



(12) 发明专利

(10) 授权公告号 CN 112818221 B

(45) 授权公告日 2024.06.14

(21) 申请号 202110104022.2

G06F 40/284 (2020.01)

(22) 申请日 2021.01.26

G06F 16/35 (2019.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 112818221 A

(56) 对比文件

CN 111949898 A, 2020.11.17

CN 101262494 A, 2008.09.10

(43) 申请公布日 2021.05.18

审查员 王克

(73) 专利权人 北京百度网讯科技有限公司

地址 100085 北京市海淀区上地十街10号

百度大厦2层

(72) 发明人 朱嘉琪 卢佳俊 柴春光

(74) 专利代理机构 北京清亦华知识产权代理事

务所(普通合伙) 11201

专利代理师 韩海花

(51) Int. Cl.

G06F 16/9535 (2019.01)

G06F 40/30 (2020.01)

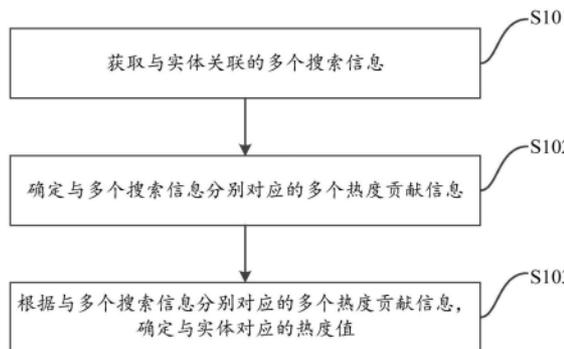
权利要求书2页 说明书10页 附图5页

(54) 发明名称

实体的热度确定方法、装置、电子设备及存储介质

(57) 摘要

本申请公开了实体的热度确定方法、装置、电子设备及存储介质,涉及计算机技术领域,具体涉及深度学习、知识图谱、自然语言处理等人工智能技术领域。具体实施方案为:获取与实体关联的多个搜索信息;确定与多个搜索信息分别对应的多个热度贡献信息;以及根据与多个搜索信息分别对应的多个热度贡献信息,确定与实体对应的热度值,从而利用少量的特征即能够准确地确定出实体的热度值,有效提升实体热度确定的合理性和效率,能够有效地提升实体的热度分析效果。



1. 一种实体的热度确定方法,包括:
 - 获取与实体关联的多个搜索信息;
 - 确定与所述多个搜索信息分别对应的多个热度贡献信息;以及
 - 根据与所述多个搜索信息分别对应的多个热度贡献信息,确定与所述实体对应的热度值;
 - 所述确定与所述多个搜索信息分别对应的多个热度贡献信息,包括:
 - 确定与所述多个搜索信息分别对应的多个搜索结果;
 - 确定与所述多个搜索结果分别对应的多个搜索需求信息;以及
 - 根据所述多个搜索需求信息,确定与所述多个搜索信息分别对应的多个热度贡献信息;
 - 所述根据所述多个搜索需求信息,确定与所述多个搜索信息分别对应的多个热度贡献信息,包括:
 - 确定与所述多个搜索结果分别对应的多个页面浏览量;
 - 根据与所述搜索结果对应的搜索需求信息,和与其对应的所述页面浏览量确定与所述搜索信息对应的热度贡献信息。
2. 根据权利要求1所述的方法,其中,所述确定与所述多个搜索结果分别对应的多个搜索需求信息,包括:
 - 确定与所述多个搜索结果分别对应的多个搜索需求比例值,并将所述搜索需求比例值作为所述搜索需求信息;
 - 其中,所述搜索需求比例值,是采用预设函数对所述搜索结果在搜索结果页面中的显示位置数据进行量化处理后得到的函数值。
3. 根据权利要求2所述的方法,其中,所述根据与所述搜索结果对应的搜索需求信息,和与其对应的所述页面浏览量确定与所述搜索信息对应的热度贡献信息,包括:
 - 如果所述搜索需求信息是搜索需求比例值,则确定所述搜索需求比例值与所述页面浏览量的乘积值,并将所述乘积值作为与所述搜索信息对应的热度贡献信息。
4. 根据权利要求1所述的方法,其中,所述获取与实体关联的多个搜索信息,包括:
 - 确定与所述实体对应的行业领域类型;
 - 获取与所述行业领域类型对应的搜索信息集合,并将所述搜索信息集合之中的多个搜索信息作为所述关联的多个搜索信息;
 - 其中,所述搜索信息集合,是预先对海量搜索信息基于行业领域类型的维度,并结合自然语言处理方法分类得到的。
5. 一种实体的热度确定装置,包括:
 - 获取模块,用于获取与实体关联的多个搜索信息;
 - 第一确定模块,用于确定与所述多个搜索信息分别对应的多个热度贡献信息;以及
 - 第二确定模块,用于根据与所述多个搜索信息分别对应的多个热度贡献信息,确定与所述实体对应的热度值;
 - 所述第一确定模块,包括:
 - 第一确定子模块,用于确定与所述多个搜索信息分别对应的多个搜索结果;
 - 第二确定子模块,用于确定与所述多个搜索结果分别对应的多个搜索需求信息;以及

第三确定子模块,用于根据所述多个搜索需求信息,确定与所述多个搜索信息分别对应的多个热度贡献信息;

所述第三确定子模块,具体用于:

确定与所述多个搜索结果分别对应的多个页面浏览量;

根据与所述搜索结果对应的搜索需求信息,和与其对应的所述页面浏览量确定与所述搜索信息对应的热度贡献信息。

6. 根据权利要求5所述的装置,其中,所述第二确定子模块,用于:

确定与所述多个搜索结果分别对应的多个搜索需求比例值,并将所述搜索需求比例值作为所述搜索需求信息;

其中,所述搜索需求比例值,是采用预设函数对所述搜索结果在搜索结果页面中的显示位置数据进行量化处理后得到的函数值。

7. 根据权利要求5所述的装置,其中,所述第三确定子模块,具体用于:

如果所述搜索需求信息是搜索需求比例值,则确定所述搜索需求比例值与所述页面浏览量的乘积值,并将所述乘积值作为与所述搜索信息对应的热度贡献信息。

8. 根据权利要求5所述的装置,其中,所述获取模块,具体用于:

确定与所述实体对应的行业领域类型;

获取与所述行业领域类型对应的搜索信息集合,并将所述搜索信息集合之中的多个搜索信息作为所述关联的多个搜索信息;

