计算所述包括关键属性信息与非关键属性信息的特征值向量的训练样本集的归类条件；
利用所述归类条件，对所述产品信息中的属性信息进行归类，得到所述属性信息中的关键属性信息。
- 5、如权利要求 1 所述的一种产品信息处理的方法，其特征在于，所述向客户端发送所述查询的关键属性信息具体包括：

预先设置网页模板；

将所述查询的关键属性信息填充至所述网页模板中；

向客户端发送填充关键属性信息后的网页。

6、如权利要求 1 所述的一种产品信息处理的方法，其特征在于，在所述向客户端发送所述查询的关键属性信息步骤之前，所述方法还包括：

根据用户偏好，对所述查询的关键属性信息进行筛选。

7、一种产品信息处理的服务器，其特征在于，包括：产品信息获取单元，关键属性信息提取单元，关联关系建立单元，产品标识获取单元，关键属性信息查询单元，发送单元，其中：

所述产品信息获取单元，用来获取产品信息，所述产品信息包括产品标识和属性信息；

所述关键属性信息提取单元，用来根据预设规则提取所述属性信息中的关键属性信息；

所述关联关系建立单元，用来提取所述产品信息中的产品标识，并建立所述产品标识与所述关键属性信息的关联关系；

所述产品标识获取单元，用来获取用户搜索指令并提取所述搜索指令中的产品标识；

所述关键属性信息查询单元，用来根据所述建立的产品标识与关键属性信息的关联关系，查询与所述提取的产品标识相关联的关键属性信息；

所述发送单元，用来向客户端发送所述查询的关键属性信息。

