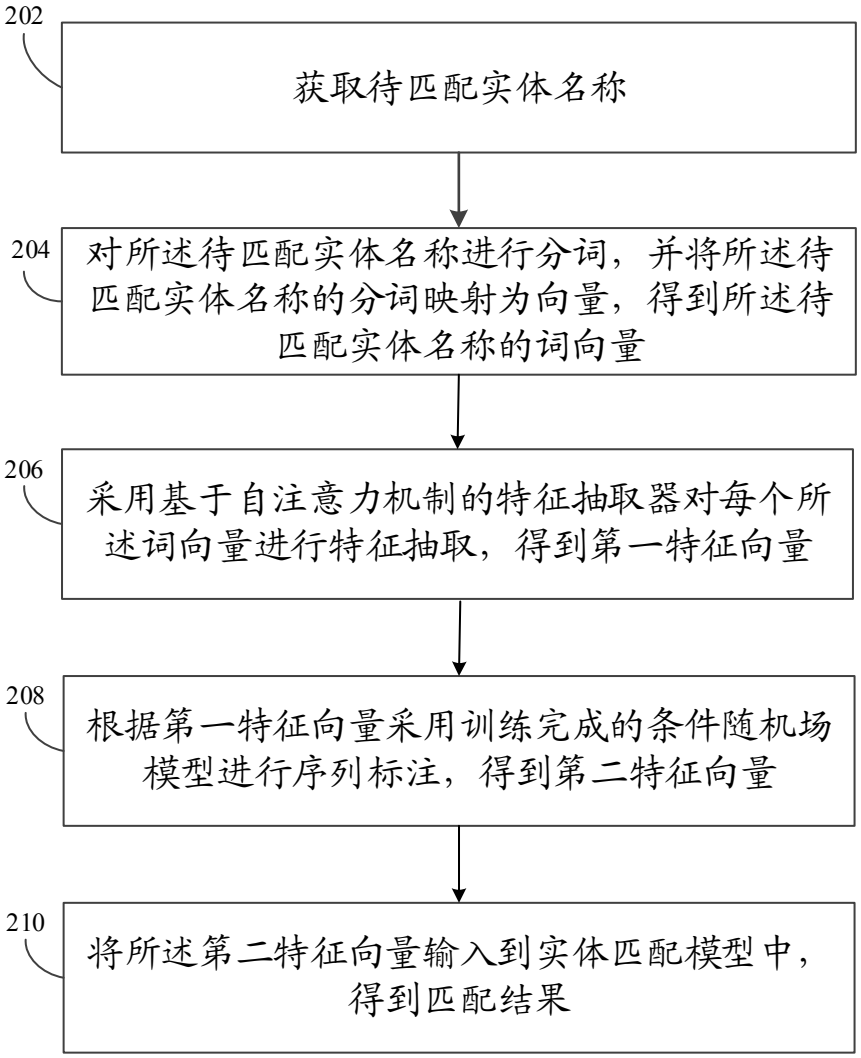
说　明　书　摘　要

本说明书实施例提供一种实体名称匹配方法、装置及设备。方案包括：获取待匹配实体名称；对待匹配实体名称进行分词，并将分词映射为向量，得到待匹配实体名称的词向量；采用基于自注意力机制的特征抽取器对每个词向量进行特征抽取，得到第一特征向量；再采用训练完成的条件随机场模型对第一特征向量进行序列标注，得到携带有域标签的第二特征向量，将第二特征向量输入到实体匹配模型中，得到匹配结果。

摘　要　附　图


202
获取待匹配实体名称
204
对所述待匹配实体名称进行分词，并将所述待匹配实体名称的分词映射为向量，得到所述待匹配实体名称的词向量
206
采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到第一特征向量
208
根据第一特征向量采用训练完成的条件随机场模型进行序列标注，得到第二特征向量
210
将所述第二特征向量输入到实体匹配模型中，得到匹配结果

权　利　要　求　书

1、一种实体名称匹配方法，包括:

获取待匹配实体名称;

对所述待匹配实体名称进行分词，并将所述待匹配实体名称的分词映射为向量，得到所述待匹配实体名称的词向量;

采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到第一特征向量;

采用训练完成的条件随机场模型对所述第一特征向量进行序列标注，得到第二特征向量，所述第二特征向量为携带有域标签的特征向量;

将所述第二特征向量输入到实体匹配模型中，得到匹配结果。

2、如权利要求1所述的方法，所述获取待匹配实体名称，具体包括:

获取待匹配交易数据;

从所述待匹配交易数据中提取出交易双方的账户实体名称，所述账户实体名称包括公司实体名称。

3、如权利要求1所述的方法，所述采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到第一特征向量，具体包括:

根据所述词向量的上下文信息采用自注意力机制计算每个词向量的权重值;

根据所述权重值对每个所述词向量进行注意力权重赋值，得到第一特征向量。

4、如权利要求3所述的方法，所述根据所述词向量的上下文信息采用自注意力机制计算每个词向量的权重值，具体包括:

对于任意一个所述词向量，计算所述实体名称中的其他词向量对该词向量的注意力权重;

对所述任意一个所述词向量的注意力权重进行归一化;

将进行归一化后的权重进行加权求和，得到每个词向量的权重值。

5、如权利要求1所述的方法，所述采用训练完成的条件随机场模型对所

述第一特征向量进行序列标注，得到第二特征向量，具体包括:

确定每个所述第一特征向量对应的域标签概率;

根据所述域标签概率确定所述第一特征向量的标签转移关系;

根据所述标签转移关系对所述第一特征向量进行序列组合排序，得到组合排序后的特征向量;

对所述组合排序后的特征向量标注域标签，得到第二特征向量。

6、如权利要求 5 所述的方法，所述域标签包括：名称标签、地址标签、领域标签、后缀标签和/或其他标签。

7、如权利要求 1 所述的方法，所述采用训练完成的条件随机场模型对所述第一特征向量进行序列标注之前，还包括:

获取域标签已知的实体名称样本;

提取所述实体名称样本对应的第三特征向量;

将所述第三特征向量输入待训练的条件随机场模型进行训练，得到所述待训练的条件随机场模型输出的对所述第三特征向量的域标签标注结果;

将全部所述实体名称样本中的域标签标注结果与全部所述实体名称样本的已知域标签进行比对，得到比对结果;

当所述比对结果表示所述全部实体名称样本中的域标签标注结果与所述实体名称样本的已知域标签相比，准确率达到预设阈值时，得到训练完成的条件随机场模型。

8、如权利要求 1 所述的方法，所述实体匹配模型中包括实体名称名单，所述实体名称名单包括公司名称。

9、如权利要求 8 所述的方法，所述将所述第二特征向量输入到实体匹配模型中，得到匹配结果，具体包括:

按照域标签分类将所述第二特征向量与所述实体名单中相同域的特征向量进行对齐匹配，得到每个域标签对应的相似度匹配分数;

将所述每个域标签对应的相似度匹配分数进行加权得到所述第二特征向

量的匹配分数;

