



(12) 发明专利

(10) 授权公告号 CN 107392781 B

(45) 授权公告日 2021. 11. 02

(21) 申请号 201710469690.9

H04M 1/72403 (2021.01)

(22) 申请日 2017.06.20

审查员 朱凤阳

(65) 同一申请的已公布的文献号

申请公布号 CN 107392781 A

(43) 申请公布日 2017.11.24

(73) 专利权人 挖财网络技术有限公司

地址 310012 浙江省杭州市西湖区华星路
96号第18层

(72) 发明人 尤志强 潘琪 车曦

(74) 专利代理机构 北京博思佳知识产权代理有

限公司 11415

代理人 林祥 李威

(51) Int. Cl.

G06Q 50/00 (2012.01)

G06K 9/62 (2006.01)

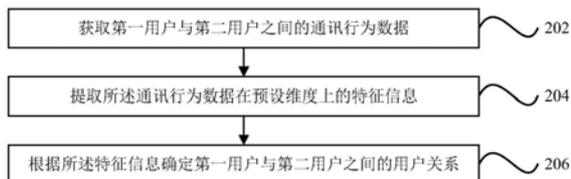
权利要求书3页 说明书16页 附图4页

(54) 发明名称

用户关系的识别方法、对象关系的识别方法及装置

(57) 摘要

本申请提供一种用户关系的识别方法、对象关系的识别方法及装置,该用户关系的识别方法可以包括:获取第一用户与第二用户之间的通讯行为数据;提取所述通讯行为数据在预设维度上的特征信息;根据所述特征信息确定所述第一用户与所述第二用户之间的用户关系。



1. 一种对象关系的识别方法,其特征在于,包括:
获取第一对象与第二对象之间的通讯行为数据;
提取所述通讯行为数据在预设维度上的特征信息;
根据所述特征信息确定所述第一对象与所述第二对象之间的对象关系,包括:通过关系召回模型确定所述第一对象与所述第二对象之间的对象关系;其中,所述关系召回模型由基于历史对象通讯行为数据生成的第一训练样本进行训练得到;所述关系召回模型包括对应于不同类型的多个分类器,每一分类器用于确定所述第一对象与所述第二对象是否属于特定类型的对象关系;其中,所述第一训练样本中的部分数据包含标注的所述特定类型的对象关系信息,以通过半监督算法训练得到所述分类器。
2. 根据权利要求1所述的方法,其特征在于,所述通讯行为数据包括以下至少之一:
语音通话记录数据、视频通话记录数据、通讯消息收发记录数据。
3. 根据权利要求1所述的方法,其特征在于,所述通讯行为数据包括所述第一对象与所述第二对象在预设时间段内的通讯行为的描述信息。
4. 根据权利要求1所述的方法,其特征在于,所述第一训练样本包括所述历史对象通讯行为数据在所述预设维度上的特征信息。
5. 根据权利要求4所述的方法,其特征在于,所述第一训练样本中的特征信息包括:
所述历史对象通讯行为数据在所述预设维度上的原始特征;
或者,由所述原始特征经过压缩转换得到的压缩特征。
6. 根据权利要求5所述的方法,其特征在于,所述原始特征通过变分自编码器被压缩转换为相应的压缩特征。
7. 根据权利要求1、5或6中任一项所述的方法,其特征在于,所述预设维度与下述信息中至少之一相关:
通讯发起时刻、通讯时长、通讯角色、通讯类型。
8. 根据权利要求1所述的方法,其特征在于,所述半监督算法基于深度生成模型对所述第一训练样本进行训练,以得到所述分类器。
9. 根据权利要求1所述的方法,其特征在于,所述根据所述特征信息确定所述第一对象与所述第二对象之间的对象关系,包括:
根据所述特征信息确定所述第一对象与所述第二对象之间的关系重要程度;
当所述关系重要程度大于预设重要程度时,根据所述特征信息确定所述第一对象与所述第二对象之间的对象关系。
10. 根据权利要求9所述的方法,其特征在于,所述根据所述特征信息确定所述第一对象与所述第二对象之间的关系重要程度,包括:
通过重要性识别模型确定所述第一对象与所述第二对象之间的关系重要程度;其中,所述重要性识别模型由基于历史对象通讯行为数据生成的第二训练样本进行训练得到。
11. 根据权利要求10所述的方法,其特征在于,所述第二训练样本包含标注的对象关系信息,以通过监督算法训练得到所述重要性识别模型。
12. 根据权利要求11所述的方法,其特征在于,所述监督算法包括梯度提升决策树算法。
13. 一种对象关系的识别装置,其特征在于,包括:

获取单元,获取第一对象与第二对象之间的通讯行为数据;

提取单元,提取所述通讯行为数据在预设维度上的特征信息;

确定单元,根据所述特征信息确定所述第一对象与所述第二对象之间的对象关系,包括:通过关系召回模型确定所述第一对象与所述第二对象之间的对象关系;其中,所述关系召回模型由基于历史对象通讯行为数据生成的第一训练样本进行训练得到;所述关系召回模型包括对应于不同类型的多个分类器,每一分类器用于确定所述第一对象与所述第二对象是否属于特定类型的对象关系;其中,所述第一训练样本中的部分数据包含标注的所述特定类型的对象关系信息,以通过半监督算法训练得到所述分类器。

14. 根据权利要求13所述的装置,其特征在於,所述通讯行为数据包括以下至少之一:

语音通话记录数据、视频通话记录数据、通讯消息收发记录数据。

15. 根据权利要求13所述的装置,其特征在於,所述通讯行为数据包括所述第一对象与所述第二对象在预设时间段内的通讯行为的描述信息。

16. 根据权利要求13所述的装置,其特征在於,所述第一训练样本包括所述历史对象通讯行为数据在所述预设维度上的特征信息。

17. 根据权利要求16所述的装置,其特征在於,所述第一训练样本中的特征信息包括:

所述历史对象通讯行为数据在所述预设维度上的原始特征;

或者,由所述原始特征经过压缩转换得到的压缩特征。

18. 根据权利要求17所述的装置,其特征在於,所述原始特征通过变分自编码器被压缩转换为相应的压缩特征。

19. 根据权利要求13、17或18中任一项所述的装置,其特征在於,所述预设维度与下述信息中至少之一相关:

通讯发起时刻、通讯时长、通讯角色、通讯类型。

20. 根据权利要求13所述的装置,其特征在於,所述半监督算法基于深度生成模型对所述第一训练样本进行训练,以得到所述分类器。

21. 根据权利要求13所述的装置,其特征在於,所述确定单元具体用于:

根据所述特征信息确定所述第一对象与所述第二对象之间的关系重要程度;

当所述关系重要程度大于预设重要程度时,根据所述特征信息确定所述第一对象与所述第二对象之间的对象关系。

22. 根据权利要求21所述的装置,其特征在於,所述确定单元通过下述方式,实现根据所述特征信息确定所述第一对象与所述第二对象之间的关系重要程度:

通过重要性识别模型确定所述第一对象与所述第二对象之间的关系重要程度;其中,所述重要性识别模型由基于历史对象通讯行为数据生成的第二训练样本进行训练得到。

23. 根据权利要求22所述的装置,其特征在於,所述第二训练样本包含标注的对象关系信息,以通过监督算法训练得到所述重要性识别模型。

24. 根据权利要求23所述的装置,其特征在於,所述监督算法包括梯度提升决策树算法。

25. 一种对象关系的识别装置,其特征在於,包括:

处理器;

用于存储处理器可执行指令的存储器;

其中,所述处理器被配置为实现如权利要求1-12中任一项所述的方法。

用户关系的识别方法、对象关系的识别方法及装置

技术领域

[0001] 本申请涉及信息处理技术领域,尤其涉及一种用户关系的识别方法、对象关系的识别方法及装置。

背景技术

[0002] 在用户之间可以通过社交等方式建立起相应的用户关系(对于用户之外的其他类型的对象,可以称为对象关系)。用户关系可以被应用于各种场景中,比如通过用户A的信息以及该用户A与另一用户B之间的用户关系,可以推测出用户B的信息。

[0003] 在相关技术中,用户可以在通讯录中记录其他用户的信息,因而通过获取用户的通讯录数据,即可基于该通讯录数据包含的用户信息,确定用户之间的用户关系。比如,当用户A在通讯录中记录了用户B的信息为“爸”时,可以确定用户A与用户B之间属于“直系亲属”的用户关系。

[0004] 但是,用户并不会在通讯录中记录所有用户的信息,并且记录的信息并不总是能够表现出用户关系;比如,当用户A将用户B的信息记录为姓名时,就无法据此确定用户A与用户B之间的用户关系。

发明内容

[0005] 本申请提供了用户关系、对象关系的识别方案,该方案可以包括如下内容:

[0006] 根据本申请的第一方面,提出了一种用户关系的识别方法,包括:

[0007] 获取第一用户与第二用户之间的通讯行为数据;

[0008] 提取所述通讯行为数据在预设维度上的特征信息;

[0009] 根据所述特征信息确定所述第一用户与所述第二用户之间的用户关系。

[0010] 根据本申请的第二方面,提出了一种对象关系的识别方法,包括:

[0011] 获取第一对象与第二对象之间的通讯行为数据;

[0012] 提取所述通讯行为数据在预设维度上的特征信息;

[0013] 根据所述特征信息确定所述第一对象与所述第二对象之间的对象关系。

[0014] 根据本申请的第三方面,提出了一种用户关系的识别装置,包括:

[0015] 获取单元,获取第一用户与第二用户之间的通讯行为数据;

[0016] 提取单元,提取所述通讯行为数据在预设维度上的特征信息;

[0017] 确定单元,根据所述特征信息确定所述第一用户与所述第二用户之间的用户关系。

[0018] 根据本申请的第四方面,提出了一种对象关系的识别装置,包括:

[0019] 获取单元,获取第一对象与第二对象之间的通讯行为数据;

[0020] 提取单元,提取所述通讯行为数据在预设维度上的特征信息;

[0021] 确定单元,根据所述特征信息确定所述第一对象与所述第二对象之间的对象关系。

[0022] 由以上技术方案可见,本申请通过分析用户或其他类型的对象之间的通讯行为数据,可以根据该通讯行为数据包含的特征信息,准确识别出相应的用户关系或对象关系,而无需用户之间或对象之间相互标注的信息(如通讯录等处标注的信息)。