8、如权利要求 7 所述的一种产品信息处理的服务器，其特征在于，所述关键属性信息提取单元具体包括：

搜索指令获取模块，用来获取预设数量的携带属性信息的产品搜索指令；

比重值计算模块，用来计算所述获取的产品搜索指令中各个属性信息所占的比重值；

第一阈值设置模块，用来设置第一阈值；

第一关键属性信息确定模块，用来将所述比重值大于或者等于所述第一阈值的属性信息确定为所述产品的关键属性信息。

9、如权利要求 8 所述的一种产品信息处理的服务器，其特征在于，所述关键属性信息提取单元还包括：

参考样本设置模块，用来将所述确定的关键属性信息作为参考样本；

相似度计算模块，用来计算所述产品的各个属性信息与所述参考样本中关键属性信息的相似度；

第二阈值设置模块，用来设置第二阈值；

第二关键属性信息确定模块，用来将所述相似度大于或者等于第二阈值的属性信息确定

为所述产品的关键属性信息。

10、如权利要求 7 所述的一种产品信息处理的服务器，其特征在于，所述关键属性信息提取单元具体包括：

训练样本集建立模块，用来建立训练样本集，所述训练样本集包括关键属性信息与非关键属性信息的特征值向量；

归类条件计算模块，用来计算所述包括关键属性信息与非关键属性信息的特征值向量的训练样本集的归类条件；

关键属性信息获取模块，用来利用所述归类条件，对所述产品信息中的属性信息进行归类，得到所述属性信息中的关键属性信息。