当所述匹配分数大于预设分值时，将匹配分数最高的实体名称作为所述待匹配实体名称的匹配结果。

10、如权利要求9所述的方法，所述将所述第二特征向量输入到实体匹配模型中，得到匹配结果之后，还包括:

当所述匹配结果表示所述匹配分数大于预设分值时，停止对于所述待匹配实体名称对应的交易数据的处理过程。

11、一种实体名称匹配装置，包括:

待匹配实体名称获取模块，用于获取待匹配实体名称;

词向量确定模块，用于对所述待匹配实体名称进行分词，并将所述待匹配实体名称的分词映射为向量，得到所述待匹配实体名称的词向量;

特征抽取模块，用于采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到第一特征向量;

序列标注模块，用于采用训练完成的条件随机场模型对所述第一特征向量进行序列标注，得到第二特征向量，所述第二特征向量为携带有域标签的特征向量;

匹配模块，用于将所述第二特征向量输入到实体匹配模型中，得到匹配结果。

12、一种实体名称匹配设备，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:

获取待匹配实体名称;

对所述待匹配实体名称进行分词，并将所述待匹配实体名称的分词映射为向量，得到所述待匹配实体名称的词向量;

采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到第一特征向量;

采用训练完成的条件随机场模型对所述第一特征向量进行序列标注，得到第二特征向量，所述第二特征向量为携带有域标签的特征向量;

将所述第二特征向量输入到实体匹配模型中，得到匹配结果。

13、一种计算机可读介质，其上存储有计算机可读指令，所述计算机可读指令可被处理器执行以实现权利要求 1 至 10 中任一项所述的实体名称匹配方法。

说　明　书

一种实体名称匹配方法、装置及设备

技术领域

本说明书一个或多个实施例涉及计算机技术领域，尤其涉及一种实体名称匹配方法、装置及设备。

背景技术

目前，命名实体识别(Named Entities Recognition, NER)是自然语言处理(Natural Language Processing, NLP)的一个基础任务。其目的是识别语料中人名、地名、组织机构名等命名实体。由于这些命名实体数量不断增加，通常不可能在词典中穷尽列出，且其构成方法具有各自的一些规律性，因而，通常把对这些词的识别从词汇形态处理(如汉语切分)任务中独立处理，称为命名实体识别。命名实体识别技术是信息抽取、信息检索、机器翻译等多种自然语言处理技术必不可少的组成部分。

现有技术中，在进行命名实体识别时，一般采用全词对齐匹配的方法，比如：直接基于字符串匹配算法。但是现有技术中的方法，并没有考虑到实体文本中各个字词之间的语义关联，比如公司名作为文本的语义关联。也无法区分企业/机构名中的各部分在实体文本匹配中的重要程度，在面向大量企业业务时，会因为着重匹配非关键部分（如后缀，地区等）、简称部分等而导致匹配到错误的实体对象，难以保证匹配准确性，对于含有非常见词、非登录词的文本泛化能力也较差，导致系统打扰率的升高。

因此，需要提供一种更可靠的实体名称匹配方案。

发明内容

有鉴于此，本说明书一个或多个实施例提供了一种实体名称匹配方法、装置及设备，用于提高实体名称匹配的准确率。

为解决上述技术问题，本说明书实施例是这样实现的:

本说明书实施例提供的一种实体名称匹配方法，包括:

获取待匹配实体名称;

对所述待匹配实体名称进行分词，并将所述待匹配实体名称的分词映射为向量，得到所述待匹配实体名称的词向量;

采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到第一特征向量;

采用训练完成的条件随机场模型对所述第一特征向量进行序列标注，得到第二特征向量，所述第二特征向量为携带有域标签的特征向量;

将所述第二特征向量输入到实体匹配模型中，得到匹配结果。

本说明书实施例提供的一种实体名称匹配装置，包括:

待匹配实体名称获取模块，用于获取待匹配实体名称;

词向量确定模块，用于对所述待匹配实体名称进行分词，并将所述待匹配实体名称的分词映射为向量，得到所述待匹配实体名称的词向量;

特征抽取模块，用于采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到第一特征向量;

序列标注模块，用于采用训练完成的条件随机场模型对所述第一特征向量进行序列标注，得到第二特征向量，所述第二特征向量为携带有域标签的特征向量;

匹配模块，用于将所述第二特征向量输入到实体匹配模型中，得到匹配结果。

本说明书实施例提供的一种实体名称匹配设备，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:

获取待匹配实体名称;

对所述待匹配实体名称进行分词，并将所述待匹配实体名称的分词映射为向量，得到所述待匹配实体名称的词向量;

采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到第一特征向量;

采用训练完成的条件随机场模型对所述第一特征向量进行序列标注，得到第二特征向量，所述第二特征向量为携带有域标签的特征向量;

将所述第二特征向量输入到实体匹配模型中，得到匹配结果。

本说明书实施例提供的一种计算机可读介质，其上存储有计算机可读指令，所述计算机可读指令可被处理器执行以实现一种实体名称匹配方法。

本说明书一个实施例实现了能够达到以下有益效果：通过获取待匹配实体名称；对所述待匹配实体名称进行分词，并将所述待匹配实体名称的分词映射为向量，得到所述待匹配实体名称的词向量，采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，采用训练完成的条件随机场模型对所述第一特征向量进行序列标注，得到携带有域标签第二特征向量，能够捕捉字词的上下文依赖关系以及标签序列的概率转移关系，降低实体名称的匹配失误率，提高实体名称的匹配效率。

附图说明

此处所说明的附图用来提供对本说明书一个或多个实施例的进一步理解，构成本说明书一个或多个实施例的一部分，本说明书的示意性实施例及其说明用于解释本说明书一个或多个实施例，并不构成对本说明书一个或多个实施例的不当限定。在附图中：

图 1 为本说明书实施例中一种实体名称匹配方法的模型结构示意图;

图 2 为本说明书实施例提供的一种实体名称匹配方法的流程示意图;

图 3 为本说明书实施例提供的对应于图 2 的一种实体名称匹配装置的结构

示意图;

图 4 为本说明书实施例提供的对应于图 2 的一种实体名称匹配设备的结构示意图。

具体实施方式

为使本说明书一个或多个实施例的目的、技术方案和优点更加清楚，下面将结合本说明书具体实施例及相应的附图对本说明书一个或多个实施例的技术方案进行清楚、完整地描述。显然，所描述的实施例仅是本说明书的一部分实施例，而不是全部的实施例。基于本说明中的实施例，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他实施例，都属于本说明书一个或多个实施例保护的范围。