附图说明

- [0023] 图1是本申请一示例性实施例提供的一种用户关系的识别系统的架构示意图。
- [0024] 图2是本申请一示例性实施例提供的一种用户关系的识别方法的流程图。
- [0025] 图3是本申请一示例性实施例提供的一种对象关系的识别方法的流程图。
- [0026] 图4是本申请一示例性实施例提供的一种基于通话行为数据识别用户关系的示意图。
- [0027] 图5是本申请一示例性实施例提供的一种识别用户关系的示意图。
- [0028] 图6是本申请一示例性实施例提供的一种电子设备的结构示意图。
- [0029] 图7是本申请一示例性实施例提供的一种用户关系的识别装置的框图。
- [0030] 图8是本申请一示例性实施例提供的另一种电子设备的结构示意图。
- [0031] 图9是本申请一示例性实施例提供的一种对象关系的识别装置的框图。

具体实施方式

[0032] 本申请通过分析用户之间的通讯行为数据,可以根据该通讯行为数据包含的特征信息,准确识别出相应的用户关系;类似地,通过分析对象之间的通讯行为数据,可以根据该通讯行为数据包含的特征信息,准确识别出相应的对象关系。

[0033] 为对本申请进行进一步说明,提供下列实施例:

[0034] 图1是本申请一示例性实施例提供的一种用户关系的识别系统的架构示意图。如图1所示,该系统可以包括服务器11、网络12、若干电子设备,比如手机13、PC14、手机15和PC16等。

[0035] 服务器11可以为包含一独立主机的物理服务器,或者该服务器11可以为主机集群承载的虚拟服务器,或者该服务器11可以为云服务器。在运行过程中,服务器11可以运行某一应用的服务器侧的程序,以实现该应用的相关功能,比如对用户关系或对象关系的识别功能等。

[0036] 手机13、PC14、手机15、PC16均为用户可以使用的部分类型的电子设备。实际上,用户显然还可以使用诸如下述类型的电子设备:平板设备、笔记本电脑、掌上电脑(PDAs, Personal Digital Assistants)、可穿戴设备(如智能眼镜、智能手表等)等,本申请并不对此进行限制。在运行过程中,该电子设备可以运行某一应用的客户端侧的程序,以实现该应用的相关功能,比如通讯功能等,以用于在用户之间或对象之间实现相应的通讯行为。

[0037] 而对于手机13(或PC14、手机15、PC16等)与服务器11之间进行交互的网络12,可以包括多种类型的有线或无线网络。在一实施例中,该网络12可以包括公共交换电话网络(Public Switched Telephone Network, PSTN)和因特网。同时,手机13(或PC14)、手机15(或PC16)等电子设备之间也可以通过该网络12进行通讯交互,比如在任意两台电子设备之间实现通讯,或者在两台以上的电子设备之间实现群体通讯等。

[0038] 图2是本申请一示例性实施例提供的一种用户关系的识别方法的流程图。假定第

一用户使用图1所示的手机13或PC14、第二用户使用图1所示的手机15或PC16,以实现第一用户与第二用户之间的通讯行为,图2所示的方法可以应用于该第一用户使用的手机13或PC14等电子设备,也可以应用于该第二用户使用的手机15或PC16等电子设备,还可以应用于诸如图1所示的服务器11,或者应用于其他任意电子设备;如图2所示,该方法可以包括以下步骤:

[0039] 步骤202,获取第一用户与第二用户之间的通讯行为数据(出于信息隐私和数据安全方面的考虑,可在经由用户授权后,基于用户授权来获取通讯行为数据,避免造成侵犯用户隐私、泄露用户数据等问题)。

[0040] 在一实施例中,步骤202中获取的通讯行为数据可以包括任意通讯行为产生的数据。例如,通讯行为可以包括语音通话行为、视频通话行为、通讯消息收发行为等,而相应的通讯行为数据包括语音通话记录数据、视频通话记录数据、通讯消息收发记录数据等,本申请并不对此进行限制;其中,通讯消息可以包括任意类型,比如短信、邮件、即时通讯消息、社交网络消息等,本申请并不对此进行限制。

[0041] 在一实施例中,第一用户与第二用户之间的通讯行为,可以理解为第一用户与第二用户参与到同一通讯事件中的行为。在一实施方式中,该通讯事件可以为第一用户与第二用户之间的双方通讯事件,可以由第一用户和第二用户中的任一方作为发起方、另一方作为响应方,由该发起方向响应方发起通讯请求、响应方响应于该通讯请求,以创建该通讯事件。在另一实施方式中,该通讯事件可以为第一用户与第二用户之间的双方通讯事件,可以由第一用户和第二用户之外的其他用户创建该通讯事件,而第一用户和第二用户分别加入该通讯事件。在又一实施方式中,该通讯事件可以为第一用户、第二用户以及至少一个其他用户之间的多方通讯事件,可以由第一用户、第二用户或任一其他用户作为发起方,由剩余的用户作为响应方,使得该发起方向响应方发起通讯请求、响应方响应于该通讯请求,以创建该通讯事件;或者,由第一用户、第二用户或任一其他用户创建该通讯事件,而剩余的用户加入该通讯事件。

[0042] 在一实施例中,步骤202中获取的通讯行为数据可以包括上述的一种或多种类型,比如仅获取语音通话记录数据,或者同时获取语音通话记录数据和视频通话记录数据等。

[0043] 在一实施例中,步骤202中获取的通讯行为数据可以包括所述第一用户与所述第二用户在预设时间段内的通讯行为的描述信息。预设时间段可以包括预定义的任意时间段,比如最近一个月、最近三年、从A年B月C日至D年E月F日之间、从第一用户与第二用户之间首次产生通讯行为至今等,本申请并不对此进行限制。