11、如权利要求 7 所述的一种产品信息处理的服务器，其特征在于，所述发送单元具体包括：

网页模板设置模块，用来预先设置网页模板；

填充模块，用来将所述查询的关键属性信息填充至所述网页模板中；

发送模块，用来向客户端发送填充关键属性信息后的网页。

12、如权利要求 7 所述的一种产品信息处理的服务器，其特征在于，所述服务器还包括：

筛选单元，用来根据用户偏好，对所述关键属性信息查询单元发来的关键属性信息进行筛选，并将筛选后的关键属性信息发送给所述发送单元。

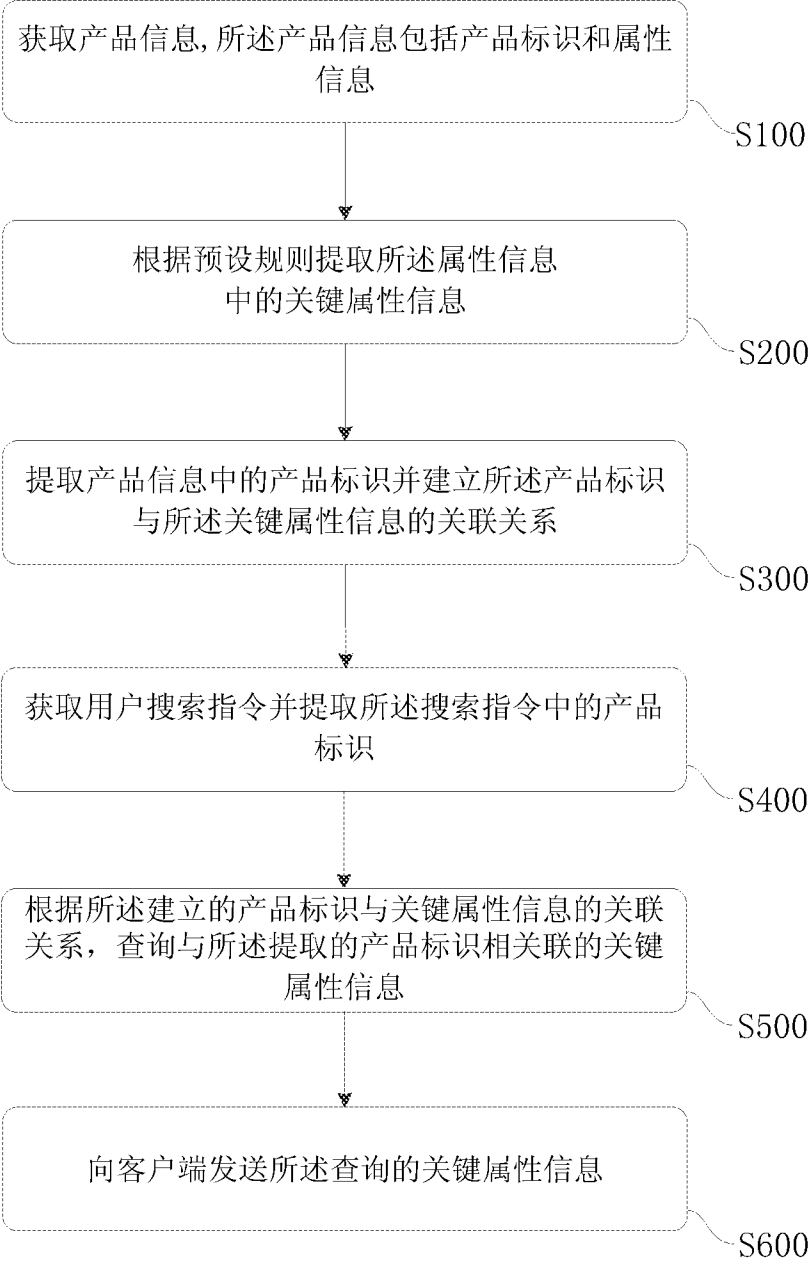


图 1

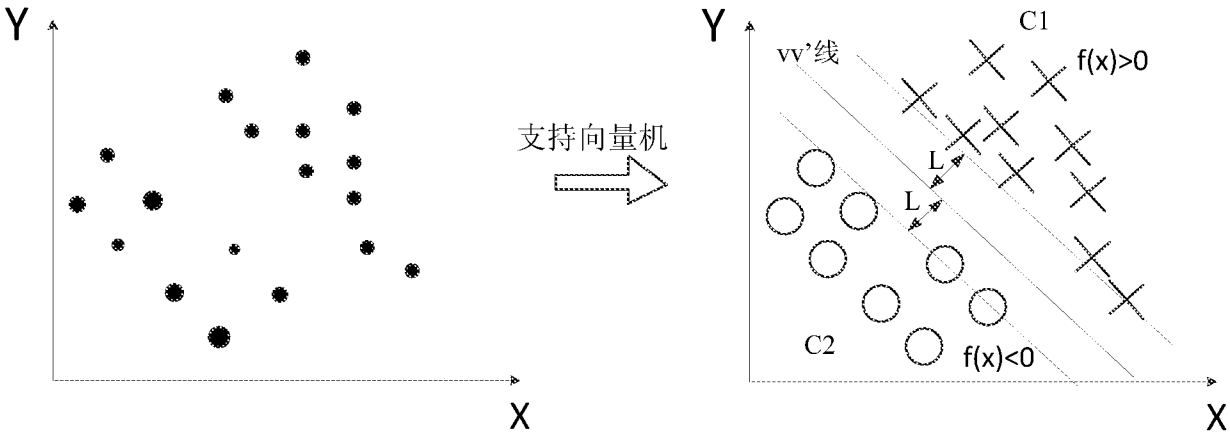


图 2

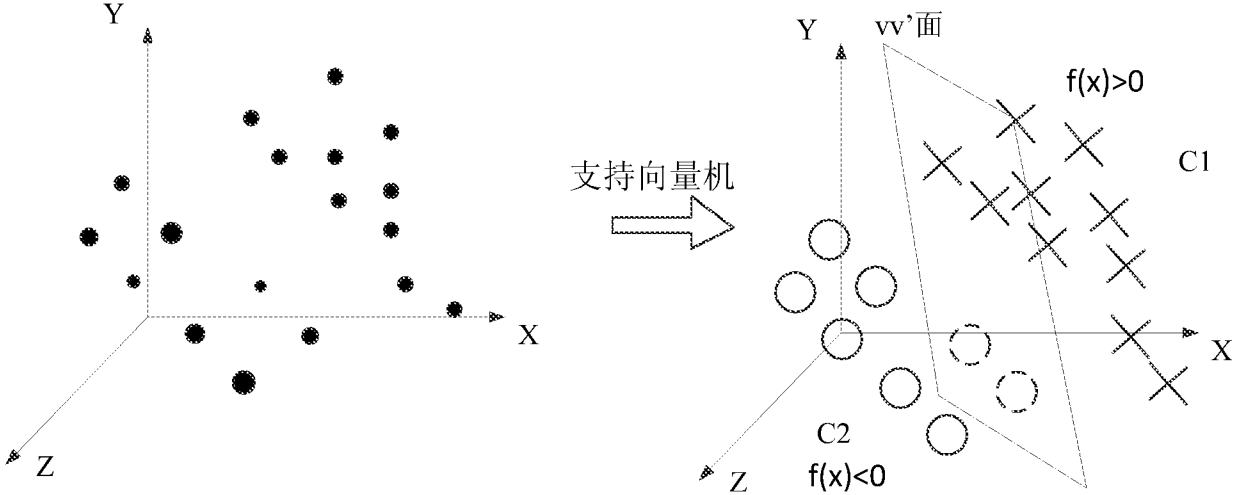


图 3

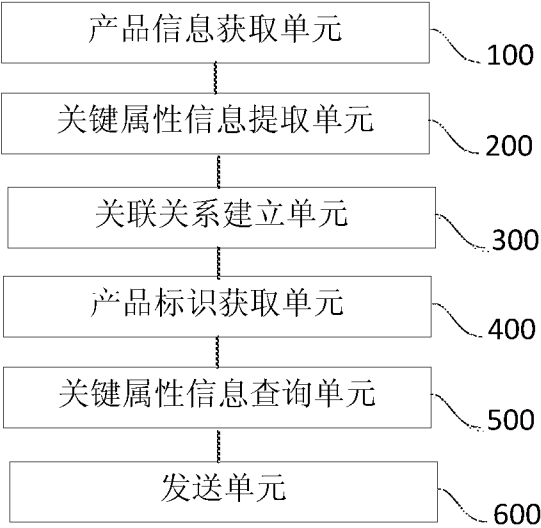


图 4

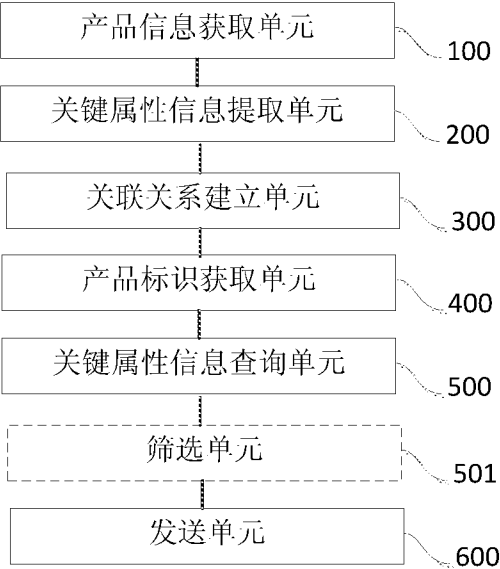


图 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/075379

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; 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)