以下结合附图，详细说明本说明书各实施例提供的技术方案。

随着线上交易的发展，对线上交易的安全性要求越来越高，比如：在反洗钱领域中，通常会通过管控网上交易的用户账户的安全性来保证交易的安全性，此时，可以通过判断交易的账户双方的安全性来管控交易行为，具体地，在金融领域，通常可以根据一些历史交易数据得到一些风险账户（比如：失信账户或者其他存在风险的账户），在风险管控系统中可以将存在风险的账户相关信息进行存储，作为判断交易账户安全性的参考。比如：交易平台接收到一条交易请求，此时，交易平台可以从交易请求对应的交易数据中提取出交易双方的账户实体名称，将交易双方的账户实体名称与数据库中保存的存在风险的账户实体名称进行匹配，如果匹配成功，可以认为该交易请求对应的至少一个账户实体属于存在风险的账户实体，此时，可以停止对于该交易数据的处理过程。并且再此基础上还可以向交易双发发送交易失败的提示信息。

以交易制裁名单匹配为例，存在一份公司制裁名单，该公司制裁名单中包括至少一个公司实体名称。此时，在进行交易制裁名单匹配时，交易方可以包含至少一个公司实体账户。因此，首先可以从交易请求对应的交易数据中提取

出公司实体名称，将提取得到的公司实体名称与交易制裁名单中的公司实体名称进行匹配，如果匹配成功，可以认为该交易请求存在风险，可以停止对于该交易数据的处理过程，在此基础上也可以向交易双发发送交易失败的提示信息。

在进行实体名称匹配时，具体可以采用下面的实施例进行实现:

图 1 为本说明书实施例中一种实体名称匹配方法的模型结构示意图。如图 1 所示，在进行实体名称匹配时，可以采用图 1 中的模型结构来实现，模型中包括嵌入层 101、特征提取层 103、序列标注层 105 和匹配层 107。其中，嵌入层 101 负责嵌入待匹配实体名称对应的词向量 102，特征提取层 103 可以是基于自注意力机制的 transformer（特征抽取器），负责对嵌入层 101 嵌入的待匹配实体名称中的词向量 102 进行特征抽取。特征抽取层 103 在进行特征抽取时，会考虑到其他向量对自身向量的注意力权重，使抽取到的第一特征向量 104 是考虑了上下文相关性的特征向量。序列标注层 105 可以采用条件随机场模型（Conditional Random Field，简称 CRF），负责对特征抽取层 103 抽取的第一特征向量 104 进行序列标注，可以理解为对第一特征向量 104 进行排序组合，并打上域标签，得到第二特征向量 106，域标签可以包括名称标签、领域标签、地址标签、后缀标签等。其中，机构可以用 ORG 表示，地点可以用 LOC 表示，名称可以用 NAME 表示， B 表示开始的字节，I 表示中间的字节，E 表示最后的字节，S 表示该实体是单字节，B-Loc 表示开始字节为地址标签。匹配层 107 可以采用机器学习模型，这一机器学习模型可以是采用已知实体名称名单训练得到的模型，负责对第二特征向量 106 进行匹配，得到匹配分数，当匹配分数大于预设分数阈值时，可以认为待匹配实体名称存在于实体名称名单中，更进一步地，可以将匹配分数最高的实体名称作为待匹配实体名称的匹配对象。

接下来，将针对说明书实施例提供的一种实体名称匹配方法结合附图进行具体说明:

图 2 为本说明书实施例提供的一种实体名称匹配方法的流程示意图。从程

序角度而言，流程的执行主体可以为搭载于应用服务器的程序或应用客户端。

如图 2 所示，该流程可以包括以下步骤:

步骤 202: 获取待匹配实体名称。

这里的待匹配实体名称可以指的是命名实体 (named entity)，命名实体一般可以分为实体类、时间类和数字类，可以指的是人名、机构名、地名、时间、日期、货币和百分比以及其他所有以名称为标识的实体。

步骤 204: 对所述待匹配实体名称进行分词，并将所述待匹配实体名称的分词映射为向量，得到所述待匹配实体名称的词向量。

对待匹配实体名称进行分词时，可以将待匹配实体名称中的每一个字都分离出来。以公司名称实体为例，例如，“北京电力 A1A2A3A4 有限公司”，对这个公司名称进行分词后的结果可以是: “/北/京/电/力/A1/ A2/ A3/ A4/有/限/公/司/，将待匹配实体名称的分词映射为向量，可以是将待匹配实体名称中分离出来的每一个字通过查找分词向量映射表得到对应的词向量。这里的分词向量映射表可以是预先存储或加载的分词向量映射表。

在确定词向量的具体过程中，为了增加模型的泛化能力，可以使用一些模型利用大量语料训练出的词向量做初始化，比如: 利用大量无监督语料得到的语料库训练得到的词向量，用来作为模型词向量的初始化，以公司名称为例，在具体场景下，基于公司实体名称继续训练调整词向量，对训练集之外的数据有很好的泛化性，能够提升词向量的覆盖度。另外，也可以通过多语种语料预训练词向量，并在训练数据中加入多语种公司名数据，模型可以具有多语种匹配能力。

步骤 206: 采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到第一特征向量。

自注意力机制会模仿生物观察行为的内部过程，提取稀疏数据的重要特征，能减少对外部信息的依赖，擅长捕捉数据或特征的内部相关性。

基于自注意力机制的特征抽取器（transform）可以对实体名称中词向量的

上下文特征进行提取，transformer 是由编码组件、解码组件和它们之间的连接组成。编码组件部分由一堆编码器（encoder）构成。解码组件部分也是由与编码器相同数量的解码器（decoder）组成的。

从编码器输入的实体名称首先会经过自注意力（self-attention）层，以帮助编码器在对每个词向量编码时关注输入实体名称中的其他词向量对该词向量本身的注意力权重影响。解码器中也有编码器的自注意力（self-attention）层和前馈（feed-forward）层。除此之外，这两个层之间可以设置注意力层，用来关注输入句子的相关部分。

第一特征向量可以包括各个词向量对应的多个特征向量。基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到的第一特征向量可以是考虑了实体名称中各词向量的上下文相关性的特征向量，比如：存在待匹配的实体名称 X，包括词向量 A1、A2、A3、A4、A5，采用基于自注意力机制的特征抽取器对每个词向量进行特征抽取，分别得到第一特征向量 a1、a2、a3、a4、a5，其中，特征向量 a1 是考虑了 A2、A3、A4、A5 对 A1 的注意力权重之后得到的特征向量，特征向量 a2 是考虑了 A1、A3、A4、A5 对 A1 的注意力权重之后得到的特征向量。其中，对每个词向量进行特征提取时，可以计算其他所有的字对自己的影响，分别得到一个向量，将向量拼接起来形成自身的向量，从而得到第一特征向量。

步骤 208：采用训练完成的条件随机场模型对所述第一特征向量进行序列标注，得到第二特征向量，所述第二特征向量为携带有域标签的特征向量。

条件随机场(conditional random fields，简称 CRF，或 CRFs)，是一种判别式概率模型，是随机场的一种，常用于标注或分析序列资料。条件随机场模型是一种可以考虑多因素，对物体进行多标签分类的模型。

以公司名称为例，公司实体名称通常可以由五个部分组成，名称、领域、地址、后缀以及无意义部分，不同部分在匹配时有不同的权重。所述域标签可以包括：名称标签、地址标签、领域标签和/或后缀标签。其中，领域标签可以