[0044] 在一实施例中,第一用户与第二用户可能在上述的预设时间段内实施了多种类型的通讯行为,步骤202中获取的通讯行为数据可以对应于其中的一种或多种类型的通讯行为,本申请并不对此进行限制。

[0045] 步骤204,提取所述通讯行为数据在预设维度上的特征信息。

[0046] 在一实施例中,预设维度可以包括通讯行为数据包含的任意维度,本申请并不对此进行限制。第一用户与第二用户之间的通讯行为,往往与两者之间的用户关系存在一定程度的关联,因而通过获取通讯行为数据在该预设维度上的特征信息,可以据此分析和确定出第一用户与第二用户的用户关系,且无需第一用户或第二用户主动在通讯录等处描述该用户关系。

[0047] 在一实施例中,预设维度可以与下述信息中至少之一相关:通讯发起时刻、通讯时长、通讯角色、通讯类型;其中,根据通讯行为的不同,上述各个预设维度的解释也可能存在一定变化。

[0048] 对于通讯发起时刻:通讯行为对应于通讯事件,当该通讯事件为第一用户与第二用户中的任一方作为发起方、另一方作为响应方时,通讯发起时刻可以包括该发起方向响应方发起通讯请求的时刻,或者该响应方响应于该通讯请求的时刻等;当该通讯事件为第一用户、第二用户以及至少一个其他用户之间的多方通讯事件时,通讯发起时刻可以包括该通讯事件对应的通讯请求被发起的时刻,或者通讯发起时刻可以包括第一用户与第二用户同时加入该通讯事件的初始时刻等。

[0049] 对于通讯时长:通讯时长可以包括通讯发起时刻与通讯终止时刻之间的时间段长度。其中,通讯终止时刻可以包括相应的通讯事件被终止的时刻,或者第一用户与第二用户中至少一方退出所述通讯事件的时刻等。

[0050] 对于通讯角色,可以包括通讯事件的发起方、响应方、参与方等。

[0051] 对于通讯类型,可以包括上述的通讯行为的类型,比如语音通话行为、视频通话行为、通讯消息收发行为等,此处不再赘述。

[0052] 步骤206,根据所述特征信息确定所述第一用户与所述第二用户之间的用户关系。

[0053] 在一实施例中,用户关系可以包括第一用户与第二用户之间的直接关系,比如亲属关系、同事关系、同学关系、合作伙伴关系等。在另一实施例中,用户关系可以包括第一用户与第二用户之间的间接关系,比如同事的朋友(二层关系)、同学的亲属的同事(三层关系)、朋友的合作伙伴的同学(三层关系)等,间接关系可能涉及到二层、三层甚至更多层级的用户关系,本申请并不限制该间接关系涉及到的关系层级数量。

[0054] 在一实施例中,可以通过关系召回模型确定所述第一用户与所述第二用户之间的用户关系;其中,所述关系召回模型可以由基于历史用户通讯行为数据生成的第一训练样本进行训练得到。历史用户通讯行为数据可以包括所有用户对应的所有历史通讯行为数据中的至少一部分;例如,历史用户通讯行为数据可以存在一个或多个数据来源,比如该数据来源可以包括:至少一台电子设备上的历史通讯记录数据(如历史通话数据、历史收发短信、历史即时通讯消息等),或者运营商处的历史通讯记录数据等。

[0055] 在一实施例中,所述第一训练样本可以包括所述历史用户通讯行为数据在所述预设维度上的特征信息。其中,所述预设维度可以参考上文所述,比如该预设维度可以与下述信息中至少之一相关:通讯发起时刻、通讯时长、通讯角色、通讯类型,此处不再赘述。

[0056] 在一实施例中,第一训练样本可以与步骤204中提取的特征信息对应于相同的预设维度,使得由第一训练样本训练得到的关系召回模型能够对步骤204中提取的特征信息进行更准确(相比于第一训练样本与步骤204中提取的特征信息对应于不同的预设维度的情况)的分析,从而更准确地判定第一用户与第二用户之间的用户关系。

[0057] 在一实施例中,第一训练样本中的特征信息可以包括:所述历史用户通讯行为数据在所述预设维度上的原始特征。

[0058] 在一实施例中,第一训练样本中的特征信息可以包括:由所述原始特征经过压缩转换得到的压缩特征。对原始特征进行压缩转换后,可以表现为特征维度的减少,比如原始特征为50维,而相应的压缩特征为10维。通过将原始特征压缩转换为压缩特征,一方面可以

对原始特征的各个维度之间的相同或相近的维度信息进行“去重”，另一方面可以对不同维度之间的维度信息进行交叉组合，可见该压缩转换过程不仅能够提升对数据波动的应对能力，而且压缩特征能够更恰当、更全面地描述相应的第一训练样本，还能进一步使得由该第一训练样本训练得到相应的关系召回模型时，可以提升该关系召回模型对第一用户与第二用户之间的用户关系进行召回时的准确度。

[0059] 在一实施例中，可以通过自编码器 (Auto Encoder) 将原始特征压缩转换为相应的压缩特征；其中，自编码器可以包括任意类型，比如变分自编码器 (Variational Auto-Encoder, VAE) 等，使得对于压缩特征的生成过程为数据驱动的自动处理过程，无需用户参与，有助于尽可能地减少主观因素的影响。