其中,所述搜索信息集合,是预先对海量搜索信息基于行业领域类型的维度,并结合自然语言处理方法分类得到的。

9. 一种电子设备,包括:

至少一个处理器;以及

与所述至少一个处理器通信连接的存储器;其中,

所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行权利要求1-4中任一项所述的方法。

10. 一种存储有计算机指令的非瞬时计算机可读存储介质,其中,所述计算机指令用于使所述计算机执行根据权利要求1-4中任一项所述的方法。

11. 一种计算机程序产品,包括计算机程序,所述计算机程序在被处理器执行时实现根据权利要求1-4中任一项所述的方法。

实体的热度确定方法、装置、电子设备及存储介质

技术领域

[0001] 本申请涉及计算机技术领域,具体涉及深度学习、知识图谱、自然语言处理等人工智能技术领域,尤其涉及实体的热度确定方法、装置、电子设备及存储介质。

背景技术

[0002] 人工智能是研究使计算机来模拟人的某些思维过程和智能行为(如学习、推理、思考、规划等)的学科,既有硬件层面的技术也有软件层面的技术。人工智能硬件技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理等技术;人工智能软件技术主要包括计算机视觉技术、语音识别技术、自然语言处理技术,以及机器学习、深度学习、大数据处理技术、知识图谱技术等几大方向。

[0003] 在搜索场景中,热度是实体的一种重要的基础特征,有多种应用场景需要通过热度来判断实体的排序位置,主需求等。例如:当用户搜索实体“张三”时,需要通过其热度来判断大多数用户搜索的是演员张三,还是主持人张三。当用户搜索“2020年热门电影”时,需要将2020年上映的电影按热度顺序排列展示。

发明内容

[0004] 提供了一种实体的热度确定方法、装置、电子设备、存储介质及计算机程序产品。

[0005] 根据第一方面,提供了一种实体的热度确定方法,包括:获取与实体关联的多个搜索信息;确定与所述多个搜索信息分别对应的多个热度贡献信息;以及根据与所述多个搜索信息分别对应的多个热度贡献信息,确定与所述实体对应的热度值。

[0006] 根据第二方面,提供了一种实体的热度确定装置,包括:获取模块,用于获取与实体关联的多个搜索信息;第一确定模块,用于确定与所述多个搜索信息分别对应的多个热度贡献信息;以及第二确定模块,用于根据与所述多个搜索信息分别对应的多个热度贡献信息,确定与所述实体对应的热度值。

[0007] 根据第三方面,提供了一种电子设备,包括:至少一个处理器;以及与所述至少一个处理器通信连接的存储器;其中,所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行本申请实施例的实体的热度确定方法。

[0008] 根据第四方面,提出了一种存储有计算机指令的非瞬时计算机可读存储介质,所述计算机指令用于使所述计算机执行本申请实施例公开的实体的热度确定方法。

[0009] 根据第五方面,提出了一种计算机程序产品,包括计算机程序,当所述计算机程序由处理器执行时实现本申请实施例公开的实体的热度确定方法。

[0010] 应当理解,本部分所描述的内容并非旨在标识本申请的实施例的关键或重要特征,也不用于限制本申请的范围。本申请的其它特征将通过以下的说明书而变得容易理解。

附图说明

- [0011] 附图用于更好地理解本方案,不构成对本申请的限定。其中:
- [0012] 图1是根据本申请第一实施例的示意图;
- [0013] 图2是根据本申请第二实施例的示意图;
- [0014] 图3是本申请实施例中搜索需求信息的确定方式示意图;
- [0015] 图4为本申请实施例中的应用示意图;
- [0016] 图5是根据本申请第三实施例的示意图;
- [0017] 图6是根据本申请第四实施例的示意图;
- [0018] 图7是用来实现本申请实施例的实体的热度确定方法的电子设备的框图。

具体实施方式

[0019] 以下结合附图对本申请的示范性实施例做出说明,其中包括本申请实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本申请的范围和精神。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

[0020] 图1是根据本申请第一实施例的示意图。

[0021] 其中,需要说明的是,本实施例的实体的热度确定方法的执行主体为实体的热度确定装置,该装置可以由软件和/或硬件的方式实现,该装置可以配置在电子设备中,电子设备可以包括但不限于终端、服务器端等。

[0022] 本申请实施例涉及深度学习、知识图谱、自然语言处理等人工智能技术领域。

[0023] 其中,人工智能(Artificial Intelligence),英文缩写为AI。它是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。

[0024] 深度学习是学习样本数据的内在规律和表示层次,这些学习过程中获得的信息对诸如文字,图像和声音等数据的解释有很大的帮助。深度学习的最终目标是让机器能够像人一样具有分析学习能力,能够识别文字、图像和声音等数据。

[0025] 自然语言处理(Natural Language Processing,NLP),能实现人与计算机之间用自然语言进行有效通信的各种理论和方法。深度学习是学习样本数据的内在规律和表示层次,这些学习过程中获得的信息对诸如文字,图像和声音等数据的解释有很大的帮助。深度学习的最终目标是让机器能够像人一样具有分析学习能力,能够识别文字、图像和声音等数据。

[0026] 而知识图谱,是通过将应用数学、图形学、信息可视化技术、信息科学等学科的理论与方法与计量学引文分析、共现分析等方法结合,并利用可视化的图谱形象地展示学科的核心结构、发展历史、前沿领域以及整体知识架构达到多学科融合目的的现代理论。

[0027] 如图1所示,该实体的热度确定方法包括:

[0028] S101:获取与实体关联的多个搜索信息。

[0029] 其中,该实体可以是当前待确定其热度值的实体。

[0030] 其中,与该实体的特征具有一定的关联关系的搜索信息,可以被称为与该实体关联的搜索信息,该搜索信息通常可以被用于输入至搜索引擎中,以触发搜索。

[0031] 实体具体例如,林一版半生缘,刘一版半生缘,电影版半生缘等,对此不做限制。

[0032] 而与该实体相关联的搜索信息,可以具体是采用搜索信息在搜索引擎当中进行搜索,得到的搜索结果中能够链指至该实体。

[0033] 其中,搜索信息例如可以是搜索文本,或者是搜索词,对此不做限制。

[0034] 例如,如果实体是林一版半生缘,则关联的搜索信息可以是半生缘、半生缘在线观看,也即是说,当用户采用半生缘或者半生缘在线观看作为搜索信息时,得到的搜索结果中的具体内容能够链指出实体“林一版半生缘”,对此不做限制。