是公司的经营类型、公司所属的领域等，比如：知识产权领域、服装领域、医疗领域等等。

在进行序列标注时，可以对分词之后的词向量进行单独标注标签，也可以结合实体名称中各个词向量之间的相关性进行重新组合排序之后，再进行标签标注，标注完之后的特征向量是携带有域标签的向量，即第二特征向量可以是对第一特征向量进行分域后的特征向量。比如：将公司名称“北京 X1X2Y1Y2 有限公司”，进行分词后得到北京/X1X2 /Y1Y2/有限公司/，将分词映射为词向量:C1C2/C3C4/C5 C6/C7C8C9C10/，采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，可以得到第一特征向量：C1C2→d1、C3C4→d2、C5 C6→d3、C7C8C9C10→d4，第一特征向量中各个向量具有对应的注意力权重，可以获取上下文特征。采用训练完成的条件随机场模型对第一特征向量中的各个向量进行序列标注，比如：将 d1 打上地域标签，将 d2 打上名称标签，将 d3 打上领域标签，将 d4 打上后缀标签。

步骤 210：将所述第二特征向量输入到实体匹配模型中，得到匹配结果。

实体匹配模型可以是机器学习模型，具体可以是提前训练完成的模型，将公司名称对应的特征向量进行分域之后，输入实体匹配模型中可以按照域里面的词去进行对齐和匹配，针对每个域都可以得到一个匹配得分，最后进行加权得到待匹配实体名称的匹配得分。

图 2 中的方法，通过获取待匹配实体名称，对所述待匹配实体名称进行分词，并将所述待匹配实体名称的分词映射为向量，得到所述待匹配实体名称的词向量，采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，采用训练完成的条件随机场模型对所述第一特征向量进行序列标注，得到携带有域标签第二特征向量，能够捕捉字词的上下文依赖关系以及字词标签序列的概率转移关系，能降低实体名称的匹配失误率，提升特征的抽取能力，提高实体名称的匹配效率。

基于图 2 的方法，本说明书实施例还提供了该方法的一些具体实施方案，

下面进行说明。

图 2 中的方法可以具体用在交易场景中，通过判断交易的账户双方是否属于危险账户，来判断是否允许交易；还可以用于制裁扫描场景（比如：反洗钱制裁扫描）或者需要搜索的场景中。其中，制裁扫描，可以理解为根据已有名单扫描匹配公司名称，匹配到的公司即为被制裁的公司。匹配成功的可以暂停交易、继续审核或者直接冻结交易。

实体名称的获取具体可以从待匹配交易数据提取，待匹配交易数据中交易双方的账户实体名称可以作为待匹配的实体名称。

实体名称匹配中，一般会采用名称字符串匹配、基于词典/规则进行匹配、根据提取的关键词进行匹配或者利用机器学习或者深度学习的方法进行匹配等方案。但是这些方法忽略了实体名称中各个字的重要程度（例如：公司名称中的名称部分相比于其他部分在匹配时重要性更高）、覆盖度不够，泛化能力差。实体名称的重要性与出现频次不成正比，比如：以匹配公司名称为例，为了保证公司名称的独特性，经常会将不常见的词作为公司名称，导致匹配时可能出现的误匹配情况。通用的分类模型没有考虑到上下文相关性、一字/词多义等现象。为了克服上述缺陷，可以采用以下技术方案：

在实际应用中，所述采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到第一特征向量，具体可以包括：

根据所述词向量的上下文信息采用自注意力机制计算每个词向量的权重值；

根据所述权重值对每个所述词向量进行注意力权重赋值，得到第一特征向量。

其中，所述根据所述词向量的上下文信息采用自注意力机制计算每个词向量的权重值，具体可以包括：

对于任意一个所述词向量，计算所述实体名称中的其他词向量对该词向量的注意力权重；

对所述任意一个所述词向量的注意力权重进行归一化;

将进行归一化后的权重进行加权求和，得到每个词向量的权重值。

需要说明的是，在采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取时，需要计算每个词向量的注意力权重，在计算特征向量的注意力权重时，可以引入不同的函数和计算机制来计算该向量与其他任意向量两者的相似性或者相关性。其中，最常见的方法包括：求两者的向量点积、求两者的向量 Cosine 相似性。得到的分值根据具体产生的方法不同其数值取值范围也不一样，接着可以采用 SoftMax 的计算方式对得到的分值进行数值转换，一方面可以进行归一化，将原始计算分值整理成所有元素权重之和为 1 的概率分布；另一方面也可以通过 SoftMax 的内在机制更加突出重要元素的权重，然后进行加权求和即可得到注意力权重。

上述方法中，在进行特征提取时，除了采用基于自注意力机制的特征提取器 Transformer，还可以考虑进行其他模型的组合，如在编解码层中加入 CNN，在解码器后再加入一层 Bi-LSTM 等，对特征提取能力的小幅度的提升。

通过上述方法，采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，可以捕获同一个实体名称中词向量之间的上下文特征，更容易捕获句子中长距离的相互依赖的特征，相对于循环神经网络（Recurrent Neural Network，简称 RNN）或者长短期记忆网络（Long Short Term Memory，简称 LSTM），不需要依次序序列计算对于远距离的相互依赖的特征，不需要经过若干时间步步骤的信息累积就能将两者联系起来，不管两个词向量中间有多长距离，最大的路径长度也都只是 1。因此，采用 CRF 进行序列标注能够无视词之间的距离直接计算依赖关系，能够学习一个句子的内部结构，能够有效捕捉词向量之间的上下文特征。

在具体的应用场景中，以公司名称为例，公司名称之间存在一定上下文依赖。例如“巴”出现在公司名称中时，属于名称领域，在“古巴”、“巴东”等地址名称领域中属于地址，这些中文分词在公司名这类外词或非登录词（out of

vocabulary，简称 OOV）较多的场景下分词准确率不足。另外，公司名称中的各个部分之间存在一定的依赖与转移关系，比如，在中文语境下，公司名称中，名称部分一般会在后面接上其他名称或者行业。

考虑到上述实体名称中可能存在的问题，可以采用以下技术方案对实体名称进行序列标注：

所述采用训练完成的条件随机场模型对所述第一特征向量进行序列标注，得到第二特征向量，具体可以包括：

确定每个所述第一特征向量对应的域标签概率；

根据所述域标签概率确定所述第一特征向量的标签转移关系；

根据所述标签转移关系对所述第一特征向量进行序列组合排序，得到组合排序后的特征向量；

对所述组合排序后的特征向量标注域标签，得到第二特征向量。

在标注时，可以对分词后的每一个词向量分别进行域标签标注，也可以结合上下文特征，对分词后的向量进行重新组合并排序，进行域标签标注。为了关注到序列标注之间的关联性，可以采用概率图模型。具体地，可以采用条件随机场（Conditional Random Field，CRF）执行序列标注任务。相对于隐马尔可夫模型（Hidden Markov Model，HMM），最大熵马尔可夫模型（Maximum Entropy Markov Model，MEMM）来说，CRF （条件随机场）的能够解决标签偏置问题。