[0060] 在一实施例中，所述关系召回模型可以包括分类器，所述分类器用于确定所述第一用户与所述第二用户是否属于特定类型的用户关系；其中，所述第一训练样本中的部分数据包含标注的所述特定类型的用户关系信息，以通过半监督算法训练得到所述分类器。由于不同用户对同一类型的用户关系可能存在很多种描述方式，因而通过采用半监督算法进行训练时，可以在第一训练样本仅包含部分标注数据的情况下，自动学习所述特定类型下的其他可能的用户关系信息，极大地简化了对第一训练样本的标注操作。

[0061] 在一实施例中，所述半监督算法可以基于深度生成模型对所述第一训练样本进行训练，以得到所述分类器；相比于未采用深度生成模型的半监督算法，基于深度生成模型的半监督算法训练得到的分类器具有更优的分类效果。

[0062] 在一实施例中，所述关系召回模型可以包括分类器，所述分类器用于确定所述第一用户与所述第二用户是否属于特定类型的用户关系；其中，所述第一训练样本中的数据均包含标注的所述特定类型的用户关系信息，以通过监督算法训练得到所述分类器。

[0063] 在一实施例中，关系召回模型可以包括一个或多个分类器，每个分类器用于确定所述第一用户与所述第二用户是否属于一个或多个特定类型的用户关系。例如，关系召回模型可以包括一个工作关系类型的第一分类器、一个亲属关系类型的第二分类器，该第一分类器用于确定所述第一用户与所述第二用户是否属于工作关系类型的用户关系，该第二分类器用于确定所述第一用户与所述第二用户是否属于亲属关系类型的用户关系；再例如，关系召回模型可以包括同时对应于工作关系类型和亲属关系类型的分类器，该分类器用于确定所述第一用户与所述第二用户是否属于工作关系类型或亲属关系类型的用户关系。

[0064] 在一实施例中，可以根据所述特征信息确定所述第一用户与所述第二用户之间的关系重要程度；然后，当所述关系重要程度大于预设重要程度 (即重要关系) 时，根据所述特征信息确定所述第一用户与所述第二用户之间的用户关系。

[0065] 在一实施例中，当所述关系重要程度不大于预设重要程度 (即非重要关系) 时，可以避免根据所述特征信息确定所述第一用户与所述第二用户之间的用户关系，可以节省对非重要关系的特征信息进行识别的处理量。

[0066] 在一实施例中，可以通过重要性识别模型确定所述第一用户与所述第二用户之间的关系重要程度；其中，所述重要性识别模型由基于历史用户通讯行为数据生成的第二训练样本进行训练得到。

[0067] 在一实施例中，所述第二训练样本包含标注的用户关系信息，以通过监督算法训

练得到所述重要性识别模型。例如,监督算法包括梯度提升决策树(Gradient Boosting Decision Tree,GBDT)算法或相关衍生算法(如eXtreme Gradient Boosting,简称为Xgboost)。

[0068] 在一实施例中,所述第二训练样本中的部分数据包含标注的用户关系信息,以通过半监督算法训练得到所述重要性识别模型。

[0069] 图3是本申请一示例性实施例提供的一种对象关系的识别方法的流程图。假定第一对象使用图1所示的手机13或PC14、第二对象使用图1所示的手机15或PC16,以实现第一对象与第二对象之间的通讯行为,图3所示的方法可以应用于该第一对象使用的手机13或PC14等电子设备,也可以应用于该第二对象使用的手机15或PC16等电子设备,还可以应用于诸如图1所示的服务器11,或者应用于其他任意电子设备;如图3所示,该方法可以包括以下步骤:

[0070] 步骤302,获取第一对象与第二对象之间的通讯行为数据。

[0071] 在一实施例中,对象可以包括任意类型的操作体,多个操作体之间可以实施通讯行为,且相应的通讯行为数据可以表现出这些操作体之间存在的对象关系。例如,对象可以包括用户、团体(比如企业、学校、事业单位、政府机关、社会组织等,本申请并不对此进行限制)、人工智能等类型,本申请并不对此进行限制。

[0072] 在一实施例中,步骤302中获取的通讯行为数据可以包括任意通讯行为产生的数据。例如,通讯行为可以包括语音通话行为、视频通话行为、通讯消息收发行为等,而相应的通讯行为数据包括语音通话记录数据、视频通话记录数据、通讯消息收发记录数据等,本申请并不对此进行限制;其中,通讯消息可以包括任意类型,比如短信、邮件、即时通讯消息、社交网络消息等,本申请并不对此进行限制。

[0073] 在一实施例中,第一对象与第二对象之间的通讯行为,可以理解为第一对象与第二对象参与到同一通讯事件中的行为。在一实施方式中,该通讯事件可以为第一对象与第二对象之间的双方通讯事件,可以由第一对象和第二对象中的任一方作为发起方、另一方作为响应方,由该发起方向响应方发起通讯请求、响应方响应于该通讯请求,以创建该通讯事件。在另一实施方式中,该通讯事件可以为第一对象与第二对象之间的双方通讯事件,可以由第一对象和第二对象之外的其他对象创建该通讯事件,而第一对象和第二对象分别加入该通讯事件。在又一实施方式中,该通讯事件可以为第一对象、第二对象以及至少一个其他对象之间的多方通讯事件,可以由第一对象、第二对象或任一其他对象作为发起方,由剩余的对象作为响应方,使得该发起方向响应方发起通讯请求、响应方响应于该通讯请求,以创建该通讯事件;或者,由第一对象、第二对象或任一其他对象创建该通讯事件,而剩余的对象加入该通讯事件。