[0035] 本申请在具体执行的过程中,可以预先基于一些分类配置规则,针对各种特征的实体,确定相应关联的多个搜索信息,从而使得所关联的搜索信息,在搜索引擎当中进行搜索,得到的搜索结果中能够链指至该实体。

[0036] 而另外一些实施例中,还可以在根据搜索信息,确定出能够链指至的实体时,还可以采用以下方式实现:海量搜索信息(query)→搜索信息(query)对应的展现内容→各个展现内容的位置链指出实体→计算各个搜索信息query分配到各个位置的热度贡献信息→对链指出同一个实体的搜索信息query的热度贡献信息进行聚合,从而得到实体的热度值,对此不做限制。

[0037] 其中,所关联的搜索信息,在搜索引擎当中进行搜索,得到的搜索结果中能够链指至该实体,可以具体是对该搜索结果中的文本内容进行分词,得到多个分词,并且基于该分词能够链指至的实体。

[0038] 另外一些实施例中,在对该搜索结果中的文本内容进行分词,得到多个分词,并基于该分词能够链指至的实体之后,还可以对链指到的实体进行主要的实体的筛选。

[0039] 例如,采用搜索结果的展现内容的标题title辅助进行链指,可以对该标题title分词后,匹配出多个实体,还可以基于自然语言处理及统计方法,确定该展现内容页面主要的实体,具体例如:搜索半生缘,展现内容的标题title为《半生缘》开播,刘一,蒋一主演…,则会得到“半生缘-身份标识(Identity document, ID) Id1”,“刘一-Id2”,“蒋一-Id3”,但是,该展现内容页面主要描述是半生缘电视剧,则可以通过一些筛选方法,以确定该位置的实体为“半生缘-Id1”,对此不做限制。

[0040] 本申请实施例中的实体的特征,可以具体是行业领域类型,该行业领域类型能够用于描述当前针对该实体,所要分析的热度的领域特征,例如,行业领域类型例如电影类型,则表征该实体是电影领域的实体,行业领域类型还可以例如电视剧,动漫,综艺,小说,书籍,则相应的,行业领域类型,可以被用于描述该实体是电视剧领域的实体、或者是动漫领域的实体,以此类推。

[0041] 需要说明的是,上述的实体可以是一种结构化的数据,包含了实体相关的各种维度信息(比如,该实体可以具体例如,一个百科词条),对此不做限制。

[0042] 热度需求信息的划分不限于此,使得划分得到的热度需求信息,能够被用于区分出存在歧义的实体名称即可。

[0043] 可选地,一些实施例中,获取与实体关联的多个搜索信息,可以是确定与实体对应的行业领域类型,并获取与行业领域类型对应的搜索信息集合,并将搜索信息集合之中的多个搜索信息作为关联的多个搜索信息,其中,搜索信息集合,是预先对海量搜索信息基于行业领域类型的维度进行自然语言处理得到的,从而在有效提升关联的搜索信息的获取效率的同时,使得所确定的实体的热度值能够较好标识出歧义实体名称的热度情况,有效地

解决相关技术中存在的歧义实体名称的热度区分效果不佳的技术问题。

[0044] 举例而言,上述在形成多个搜索信息集合时,可以预先获取海量搜索信息(query),通过一个预训练的领域分类模型对海量搜索信息进行领域特征分类,从而形成多个搜索信息集合,并且还可以采用基于领域特征分类得到领域特征,针对各个搜索信息集合关联领域特征标签(该领域特征标签可以被用于匹配实体对应的行业领域类型)。

[0045] 上述的领域分类模型可以是一个多标签分类模型,输入为搜索信息(query)以及搜索信息(query)对应搜索结果所显示内容的文本语义向量,输出为多个类别的0,1的领域特征。

[0046] 具体例如,用户输入q=楚乔传,输出为1的类别有:小说,电视剧,筛选后,得到多个搜索信息集合,以及与各个搜索信息集合对应的领域特征标签,领域特征标签可以包括:电影,电视剧,动漫,综艺,小说,书籍等。

[0047] 在另外一些实施例中,还可以采用其他任意可能的方式获得与该实体相关联的搜索信息,比如,采用工程的方式、数学的方式,建模的方式等等,对此不做限制。

[0048] S102:确定与多个搜索信息分别对应的多个热度贡献信息。

[0049] 上述在获取与实体关联的多个搜索信息之后,可以确定与各个搜索信息对应的热度贡献信息。

[0050] 其中,该热度贡献信息,能够用于描述搜索信息的热度贡献情况,该热度贡献情况,具体是采用该搜索信息进行搜索得到的各个搜索结果的位置对应的实体,在该搜索信息下的贡献热度,该搜索信息的热度贡献情况,可以用对应的搜索结果的页面浏览量来衡量,对此不做限制。

[0051] 则本申请实施例中,可以确定与各个搜索信息分别对应的热度贡献情况,而后,基于各个搜索信息对应的热度贡献情况,运算得出该实体的热度值,且由于各个搜索信息是与该实体具有较好的关联关系的,从而实现仅利用了具有关联关系的多个搜索信息的热度贡献信息,即能够确定实体的热度值。

[0052] 在本申请的一些实施例中,确定与多个搜索信息分别对应的多个热度贡献信息,可以是确定与多个搜索信息分别对应的多个搜索结果,并确定与多个搜索结果分别对应的多个搜索需求信息,以及根据多个搜索需求信息,确定与多个搜索信息分别对应的多个热度贡献信息。

[0053] 也即是说,可以采用各个搜索信息在搜索引擎当中进行搜索,得到与各个搜索信息对应的搜索结果(搜索结果比如,至少一个的搜索结果页面、链接等),而后分析各个搜索结果对应的搜索需求信息(该搜索需求信息,可以用于描述搜索用户基于该搜索信息进行搜索后,针对各个搜索结果的需求偏好情况),而后,可以结合用户针对各个搜索结果的需求偏好情况,来拟合得到搜索信息的热度贡献信息。

[0054] 举例而言,可以预先对搜索信息进行搜索结果获取,获取相应搜索结果,而后,还可以反推出该搜索结果可以同样的被哪些搜索信息搜索到,从而将基于这些搜索信息的搜索量来辅助确定热度贡献信息,对此不做限制。