CRF 是无向图模型，是在给定需要标记的观察序列的条件下，计算整个标记序列的联合概率分布，从而决定最佳标记序列，而不是在给定当前状态条件下，定义下一个状态的状态分布。与 HMM 比较，CRF 没有 HMM 那样严格的独立性假设条件，因而可以容纳任意的上下文信息。与 MEMM 比较，由于 CRF 计算全局最优输出节点的条件概率，因此，可以克服 MEMM 的标签偏置问题。

CRF 层能从训练数据中获得约束性的规则，CRF 层可以为最后预测的标签

添加一些约束来保证预测的标签是合法的。在训练数据训练过程中，这些约束可以通过 CRF 层自动学习到。这些约束可以是：句子中第一个词总是以标签“B-”或 “O”开始，而不是“I-”；标签“B-label1 I-label2 I-label3 I-…”,label1, label2, label3 应该属于同一类实体。例如，“B-Person I-Person” 是合法的序列，但是“B-Person I-Organization”是非法标签序列。标签序列“O I-label” is 非法的.实体标签的首个标签应该是“B-”，而非 “I-”，换句话说,有效的标签序列应该是“O B-label”等。有了这些约束，标签序列预测中非法序列出现的概率将会大大降低。

采用 CRF 进行序列标注时，具体可以确定每个所述第一特征向量对应的域标签概率；根据域标签概率确定所述第一特征向量的标签转移关系；根据标签转移关系对所述第一特征向量进行序列组合排序，得到组合排序后的特征向量；并对组合排序后的特征向量标注域标签。

其中，确定每个所述第一特征向量对应的域标签概率可以理解为对每一个第一特征向量，都计算得到其对应的域标签概率。例如：第一特征向量为{A1、A2、A3、A4}标签类型可以包括：名称标签、地址标签、领域标签和后缀标签，此时，针对 A1，可以计算得到 A1 对应所有标签的概率分布，比如，计算得到 A1 对应的所有标签概率分布可以为： 1.5 (B-Loc)，0.9 (I-Loc) ，0.2 (B- Name)，0.4(I-Name)，0.05 (O)，0.8 (B-Suffix)，0.9 (I- Suffix)。针对其他的特征向量可以采用相同的方法计算概率分布。

例如：将公司名称“北京 X1X2Y1Y2 有限公司”，进行分词后得到北/京/X1/X2 /Y1/Y2/有/限/公/司/，将分词映射为词向量:C1/C2/C3/C4/C5/C6/C7/C8/C9/C10，采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，可以得到第一特征向量：c1/c2/c3/c4/ c5/c6/c7/c8/c9/c10，第一特征向量中各个向量注意力权重是根据上下文特征计算得到的。采用训练完成的条件随机场模型对所述第一特征向量进行序列标注时，可以结合上下文特征对第一特征向量组合排序，比如：结合上下文特征将 c1/c2 合并，打上地域标签，

将 c3/c4 合并，打上名称标签，将 c5/c6 合并，打上领域标签，将 C7/C8/C9/C10 合并，打上后缀标签。

通过上述方法，采用基于概率图模型的 CRF 模型对待匹配的公司名称对应的特征向量进行序列标注，得到携带域标签的特征向量，通过优化标签序列和文本序列的联合概率，可以捕捉实体域标签之间的概率转移关系。

上述方法中，采用条件随机场模型（CRF）来进行序列标注，在进行标注之前，需要对条件随机场模型进行训练，具体的训练过程可以采用以下方法:

所述采用训练完成的条件随机场模型对所述第一特征向量进行序列标注之前，还可以包括:

获取域标签已知的实体名称样本;

提取所述实体名称样本对应的第三特征向量;

将所述第三特征向量输入待训练的条件随机场模型进行训练，得到所述待训练的条件随机场模型输出的对所述第三特征向量的域标签标注结果;

将全部所述实体名称样本中的域标签标注结果与全部所述实体名称样本的已知域标签进行比对，得到比对结果;

当所述比对结果表示所述全部实体名称样本中的域标签标注结果与所述实体名称样本的已知域标签相比，准确率达到预设阈值时，得到训练完成的条件随机场模型。

域标签已知可以理解为已经根据已有规则确定的域标签，例如：实体名称样本可以为域标签已知的公司名称集合或者其他账户名称集合。

提取实体样本对应的特征向量，将得到的特征向量作为输入量，输入待训练的条件随机场模型进行训练，得到所述待训练的条件随机场模型输出的对所述第三特征向量的域标签标注结果；将全部所述实体名称样本中的域标签标注结果与全部所述实体名称样本的已知域标签进行比对，当比对结果表示所述全部实体名称样本中的域标签标注结果与所述实体名称样本的已知域标签相比，准确率达到预设阈值时，可以认为训练结果收敛，完成训练，得到训练完成的

条件随机场模型。

具体地，假设实体样本为[A,B]，则输入量可以为 A = a1,a2,a3,……,an，输出可以为 B = b1,b2,b3,……,bn，其中，an 可以表示向量序列中的第 n 个词向量，A 可以表示由词向量构成的实体名称，bn 可以表示 an 对应的标注的域标签，B 可以表示标注域标签中构成的序列，从实体样本中提取特征向量对待训练的命名实体识别模型进行训练，直到待训练的命名实体识别模型收敛，从而得到命名实体识别模型。

在采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到第一特征向量，并采用训练完成的条件随机场模型对所述第一特征向量进行序列标注，得到携带有域标签的第二特征向量之后，可以采用实体匹配模型进行匹配：具体的匹配过程可以包括：

按照域标签分类将所述第二特征向量与所述实体名单中相同域的特征向量进行对齐匹配，得到每个域标签对应的相似度匹配分数；

将所述每个域标签对应的相似度匹配分数进行加权得到所述第二特征向量的匹配分数；

当所述匹配分数大于预设分值时，将匹配分数最高的实体名称作为所述待匹配实体名称的匹配结果。

需要说明的是，上述方法中的实体名单可以是已知的实体名单，比如：已知的制裁名单、已知的风险名单等。通过比对待匹配的实体名称与已知名单中的实体名称是否匹配，来完成匹配过程。

在实际的实施过程中，可以按照分域结果，按照域进行匹配，比如：将携带有名称域标签的特征向量与实体名单中的名称域中的特征向量进行对齐匹配，携带有地址域标签的特征向量与实体名单中的地址域中的特征向量进行对齐匹配，携带有领域标签的特征向量与实体名单中的领域中的特征向量进行对齐匹配等等。

在进行实体匹配时，具体可以采用多种方式进行匹配，例如：距离度量、

相似性评价、误写纠正、音节比较、翻译比较以及简称匹配等，本方案对此不进行限定。

所述将所述第二特征向量输入到实体匹配模型中，得到匹配结果之后，还可以包括:

当所述匹配结果表示所述匹配分数大于预设分值时，停止对于所述待匹配实体名称对应的交易数据的处理过程。

当待匹配实体名称与已知的实体名单中的实体名称匹配成功，可以认为待匹配实体名称存在风险或者该名称应当遵循已知的实体名单遵循权限设置。在交易场景中，匹配成功的实体名称作为账户不能继续进行交易，涉及的交易数据应当被停止，进一步审核或者直接限制交易。在实际的交易过程中，当交易双方的账户名称与已知的实体名称名单中的实体名称匹配成功时，该交易停止，系统会向交易双方发送用于提示交易失败的提示信息。

通过上述方法，通过 基于 Transformer 特征抽取层,使用自注意力机制代替经典的 CNN/RNN，自注意力机制通过分词后的词向量学习注意力权重，直接捕捉字词上下文的依赖关系；基于概率图模型的 CRF 模型，通过优化标签序列和文本序列的联合概率，可以捕捉实体域标签之间的概率转移关系，从而降低了实体名称的匹配失误率，提高了实体名称的匹配效率。

上述实施例中，通过引入实体识别算法来解决公司名匹配问题中的非关键匹配、误写匹配等问题，对输入的词向量进行上下文特征提取以及序列标注的方式可以是采用以下方式：包括把常见的 NER 方案如 CRF、HMM、MEMM、CNN 模型+CRF，RNN 模型（如 RNN、LSTM、GRU、Bi-LSTM、Bi-GRU 等）+CRF 以及其它的命名实体匹配算法移植到企业名匹配问题中。比如:
Bi-LSTM+CRF 的实体识别模型结构。

上述实施例中的特征抽取层采用基于自注意力机制的 Transformer。除此之外，还可以考虑基于 Transformer 进行一些优化，比如：除了调节参数之外，进行 Transformer 与 CNN、RNN 系列模型的组合，如在编码器和解码器之间加

入 CNN 层、在解码器后再加入 RNN 模型等，以更好的捕捉上下文联系。

本说明书实施例中的技术方案具有广泛的适应性，不仅适用于公司实体名称的匹配，还广泛适用于以命名实体识别任务为代表的序列标注相关的其它任务，如词性标注、文本翻译、各种实体名称匹配等。

基于同样的思路，本说明书实施例还提供了上述方法对应的装置。图 3 为本说明书实施例提供的对应于图 2 的一种实体名称匹配装置的结构示意图。如图 3 所示，该装置可以包括:

待匹配实体名称获取模块 302，用于获取待匹配实体名称;

词向量确定模块 304，用于对所述待匹配实体名称进行分词，并将所述待匹配实体名称的分词映射为向量，得到所述待匹配实体名称的词向量;

特征抽取模块 306，用于采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到第一特征向量;

序列标注模块 308，用于采用训练完成的条件随机场模型对所述第一特征向量进行序列标注，得到第二特征向量，所述第二特征向量为携带有域标签的特征向量;

匹配模块 310，用于将所述第二特征向量输入到实体匹配模型中，得到匹配结果。

可选的，所述待匹配实体名称获取模块 302，具体可以包括:

待匹配交易数据获取单元，用于获取待匹配交易数据;

实体名称提取单元，用于从所述交易数据中提取出交易双方的账户实体名称。

可选的，所述特征抽取模块 306，具体可以用于:

根据所述词向量的上下文信息采用自注意力机制计算每个词向量的权重值;

根据所述权重值对每个所述词向量进行注意力权重赋值，得到第一特征向量。

可选的，所述特征抽取模块 306，具体可以用于:

对于任意一个所述词向量，计算所述实体名称中的其他词向量对该词向量的注意力权重;

对所述任意一个所述词向量的注意力权重进行归一化;

将进行归一化后的权重进行加权求和，得到每个词向量的权重值。

可选的，所述序列标注模块 308，具体可以包括:

域标签概率确定单元，用于确定每个所述第一特征向量对应的域标签概率;

标签转移关系确定单元，用于根据所述域标签概率确定所述第一特征向量的标签转移关系;

组合排序单元，用于根据所述标签转移关系对所述第一特征向量进行序列组合排序，得到组合排序后的特征向量;

域标签标注单元，用于对所述组合排序后的特征向量标注域标签，得到第二特征向量。

可选的，所述域标签包括：名称标签、地址标签、领域标签、后缀标签和/或其他标签。

可选的，所述装置，还可以包括:

实体名称样本获取模块，用于获取域标签已知的实体名称样本;

特征向量提取模块，用于提取所述实体名称样本对应的第三特征向量;

训练模块，用于将所述第三特征向量输入待训练的条件随机场模型进行训练，得到所述待训练的条件随机场模型输出的对所述第三特征向量的域标签标注结果;

比对模块，用于将全部所述实体名称样本中的域标签标注结果与全部所述实体名称样本的已知域标签进行比对，得到比对结果；当所述比对结果表示所述全部实体名称样本中的域标签标注结果与所述实体名称样本的已知域标签相比，准确率达到预设阈值时，得到训练完成的条件随机场模型。

可选的，所述实体匹配模型中包括实体名称名单，所述实体名称名单包括公司名称。

可选的，所述匹配模块 310，具体可以包括:

对齐匹配单元，用于按照域标签分类将所述第二特征向量与所述实体名单中相同域的特征向量进行对齐匹配，得到每个域标签对应的相似度匹配分数;

加权单元，用于将所述每个域标签对应的相似度匹配分数进行加权得到所述第二特征向量的匹配分数;

匹配结果确定单元，用于当所述匹配分数大于预设分值时，将匹配分数最高的实体名称作为所述待匹配实体名称的匹配结果。

可选的，所述装置，还可以包括:

交易停止模块，用于当所述匹配结果表示所述匹配分数大于预设分值时，停止对于所述待匹配实体名称对应的交易数据的处理过程。

基于同样的思路，本说明书实施例还提供了上述方法对应的设备。图 4 为本说明书实施例提供的对应于图 2 的一种实体名称匹配设备的结构示意图。如图 4 所示，设备 400 可以包括:

至少一个处理器 410; 以及,

与所述至少一个处理器通信连接的存储器 430; 其中,

所述存储器 430 存储有可被所述至少一个处理器 410 执行的指令 420，所述指令被所述至少一个处理器 410 执行。

所述指令可以使所述至少一个处理器 410 能够:

获取待匹配实体名称;

对所述待匹配实体名称进行分词，并将所述待匹配实体名称的分词映射为向量，得到所述待匹配实体名称的词向量;

采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到第一特征向量;

采用训练完成的条件随机场模型对所述第一特征向量进行序列标注，得到

第二特征向量，所述第二特征向量为携带有域标签的特征向量;

将所述第二特征向量输入到实体匹配模型中，得到匹配结果。

基于同样的思路，本说明书实施例还提供了上述方法对应的计算机可读介质。计算机可读介质上存储有计算机可读指令，所述计算机可读指令可被处理器执行以实现以下方法:

获取待匹配实体名称;