[0074] 在一实施例中,步骤302中获取的通讯行为数据可以包括上述的一种或多种类型,比如仅获取语音通话记录数据,或者同事获取语音通话记录数据和视频通话记录数据等。

[0075] 在一实施例中,步骤302中获取的通讯行为数据可以包括所述第一对象与所述第二对象在预设时间段内的通讯行为的描述信息。预设时间段可以包括预定义的任意时间段,比如最近一个月、最近三年、从A年B月C日至D年E月F日之间、从第一对象与第二对象之间首次产生通讯行为至今等,本申请并不对此进行限制。

[0076] 在一实施例中,第一对象与第二对象可能在上述的预设时间段内实施了多种类型

的通讯行为,步骤302中获取的通讯行为数据可以对应于其中的一种或多种类型的通讯行为,本申请并不对此进行限制。

[0077] 步骤304,提取所述通讯行为数据在预设维度上的特征信息。

[0078] 在一实施例中,预设维度可以包括通讯行为数据包含的任意维度,本申请并不对此进行限制。第一对象与第二对象之间的通讯行为,往往与两者之间的对象关系存在一定程度的关联,因而通过获取通讯行为数据在该预设维度上的特征信息,可以据此分析和确定出第一对象与第二对象的对象关系,且无需第一对象或第二对象主动在通讯录等处描述该对象关系。

[0079] 在一实施例中,预设维度可以与下述信息中至少之一相关:通讯发起时刻、通讯时长、通讯角色、通讯类型;其中,根据通讯行为的不同,上述各个预设维度的解释也可能存在一定变化。

[0080] 对于通讯发起时刻:通讯行为对应于通讯事件,当该通讯事件为第一对象与第二对象中的任一方作为发起方、另一方作为响应方时,通讯发起时刻可以包括该发起方向响应方发起通讯请求的时刻,或者该响应方响应于该通讯请求的时刻等;当该通讯事件为第一对象、第二对象以及至少一个其他对象之间的多方通讯事件时,通讯发起时刻可以包括该通讯事件对应的通讯请求被发起的时刻,或者通讯发起时刻可以包括第一对象与第二对象同时加入该通讯事件的初始时刻等。

[0081] 对于通讯时长:通讯时长可以包括通讯发起时刻与通讯终止时刻之间的时间段长度。其中,通讯终止时刻可以包括相应的通讯事件被终止的时刻,或者第一对象与第二对象中至少一方退出所述通讯事件的时刻等。

[0082] 对于通讯角色,可以包括通讯事件的发起方、响应方、参与方等。

[0083] 对于通讯类型,可以包括上述的通讯行为的类型,比如语音通话行为、视频通话行为、通讯消息收发行为等,此处不再赘述。

[0084] 步骤306,根据所述特征信息确定所述第一对象与所述第二对象之间的对象关系。

[0085] 在一实施例中,对象关系可以包括第一对象与第二对象之间的直接关系,比如亲属关系、同事关系、同学关系、合作伙伴关系等。在另一实施例中,对象关系可以包括第一对象与第二对象之间的间接关系,比如同事的朋友(二层关系)、同学的亲属的同事(三层关系)、朋友的合作伙伴的同学(三层关系)等,间接关系可能涉及到二层、三层甚至更多层级的对象关系,本申请并不限制该间接关系涉及到的关系层级数量。

[0086] 在一实施例中,可以通过关系召回模型确定所述第一对象与所述第二对象之间的对象关系;其中,所述关系召回模型可以由基于历史对象通讯行为数据生成的第一训练样本进行训练得到。历史对象通讯行为数据可以包括所有对象对应的所有历史通讯行为数据中的至少一部分;例如,历史对象通讯行为数据可以存在一个或多个数据来源,比如该数据来源可以包括:至少一台电子设备上的历史通讯记录数据(如历史通话数据、历史收发短信、历史即时通讯消息等),或者运营商处的历史通讯记录数据等。

[0087] 在一实施例中,所述第一训练样本可以包括所述历史对象通讯行为数据在所述预设维度上的特征信息。其中,所述预设维度可以参考上文所述,比如该预设维度可以与下述信息中至少之一相关:通讯发起时刻、通讯时长、通讯角色、通讯类型,此处不再赘述。

[0088] 在一实施例中,第一训练样本可以与步骤304中提取的特征信息对应于相同的预

设维度,使得由第一训练样本训练得到的关系召回模型能够对步骤304中提取的特征信息进行更准确(相比于第一训练样本与步骤304中提取的特征信息对应于不同的预设维度的情况)的分析,从而更准确地判定第一对象与第二对象之间的对象关系。