CNPAT, CNKI, WPI, EPODOC: article, data, ID, information, goods, product, item, commodity, relat+, key, identif+, attribute?, property, database, parameter

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 102456165 A (IBM) 16 May 2012 (16.05.2012) description, paragraphs [0024]-[0049], and figure 1	1-12
Y	CN 102402535 A (ALIBABA GROUP HOLDING LTD.) 04 April 2012 (04.04.2012) description, paragraphs [0048]-[0079]	1-12
Y	CN 104040540 A (TYCO FIRE & SECURITY GMBH) 10 September 2014 (10.09.2014) description, paragraphs [0020]-[0040] and figure 3	1-12
Y	CN 102402535 A (ALIBABA GROUP HOLDING LTD.) 04 April 2012 (04.04.2012) description, paragraphs [0048]-[0079]	1-12
A	CN 102890686 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 23 January 2013 (23.01.2013) the whole document	1-12

☐ 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 14 April 2016	Date of mailing of the international search report 24 May 2016
Name and mailing address of the ISA 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 ZHANG, Qian Telephone No. (86-10) 62413681

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/075379

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 102456165 A	16 May 2012	US 2012110647 A1	03 May 2012
CN 102402535 A	04 April 2012	HK 1166395 A0	26 October 2012
CN 104040540 A	10 September 2014	EP 2748727 A1	02 July 2014
		KR 20140064910 A	28 May 2014
		AU 2012298827 A1	10 April 2014
		US 2013048721 A1	28 February 2013
		WO 2013028846 A1	28 February 2013
		CA 2849882 A1	28 February 2013
CN 102890686 A	23 January 2013	WO 2013010416 A1	24 January 2013
		US 2014136524 A1	15 May 2014

国际检索报告

国际申请号

PCT/CN2016/075379

A. 主题的分类

G06F 17/30(2006.01)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)

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))

CNPAT, CNKI, WPI, EPODOC: 产品, 商品、货品, 物品, 货物, 信息, 数据, 参数, 关键, 核心, 主要, 标识, ID, information, goods, product, item, commodity, relat+, key, identif+, attribute?, property, database, parameter

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
Y	CN 102456165 A (国际商业机器公司) 2012年 5月 16日 (2012 - 05 - 16) 说明书第[0024]-[0049]段, 图1	1-12
Y	CN 102402535 A (阿里巴巴集团控股有限公司) 2012年 4月 4日 (2012 - 04 - 04) 说明书第[0048]-[0079]段	1-12
Y	CN 104040540 A (泰科消防及安全有限公司) 2014年 9月 10日 (2014 - 09 - 10) 说明书第[0020]-[0040]段, 图3	1-12
Y	CN 102402535 A (阿里巴巴集团控股有限公司) 2012年 4月 4日 (2012 - 04 - 04) 说明书第[0048]-[0079]段	1-12
A	CN 102890686 A (腾讯科技深圳有限公司) 2013年 1月 23日 (2013 - 01 - 23) 全文	1-12

☐ 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

2016年 4月 14日

国际检索报告邮寄日期

2016年 5月 24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

授权官员

张千

传真号 (86-10)62019451

电话号码 (86-10)62413681

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/075379

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	102456165	A	2012年 5月 16日	US	2012110647	A1	2012年 5月 3日
CN	102402535	A	2012年 4月 4日	HK	1166395	A0	2012年 10月 26日
CN	104040540	A	2014年 9月 10日	EP	2748727	A1	2014年 7月 2日
				KR	20140064910	A	2014年 5月 28日
				AU	2012298827	A1	2014年 4月 10日
				US	2013048721	A1	2013年 2月 28日
				WO	2013028846	A1	2013年 2月 28日
				CA	2849882	A1	2013年 2月 28日
CN	102890686	A	2013年 1月 23日	WO	2013010416	A1	2013年 1月 24日
				US	2014136524	A1	2014年 5月 15日

表 PCT/ISA/210 (同族专利附件) (2009年7月)