对所述待匹配实体名称进行分词，并将所述待匹配实体名称的分词映射为向量，得到所述待匹配实体名称的词向量;

采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到第一特征向量;

采用训练完成的条件随机场模型对所述第一特征向量进行序列标注，得到第二特征向量，所述第二特征向量为携带有域标签的特征向量;

将所述第二特征向量输入到实体匹配模型中，得到匹配结果。

上述对本说明书特定实施例进行了描述。其它实施例在所附权利要求书的范围内。在一些情况下，在权利要求书中记载的动作或步骤可以按照不同于实施例中的顺序来执行并且仍然可以实现期望的结果。另外，在附图中描绘的过程不一定要求示出的特定顺序或者连续顺序才能实现期望的结果。在某些实施方式中，多任务处理和并行处理也是可以的或者可能是有利的。

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。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、RHDL（Ruby Hardware Description Language）等，目前最普遍使用的是 VHDL（Very-High-Speed Integrated Circuit Hardware Description Language）与 Verilog。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。一种典型的实现设备为计算机。具体的，计算机例如可以为个人计算机、膝上型计算机、蜂窝电话、相机电话、智能电话、个人数字助理、媒体播放器、导航设备、电子邮件设备、游戏控制台、平板计算机、可穿戴设备或者这些设备中的任何设备的组合。

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，在实施本说明书一个或多个实施例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。

本领域内的技术人员应明白，本说明书一个或多个实施例可提供为方法、系统、或计算机程序产品。因此，本说明书一个或多个实施例可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本说明书一个或多个实施例可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

本说明书一个或多个实施例是参照根据本说明书一个或多个实施例的方法、设备（系统）、和计算机程序产品的流程图和／或方框图来描述的。应理解可由计算机程序指令实现流程图和／或方框图中的每一流程和／或方框、以及流程图和／或方框图中的流程和／或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个

流程和／或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能的步骤。

在一个典型的配置中，计算设备包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。

内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器(RAM)和/或非易失性内存等形式，如只读存储器(ROM)或闪存(flash RAM)。内存是计算机可读介质的示例。

计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括暂存电脑可读媒体(transitory media)，如调制的数据信号和载波。

还需要说明的是，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、方法、商品或者设备不仅包括那些要素，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、方法、商品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

本说明书一个或多个实施例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说明书一个或多个实施例，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。

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。

以上所述仅为本说明书的实施例而已，并不用于限制本说明书一个或多个实施例。对于本领域技术人员来说，本说明书一个或多个实施例可以有各种更改和变化。凡在本说明书一个或多个实施例的精神和原理之内所作的任何修改、等同替换、改进等，均应包含在本说明书一个或多个实施例的权利要求范围之内。