[0055] 具体例如,可以采用链指方式反推出该搜索结果可以同样的包含哪些其他的实体,比如可以分析该搜索结果的统一资源定位符,搜索结果标题,以及对应的文本内容,而后,采用实体链指模型,对各个统一资源定位符,搜索结果标题,以及对应的文本内容进行

链指,以获得该该搜索结果对应的描述内容所涉及的知识库中的实体标识。

[0056] 例如:用户搜索“西游记”,通过实体链指模型,可以获得搜索结果对应的描述内容所涉及的“西游记”实体标识(实体ID),而后,通过该实体ID可以查到该实体的结构化信息,该结构化信息中可以表征该搜索结果对应的描述内容对应的是“小说西游记”,还是“电视剧西游记”,以及是“哪一部电视剧西游记”,对此不做限制。

[0057] 上述通过确定与多个搜索信息分别对应的多个搜索结果,并确定与多个搜索结果分别对应的多个搜索需求信息,以及根据多个搜索需求信息,确定与多个搜索信息分别对应的多个热度贡献信息,能够结合用户针对各个搜索结果的需求偏好情况,来拟合得到搜索信息的热度贡献信息,从而进一步地提升热度分析的合理性,提升热度分析的准确性。

[0058] S103:根据与多个搜索信息分别对应的多个热度贡献信息,确定与实体对应的热度值。

[0059] 上述在确定与多个搜索信息分别对应的多个热度贡献信息,可以将各个搜索信息的热度贡献信息拟合至实体上,从而将拟合得到的结果作为实体的热度值。

[0060] 举例而言,可以对热度贡献信息进行量化处理,得到热度评价分值,而后采用一定的加权运算方式将各个搜索信息的热度评价分值加权运算,得到与实体对应的热度值。

[0061] 举例而言,还可以将各个搜索信息的热度贡献信息输入至预训练的热度分析模型当中,从而得到热度分析模型输出的与实体对应的热度值,对此不做限制。

[0062] 本实施例中,通过获取与实体关联的多个搜索信息,并确定与多个搜索信息分别对应的多个热度贡献信息,以及根据与多个搜索信息分别对应的多个热度贡献信息,确定与实体对应的热度值,从而利用少量的特征即能够准确地确定出实体的热度值,有效提升实体热度确定的合理性和效率,能够有效地提升实体的热度分析效果。

[0063] 图2是根据本申请第二实施例的示意图。

[0064] 如图2所示,该实体的热度确定方法包括:

[0065] S201:获取与实体关联的多个搜索信息。

[0066] S201的举例说明可以具体参见上述实施例,在此不再赘述。

[0067] S202:确定与多个搜索信息分别对应的多个搜索结果。

[0068] 也即是说,可以采用各个搜索信息在搜索引擎当中进行搜索,得到与各个搜索信息对应的搜索结果(搜索结果比如,至少一个的搜索结果页面、链接等)。

[0069] 搜索结果例如搜索页,举例而言,若在浏览器搜索“半生缘”,搜索页是<https://m.baidu.com/s?word=半生缘>,对此不做限制。

[0070] S203:确定与多个搜索结果分别对应的多个搜索需求比例值,并将搜索需求比例值作为搜索需求信息。

[0071] 其中,搜索需求比例值,是采用预设函数对搜索结果在搜索结果页面中的显示位置数据进行量化处理后得到的函数值。

[0072] 上述在确定与各个搜索信息分别对应的搜索结果之后,可以分析各个搜索结果对应的搜索需求信息(该搜索需求信息,可以用于描述搜索用户基于该搜索信息进行搜索后,针对各个搜索结果的需求偏好情况),而后,可以结合用户针对各个搜索结果的需求偏好情况,来拟合得到搜索信息的热度贡献信息。

[0073] 本实施例中的搜索需求信息,可以具体是搜索需求比例值,通常用户输入一个搜

索信息发起搜索,得到的多个搜索结果会按顺序排列在搜索结果页面当中,如图3所示,图3是本申请实施例中搜索需求信息的确定方式示意图,在搜索结果页面中排列了十个搜索结果(搜索卡片、网页2、百科、网页4, ..., 豆瓣、网页10),则基于用户的搜索习惯,通常用户会更倾向于点击、浏览排序在前的搜索结果,并针对各个搜索结果的搜索需求情况会逐渐的随着排序位置的下降而降低。

[0074] 由此,本实施例中,正是考虑到用户的搜索需求习惯,可以采用预设函数对搜索结果在搜索结果页面中的显示位置数据进行量化处理后得到的函数值,从而实现对用户针对各个搜索结果的搜索需求进行量化处理,并将量化处理得到的函数值作为与搜索结果对应的搜索需求信息,使得搜索需求信息的确定方式更具有合理性,并且通过函数拟合的方式得到搜索需求信息,便于后续搜索需求信息的使用以及模型算法输入。

[0075] 上述的预设函数可以具体是指数函数,即:

$$[0076] \quad f(x) = \lambda e^{-\lambda x}, x > 0。$$

[0077] 则本实施例中在采用预设函数对搜索结果在搜索结果页面中的显示位置数据进行量化处理时,可将指数函数在对应位置区间的面积(即,显示位置数据)作为用户需求在该展现位置对应内容的搜索需求比例值,并将该搜索需求比例值作为相应的搜索结果对应的搜索需求信息。

[0078] 举例而言,结合上述图3所示,搜索结果页面共计十个搜索结果,对应十个展现位置,将对应展示位置,依据上述指数函数的积分,确定其在相应区间段占据的面积(该面积被作为显示位置数据),以进行搜索需求比例的分配,上述指数函数的稀疏 λ 的值可自适应调节。

[0079] 具体例如,位置1的搜索需求比例值为对应区间 $[0, 1]$ 的指数函数积分。对应可得到在 $[1, 10]$ 不同位置区间中,搜索需求比例值在逐渐递减,且较大的搜索需求比例值对应展示在在排名前四的位置。

[0080] S204:确定与多个搜索结果分别对应的多个页面浏览量。

[0081] 其中,页面浏览量(Page Views, PV),通常是衡量一个网络新闻频道或网站甚至一条网络新闻的主要指标,本申请实施例中采用的页面浏览量,可以具体是日均页面浏览量,对此不做限制。