[0089] 在一实施例中,第一训练样本中的特征信息可以包括:所述历史对象通讯行为数据在所述预设维度上的原始特征。

[0090] 在一实施例中,第一训练样本中的特征信息可以包括:由所述原始特征经过压缩转换得到的压缩特征。对原始特征进行压缩转换后,可以表现为特征维度的减少,比如原始特征为50维,而相应的压缩特征为10维。通过将原始特征压缩转换为压缩特征,一方面可以对原始特征的各个维度之间的相同或相近的维度信息进行“去重”,另一方面可以对不同维度之间的维度信息进行交叉组合,可见该压缩转换过程不仅能够提升对数据波动的应对能力,而且压缩特征能够更恰当、更全面地描述相应的第一训练样本,还能进一步使得由该第一训练样本训练得到相应的关系召回模型时,可以提升该关系召回模型对第一对象与第二对象之间的对象关系进行召回时的准确度。

[0091] 在一实施例中,可以通过自编码器(Auto Encoder)将原始特征压缩转换为相应的压缩特征;其中,自编码器可以包括任意类型,比如变分自编码器(Variational Auto-Encoder,VAE)等,使得对于压缩特征的生成过程为数据驱动的自动处理过程,无需对象参与,有助于尽可能地减少主观因素的影响。

[0092] 在一实施例中,所述关系召回模型可以包括分类器,所述分类器用于确定所述第一对象与所述第二对象是否属于特定类型的对象关系;其中,所述第一训练样本中的部分数据包含标注的所述特定类型的对象关系信息,以通过半监督算法训练得到所述分类器。由于不同对象对同一类型的对象关系可能存在很多种描述方式,因而通过采用半监督算法进行训练时,可以在第一训练样本仅包含部分标注数据的情况下,自动学习所述特定类型下的其他可能的对象关系信息,极大地简化了对第一训练样本的标注操作。

[0093] 在一实施例中,所述半监督算法可以基于深度生成模型对所述第一训练样本进行训练,以得到所述分类器;相比于未采用深度生成模型的半监督算法,基于深度生成模型的半监督算法训练得到的分类器具有更优的分类效果。

[0094] 在一实施例中,所述关系召回模型可以包括分类器,所述分类器用于确定所述第一对象与所述第二对象是否属于特定类型的对象关系;其中,所述第一训练样本中的数据均包含标注的所述特定类型的对象关系信息,以通过监督算法训练得到所述分类器。

[0095] 在一实施例中,关系召回模型可以包括一个或多个分类器,每个分类器用于确定所述第一对象与所述第二对象是否属于一个或多个特定类型的对象关系。例如,关系召回模型可以包括一个工作关系类型的第一分类器、一个亲属关系类型的第二分类器,该第一分类器用于确定所述第一对象与所述第二对象是否属于工作关系类型的对象关系,该第二分类器用于确定所述第一对象与所述第二对象是否属于亲属关系类型的对象关系;再例如,关系召回模型可以包括同时对应于工作关系类型和亲属关系类型的分类器,该分类器用于确定所述第一对象与所述第二对象是否属于工作关系类型或亲属关系类型的对象关系。

[0096] 在一实施例中,可以根据所述特征信息确定所述第一对象与所述第二对象之间的关系重要程度;然后,当所述关系重要程度大于预设重要程度(即重要关系)时,根据所述特

征信息确定所述第一对象与所述第二对象之间的对象关系。

[0097] 在一实施例中,当所述关系重要程度不大于预设重要程度(即非重要关系)时,可以避免根据所述特征信息确定所述第一对象与所述第二对象之间的对象关系,可以节省对非重要关系的特征信息进行识别的处理量。

[0098] 在一实施例中,可以通过重要性识别模型确定所述第一对象与所述第二对象之间的关系重要程度;其中,所述重要性识别模型由基于历史对象通讯行为数据生成的第二训练样本进行训练得到。

[0099] 在一实施例中,所述第二训练样本包含标注的对象关系信息,以通过监督算法训练得到所述重要性识别模型。例如,监督算法包括梯度提升决策树(Gradient Boosting Decision Tree,GBDT)算法或相关衍生算法(如eXtreme Gradient Boosting,简称为Xgboost)。

[0100] 在一实施例中,所述第二训练样本中的部分数据包含标注的对象关系信息,以通过半监督算法训练得到所述重要性识别模型。

[0101] 由以上技术方案可见,本申请通过分析用户或其他类型的对象之间的通讯行为数据,可以根据该通讯行为数据包含的特征信息,准确识别出相应的用户关系或对象关系,而无需用户之间或对象之间相互标注的信息(如通讯录等处标注的信息)。

[0102] 为了便于理解,下面以基于通话行为数据来识别用户关系为例,对本申请的技术方案进行说明。结合图1,假定手机13内放置有用户A对应的SIM(Subscriber Identification Module,客户识别模块)卡、手机15内放置有用户B对应的SIM卡,使得用户A与用户B可以使用手机13、手机15实现相互通话。在本实施例中,服务器11可以根据用户A与用户B之间的通话行为数据,确定出用户A与用户B之间的用户关系。

[0103] 相应地,图4是本申请一示例性实施例提供的一种基于通话行为数据识别用户关系的示意图。如图4所示,识别用户关系的过程可以包括两个阶段:准备阶段和处理阶段;其中,准备阶段通过:①准备数据、②生成训练样本、③模型训练,可以得到基于通话行为数据来识别用户关系的模型,使得在处理阶段可以基于该模型对用户关系进行识别。

[0104] 在①准备数据的过程中,可以采用历史用户通话行为数据作为训练样本的数据来源,以用于生成训练样本。在一实施例中,历史用户通话行为数据可以包括若干用户之间的历史通话行为的描述信息,即来源于真实的历史通话行为,可以确保训练得到的模型的识别准确度。当然,在一些实施例中,也可以采用模拟的通话行为数据或者其他类型的通话行为数据,本申请并不对此进行限制。