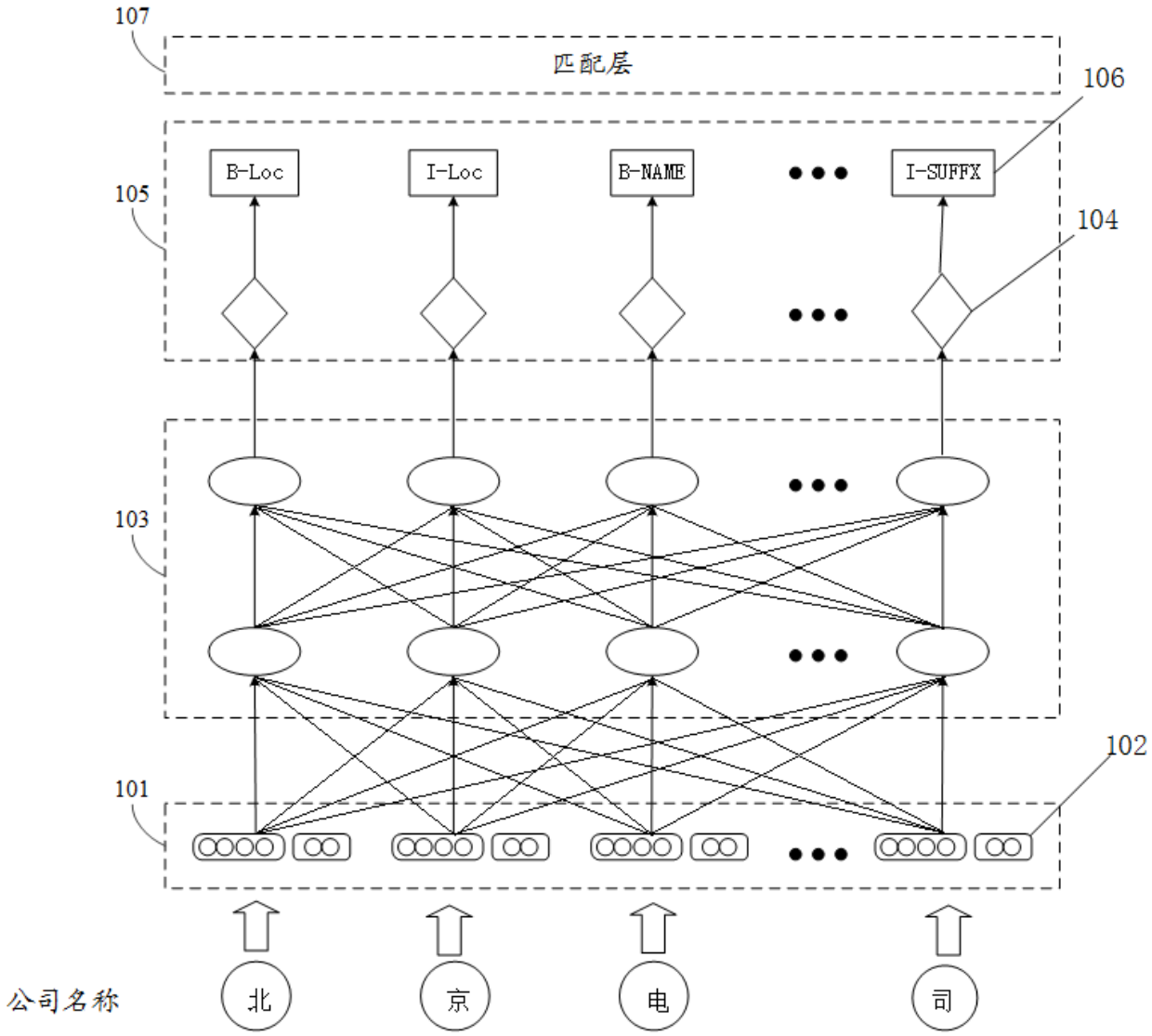

107
匹配层
106
B-Loc
I-Loc
B-NAME
I-SUFFX
105
104
103
102
101
公司名称
北
京
电
司

图 1

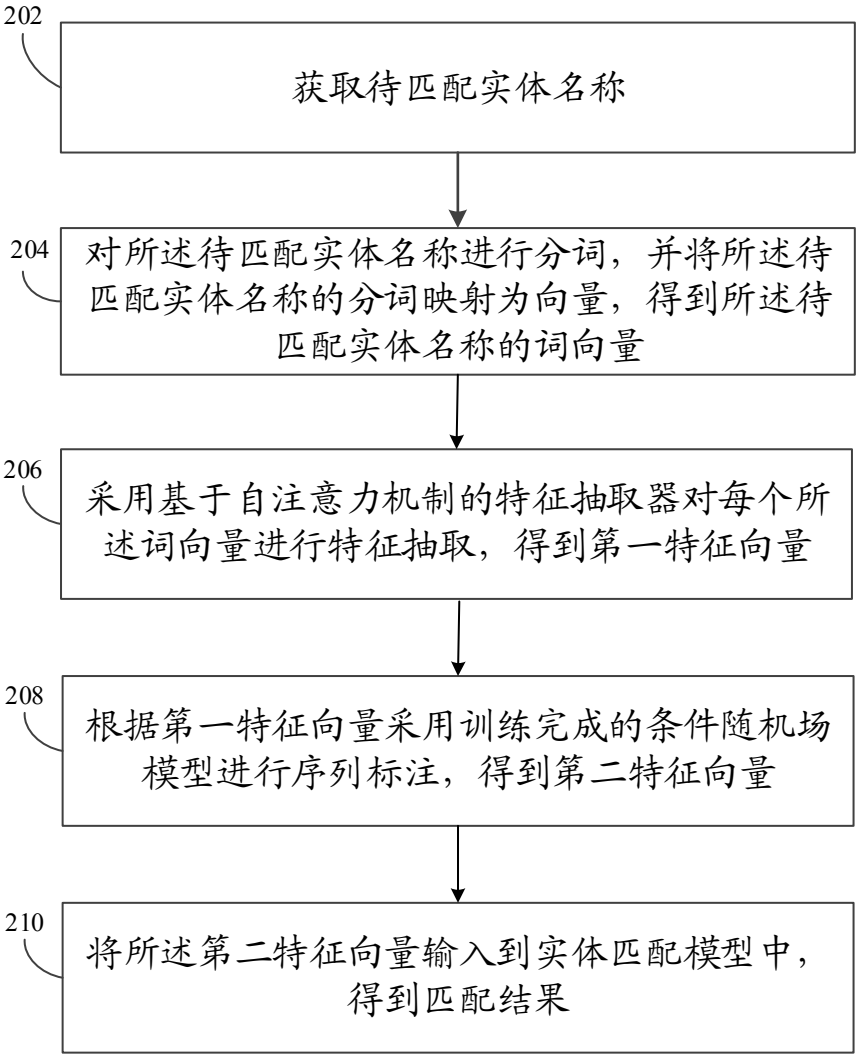

202
获取待匹配实体名称
204
对所述待匹配实体名称进行分词，并将所述待匹配实体名称的分词映射为向量，得到所述待匹配实体名称的词向量
206
采用基于自注意力机制的特征抽取器对每个所述词向量进行特征抽取，得到第一特征向量
208
根据第一特征向量采用训练完成的条件随机场模型进行序列标注，得到第二特征向量
210
将所述第二特征向量输入到实体匹配模型中，得到匹配结果

图 2

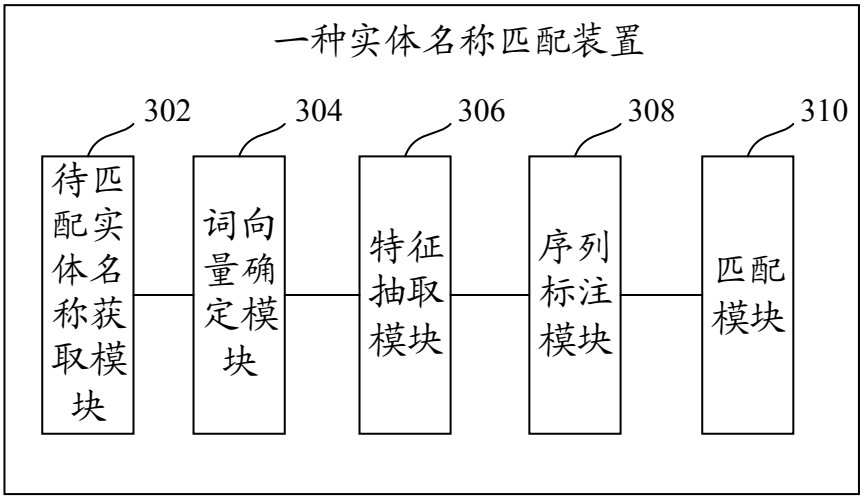

一种实体名称匹配装置
302
304
306
308
310
待匹配实体名称获取模块
词向量确定模块
特征抽取模块
序列标注模块
匹配模块

图 3

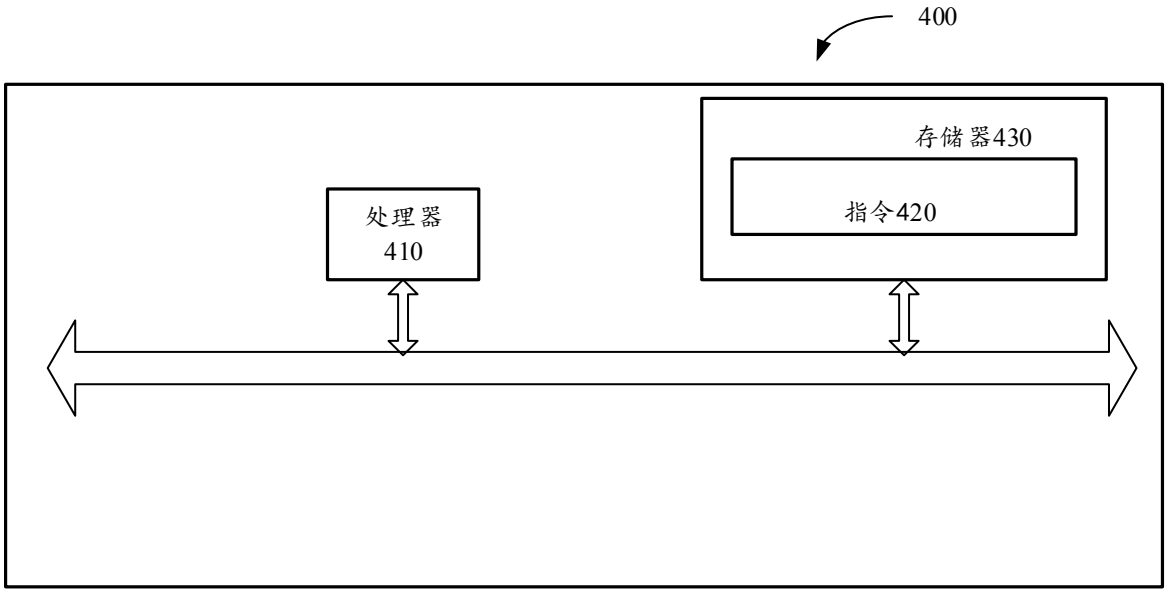

400
存储器430
指令420
处理器
410

图 4