[0082] 本实施例中,可以采用相关技术中的页面浏览量统计分析方式来确定与多个搜索结果(例如,搜索页)分别对应的多个页面浏览量,对此不做限制。

[0083] S205:根据与搜索结果对应的搜索需求信息,和与其对应的页面浏览量确定与搜索信息对应的热度贡献信息。

[0084] 一些实施例中,可以将搜索需求比例值和页面浏览量输入至预训练的热度评分模型之中,从而热度评分模型输出的与搜索结果对应的热度评价值,对此不做限制。

[0085] 而本实施例中,当搜索需求信息是搜索需求比例值时,则可以确定搜索需求比例值与页面浏览量的乘积值,并将乘积值作为与搜索信息对应的热度贡献信息。

[0086] 也即是说,本实施例支持将根据搜索结果对应的搜索需求信息,将搜索结果对应的页面浏览量挂接至对应的搜索信息上,不需要使用较难获取的实体资源站点日志信息进行统计性的热度分析,本实施例中使用的特征较为简单,且具有较好的通用性,适用于对各种类型的实体进行热度值的快速地、准确地分析。

[0087] 另外,本实施例中,还可以确定搜索需求比例值与页面浏览量的乘积值,并将乘积值作为与搜索信息对应的热度贡献信息,能够有效地提升分析算法设计的合理性,保障运算处理的便捷性和适配性。

[0088] 如图4所示,图4为本申请实施例中的应用示意图,以搜索信息query = “半生缘”为例,最终得分计算如图所示:在左侧10个搜索结果中,第1,3,4,5位分别为影视剧相关的内容。其中,第1,4位为林一版半生缘对应内容,第2位为刘一版半生缘内容,第3位为电影版半生缘内容,可以基于预先训练好的实体链指模型,可根据展示内容提取到对应的实体标识,从而辅助将搜索信息“半生缘”对应的搜索页面浏览量pv分配在对应实体上,作为该query下计算所得的实体热度。

[0089] 如4图所示,右侧表格里灰色部分所示:query半生缘的日均页面浏览量pv为987.6,可以基于上述概率函数对应的搜索需求比例值,确定各个搜索结果的热度贡献信息(热度贡献信息具体可以是热度评分),即得到分别与第1,3,4,5位对应的热度评分为171.5,61.2,39.3,20.1,并分别对应在各个实体上。

[0090] S206:根据与多个搜索信息分别对应的多个热度贡献信息,确定与实体对应的热度值。

[0091] 上述在确定与多个搜索信息分别对应的多个热度贡献信息,可以将各个搜索信息的热度贡献信息拟合至实体上,从而将拟合得到的结果作为实体的热度值。

[0092] 举例而言,可以对热度贡献信息进行量化处理,得到热度评价分值,而后采用一定的加权运算方式将各个搜索信息的热度评价分值加权运算,得到与实体对应的热度值。

[0093] 举例而言,还可以将各个搜索信息的热度贡献信息输入至预训练的热度分析模型当中,从而得到热度分析模型输出的与实体对应的热度值,对此不做限制。

[0094] 举例而言,在上述图4中的表格内,林一版半生缘对应的query还有“半生缘在线观看”,从该query中获得的热度贡献信息是42.9,电影版半生缘从该query中获得的热度贡献信息是42.4,由此,可以将同一个实体对应不同query下的所有的热度贡献信息加和,可得到实体“林一版半生缘”的热度值为253.7,“电影半生缘”的热度值为81.7,“刘一版半生缘”的热度值为61.2。

[0095] 本实施例中,通过获取与实体关联的多个搜索信息,并确定与多个搜索信息分别对应的多个热度贡献信息,以及根据与多个搜索信息分别对应的多个热度贡献信息,确定与实体对应的热度值,从而利用少量的特征即能够准确地确定出实体的热度值,有效提升实体热度确定的合理性和效率,能够有效地提升实体的热度分析效果。实现对用户针对各个搜索结果的搜索需求进行量化处理,并将量化处理得到的函数值作为与搜索结果对应的搜索需求信息,使得搜索需求信息的确定方式更具有合理性,并且通过函数拟合的方式得到搜索需求信息,便于后续搜索需求信息的使用以及模型算法输入。支持将根据搜索结果对应的搜索需求信息,将搜索结果对应的页面浏览量挂接至对应的搜索信息上,不需要使用较难获取的实体资源站点日志信息进行统计性的热度分析,本实施例中使用的特征较为简单,且具有较好的通用性,适用于对各种类型的实体进行热度值的快速地、准确地分析。

[0096] 图5是根据本申请第三实施例的示意图。

[0097] 如图5所示,该实体的热度确定装置50,包括:

[0098] 获取模块501,用于获取与实体关联的多个搜索信息;

[0099] 第一确定模块502,用于确定与多个搜索信息分别对应的多个热度贡献信息;以及
[0100] 第二确定模块503,用于根据与多个搜索信息分别对应的多个热度贡献信息,确定与实体对应的热度值。

[0101] 在本申请的一些实施例中,如图6所示,图6是根据本申请第四实施例的示意图,该实体的热度确定装置60,包括:获取模块601、第一确定模块602、第二确定模块603,其中,其中,第一确定模块602,包括:

[0102] 第一确定子模块6021,用于确定与多个搜索信息分别对应的多个搜索结果;

[0103] 第二确定子模块6022,用于确定与多个搜索结果分别对应的多个搜索需求信息;以及

[0104] 第三确定子模块6023,用于根据多个搜索需求信息,确定与多个搜索信息分别对应的多个热度贡献信息。

[0105] 在本申请的一些实施例中,其中,第三确定子模块6023,具体用于:

[0106] 确定与多个搜索结果分别对应的多个页面浏览量;

[0107] 根据与搜索结果对应的搜索需求信息,和与其对应的页面浏览量确定与搜索信息对应的热度贡献信息。

[0108] 在本申请的一些实施例中,其中,第二确定子模块6022,用于:

[0109] 确定与多个搜索结果分别对应的多个搜索需求比例值,并将搜索需求比例值作为搜索需求信息;

[0110] 其中,搜索需求比例值,是采用预设函数对搜索结果在搜索结果页面中的显示位置数据进行量化处理后得到的函数值。

[0111] 在本申请的一些实施例中,其中,第三确定子模块6023,具体用于:

[0112] 如果搜索需求信息是搜索需求比例值,则确定搜索需求比例值与页面浏览量的乘积值,并将乘积值作为与搜索信息对应的热度贡献信息。

[0113] 在本申请的一些实施例中,其中,获取模块601,具体用于:

[0114] 确定与实体对应的行业领域类型;

[0115] 获取与行业领域类型对应的搜索信息集合,并将搜索信息集合之中的多个搜索信息作为关联的多个搜索信息;

[0116] 其中,搜索信息集合,是预先对海量搜索信息基于行业领域类型的维度,并结合自然语言处理方法分类得到的。

[0117] 可以理解的是,本实施例附图6中的实体的热度确定装置60与上述实施例中的实体的热度确定装置50,获取模块601与上述实施例中的获取模块501,第一确定模块602与上述实施例中的第一确定模块502,第二确定模块603与上述实施例中的第二确定模块503,可以具有相同的功能和结构。

[0118] 需要说明的是,前述对实体的热度确定方法的解释说明也适用于本实施例的实体的热度确定装置,此处不再赘述。

[0119] 本实施例中,通过获取与实体关联的多个搜索信息,并确定与多个搜索信息分别对应的多个热度贡献信息,以及根据与多个搜索信息分别对应的多个热度贡献信息,确定与实体对应的热度值,从而利用少量的特征即能够准确地确定出实体的热度值,有效提升实体热度确定的合理性和效率,能够有效地提升实体的热度分析效果。

[0120] 根据本申请的实施例,本申请还提供了一种电子设备、一种可读存储介质和一种计算机程序产品。

[0121] 图7是用来实现本申请实施例的实体的热度确定方法的电子设备的框图。电子设备旨在表示各种形式的数字计算机,诸如,膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置,诸如,个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅仅作为示例,并且不意在限制本文中描述的和/或者要求的本申请的实现。

[0122] 如图7所示,设备700包括计算单元701,其可以根据存储在只读存储器(ROM)702中的计算机程序或者从存储单元708加载到随机访问存储器(RAM)703中的计算机程序,来执行各种适当的动作和处理。在RAM 703中,还可存储设备700操作所需的各种程序和数据。计算单元701、ROM 702以及RAM 703通过总线704彼此相连。输入/输出(I/O)接口705也连接至总线704。

[0123] 设备700中的多个部件连接至I/O接口705,包括:输入单元706,例如键盘、鼠标等;输出单元707,例如各种类型的显示器、扬声器等;存储单元708,例如磁盘、光盘等;以及通信单元709,例如网卡、调制解调器、无线通信收发机等。通信单元709允许设备700通过诸如因特网的计算机网络和/或各种电信网络与其他设备交换信息/数据。

[0124] 计算单元701可以是各种具有处理和计算能力的通用和/或专用处理组件。计算单元701的一些示例包括但不限于中央处理单元(CPU)、图形处理单元(GPU)、各种专用的人工智能(AI)计算芯片、各种运行机器学习模型算法的计算单元、数字信号处理器(DSP)、以及任何适当的处理器、控制器、微控制器等。计算单元701执行上文所描述的各个方法和处理,例如,实体的热度确定方法。

[0125] 例如,在一些实施例中,实体的热度确定方法可被实现为计算机软件程序,其被有形地包含于机器可读介质,例如存储单元708。在一些实施例中,计算机程序的部分或者全部可以经由ROM 702和/或通信单元709而被载入和/或安装到设备700上。当计算机程序加载到RAM 703并由计算单元701执行时,可以执行上文描述的实体的热度确定方法的一个或多个步骤。备选地,在其他实施例中,计算单元701可以通过其他任何适当的方式(例如,借助于固件)而被配置为执行实体的热度确定方法。

[0126] 本文中以上描述的系统和技术各种实施方式可以在数字电子电路系统、集成电路系统、场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、芯片上系统的系统(SOC)、负载可编程逻辑设备(CPLD)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括:实施在一个或者多个计算机程序中,该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释,该可编程处理器可以是专用或者通用可编程处理器,可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令,并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

[0127] 用于实施本申请的实体的热度确定方法的程序代码可以采用一个或多个编程语言的任何组合来编写。这些程序代码可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理器或控制器,使得程序代码当由处理器或控制器执行时使流程图和/或

框图中所规定的功能/操作被实施。程序代码可以完全在机器上执行、部分地在机器上执行,作为独立软件包部分地在机器上执行且部分地在远程机器上执行或完全在远程机器或服务器上执行。

[0128] 在本申请的上下文中,机器可读介质可以是有形的介质,其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备,或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器 (RAM)、只读存储器 (ROM)、可擦除可编程只读存储器 (EPROM 或快闪存储器)、光纤、便捷式紧凑盘只读存储器 (CD-ROM)、光学储存设备、磁储存设备、或上述内容的任何合适组合。

[0129] 为了提供与用户的交互,可以在计算机上实施此处描述的系统和技术,该计算机具有:用于向用户显示信息的显示装置(例如,CRT(阴极射线管)或者LCD(液晶显示器)监视器);以及键盘和指向装置(例如,鼠标或者轨迹球),用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互;例如,提供给用户的反馈可以是任何形式的传感反馈(例如,视觉反馈、听觉反馈、或者触觉反馈);并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。

[0130] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如,作为数据服务器)、或者包括中间件部件的计算系统(例如,应用服务器)、或者包括前端部件的计算系统(例如,具有图形用户界面或者网络浏览器的用户计算机,用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如,通信网络)来将系统的部件相互连接。通信网络的示例包括:局域网(LAN)、广域网(WAN)、互联网及区块链网络。

[0131] 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务端的关系。服务器可以是云服务器,又称为云计算服务器或云主机,是云计算服务体系中的一项主机产品,以解决了传统物理主机与VPS服务("Virtual Private Server",或简称"VPS")中,存在的管理难度大,业务扩展性弱的缺陷。服务器也可以为分布式系统的服务器,或者是结合了区块链的服务器。

[0132] 应该理解,可以使用上面所示的各种形式的流程,重新排序、增加或删除步骤。例如,本发申请中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行,只要能够实现本申请公开的技术方案所期望的结果,本文在此不进行限制。

[0133] 上述具体实施方式,并不构成对本申请保护范围的限制。本领域技术人员应该明白的是,根据设计要求和因素,可以进行各种修改、组合、子组合和替代。任何在本申请的精神和原则之内所作的修改、等同替换和改进等,均应包含在本申请保护范围之内。

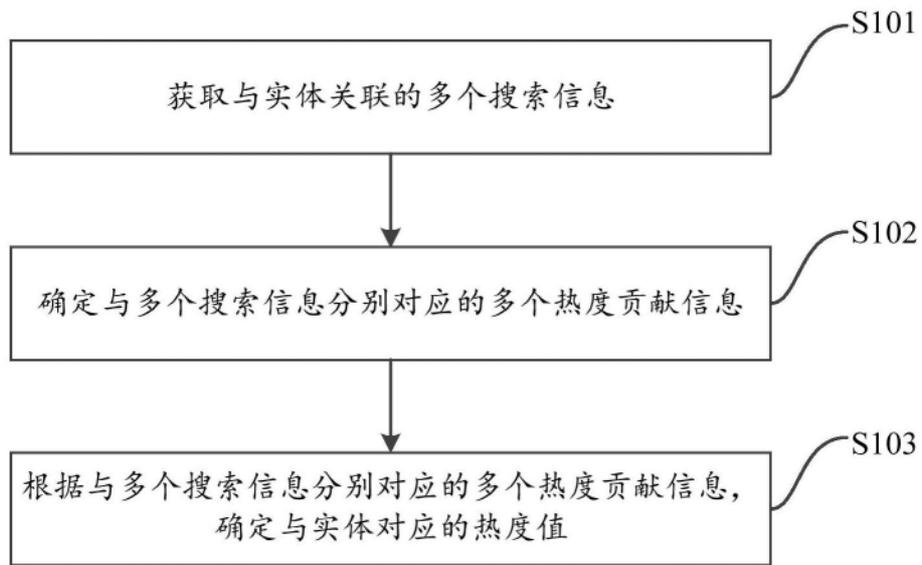


图1

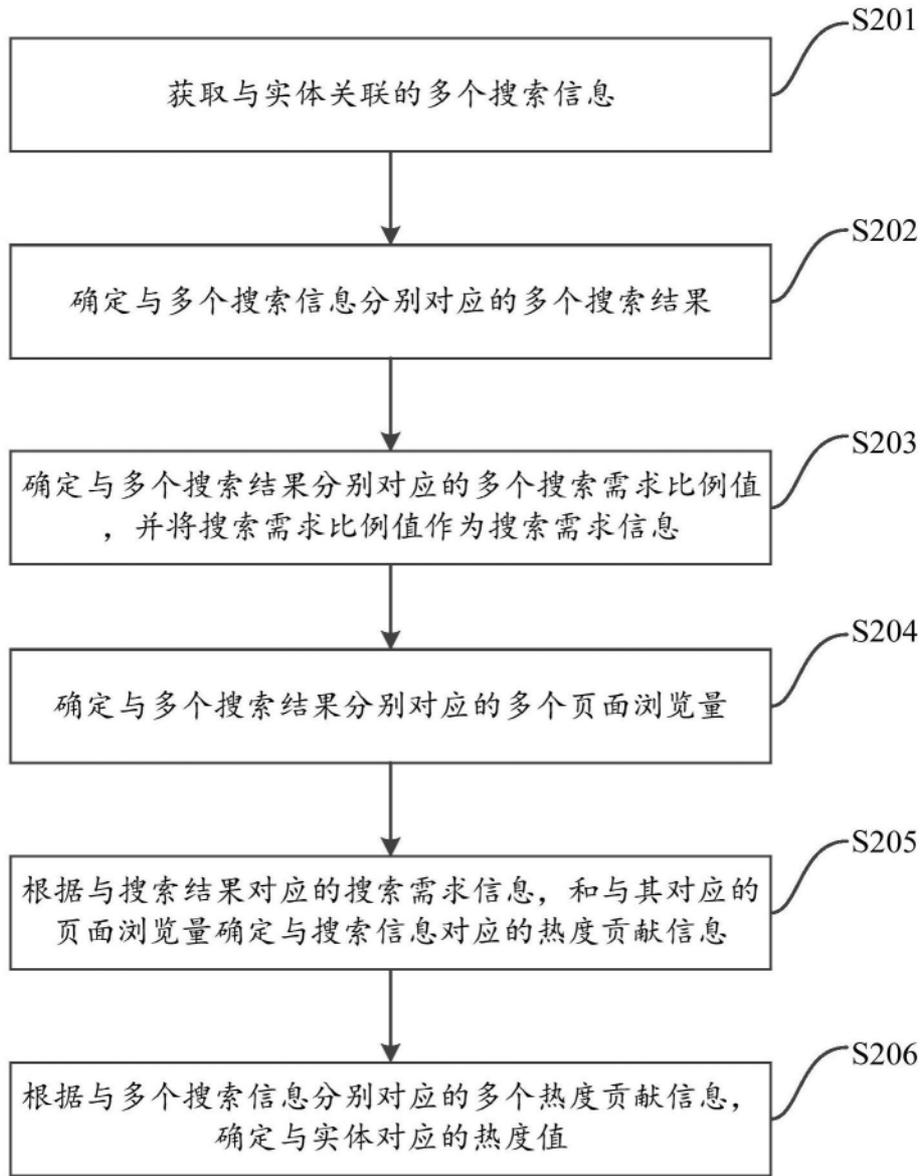


图2

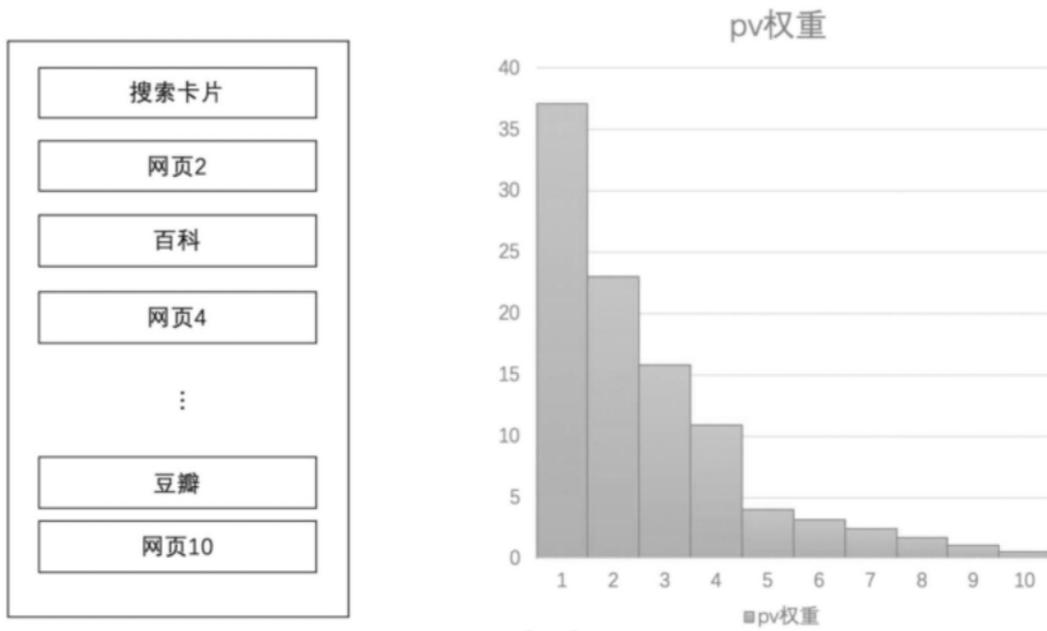


图3

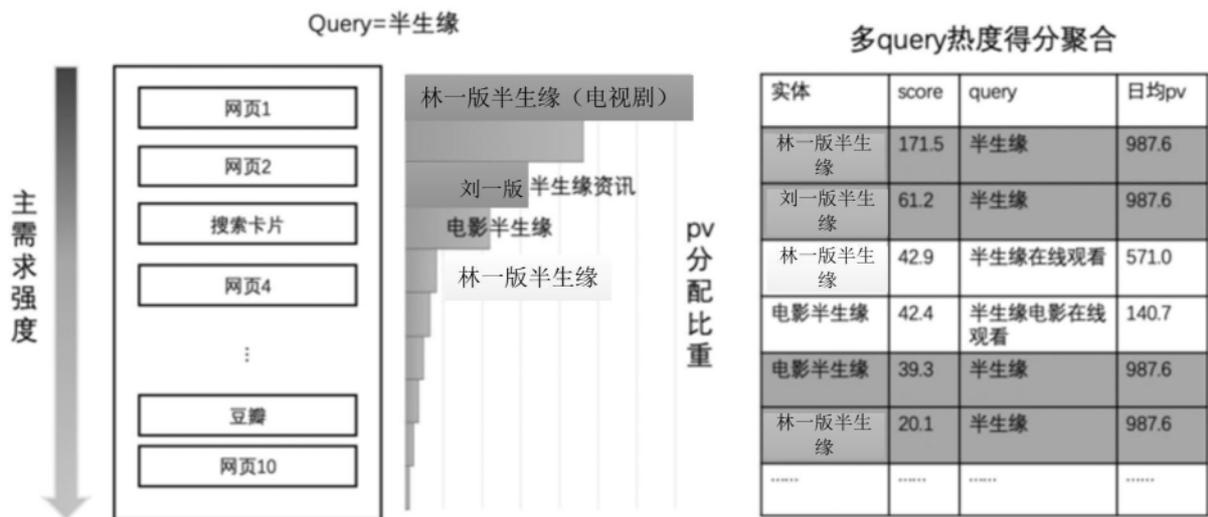


图4

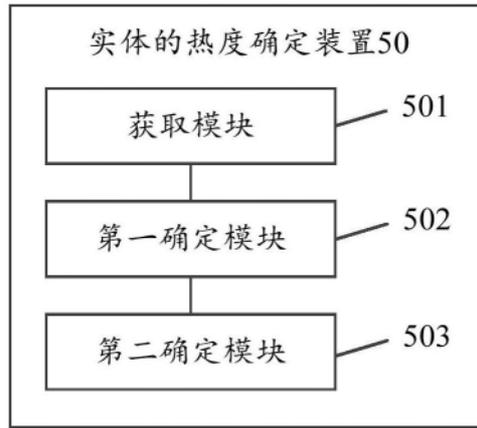


图5

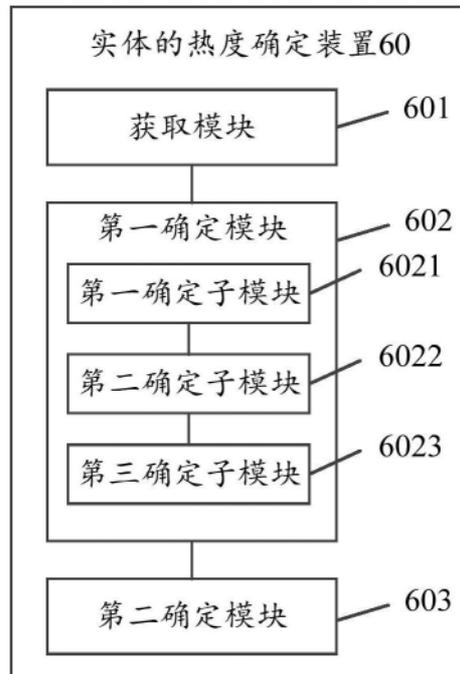


图6

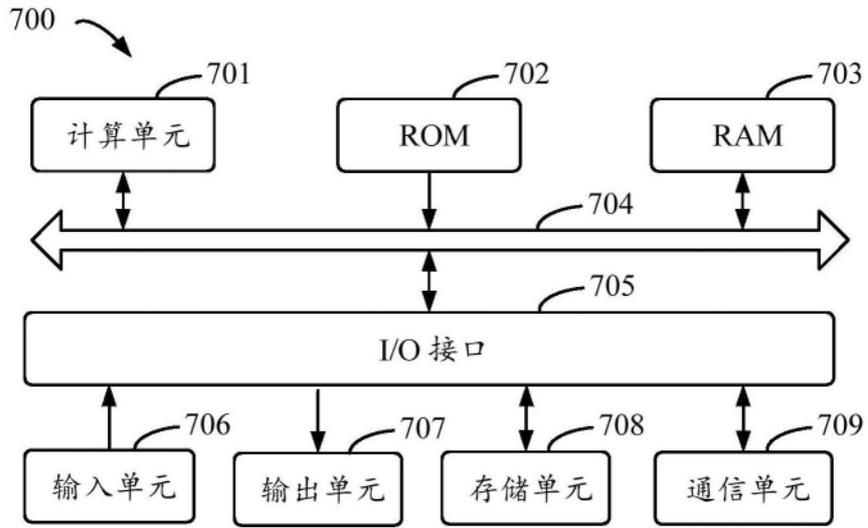


图7