[0105] 历史用户通话行为数据可以存在一个或多个数据来源,比如该数据来源可以包括:至少一台电子设备上的历史通话记录数据(可由该电子设备上传,以提供至实施本实施例的服务器等电子设备),或者运营商处的历史通话记录数据(可以通过与运营商之间的合作关系,从运营商处调取该历史通话记录数据)等。

[0106] 在②生成训练样本的过程中,可以包括下述操作:

[0107] 1)对历史用户通话行为数据进行数据标注。根据关系召回模型所需要识别的用户关系类型,对历史用户通话行为数据对应的用户关系进行标注。例如,如果希望关系召回模型对亲属类型的用户关系进行识别,可以对历史用户通话行为数据中具有亲属关系的通话行为数据进行标注,比如“爸”、“妈”、“儿子”、“女儿”、“舅舅”、“外公”等;而如果希望关系召

回模型对同学类型的用户关系进行识别,可以对历史用户通话行为数据中具有同学关系的通话行为数据进行标注,比如“同桌”、“班长”、“学习委员”等。关系召回模型可以对一种或多种类型的用户关系进行召回,通过为历史用户通话行为数据添加相应类型的标注数据,可使训练得到的关系召回模型对相应类型的用户关系进行识别。

[0108] 当模型训练算法为半监督算法时,可以仅对历史用户通话行为数据中的部分数据进行标注,以基于该历史用户通话行为数据进行模型训练;当模型训练算法为监督算法时,可以对历史用户通话行为数据进行完全标注,以基于该历史用户通话行为数据进行模型训练,或者可以选用历史用户通话行为数据中已标注的部分数据,以基于该部分数据进行模型训练。

[0109] 2) 从历史用户通话行为数据中抽取特征

[0110] 针对历史用户通话数据中的若干组用户(通话双方),分别针对每组用户对应的通话数据,抽取该通话数据中包含的预设维度上的特征。

[0111] 在不同用户关系的场景下,用户之间的通话行为数据会随之表现出若干维度上的差异。比如同事之间的通话一般发生在工作日的工作时间段内,家人之间的通话一般发生在晚上或者周末,因而可以从通话发起时刻的维度上对“同事”与“亲属”类型的用户关系进行区分。当然,通话行为数据还会在很多其他维度上表现出差异,因而通过对历史用户通话行为数据的细致分析,并尽可能地从中抽取更多维度的特征,有助于更加准确地区分不同用户关系,以及更加全面地描述用户之间的用户关系。

[0112] 在一实施例中,可以从通话发起时刻、通话持续时长、通话角色(主叫方或被叫方)等多个维度上,抽取历史用户通话行为数据中包含的特征。

[0113] 在一实施例中,通过对通话发起时刻、通话持续时长、通话角色等维度进行延伸、扩展、组合等,可以得到通话连续性及间断性、重复模式、通话时间、频率、时长、通话阵发特征、相互通讯占比等更多维度。

[0114] 在一实施例中,通过对通话连续性及间断性、重复模式、通话时间、频率、时长、通话阵发特征、相互通讯占比等维度做更小细粒度的划分,可以得到更多维度。比如,“通话连续性及间断性”可以包括:总通话间隔天数、间隔天数占总通话间隔天数的比例、最长的间隔天数等;“重复模式”可以包括:发生通话的月份数、发生通话的周数、平均周通话数、通话周数占总周数比例等;“通话时间”可以包括:工作日早/中/下午/晚上的通话数、工作日通话频率、有通话的不同时间段数等;“频率”可以包括:总通话数,周末通话数、每周呼出/呼入通话数、通话数的均值/方差/中位数;“时长”可以包括:总通话时长、周末通话时长、工作时间通话时长、平均通话时长、均值/最大值/方差/中位数等;“通话阵发特征”可以包括:发生阵发的次数、平均阵发长度、最大阵发长度等;“相互通讯占比”可以包括:呼入与呼出比值、与对应用户呼出数占总呼出数的比值、与对应用户通话占总通话比值等。

[0115] 在一实施例中,可以将按照上述维度抽取的特征直接作为训练样本。比如,假定总共设定了50个维度,那么对于历史用户通话行为数据中的每组用户而言,可以抽取相应的50维特征,以用于表征该组用户之间的用户关系,该50维特征可以作为训练样本。

[0116] 在另一实施例中,可以对上述维度抽取的特征进行压缩变换,以得到相应的压缩特征。比如对于上述50维的原始特征,可以通过压缩变换生成为10维的压缩特征。

[0117] 在一实施例中,压缩特征相对于原始特征而言,能够更好地刻画历史用户通话行

为数据表现出的用户关系。比如,当原始特征为50维特征时,压缩变换可以对50个维度之间的重新描述信息进行“去重”,以简化特征、降低模型训练过程的处理量。再比如,当原始特征为50维特征时,压缩变换可以对50个维度之间进行“组合”,从而能够表达出原始特征无法表达的信息,能够更好地表征用户关系。

[0118] 在一实施例中,可以通过自编码器对原始特征进行压缩变换为相应的压缩特征。例如,可以采用变分自编码器实施该压缩变换处理。当然,不论是自编码器或是其他方式,本申请并不对此进行限制。

[0119] 可见,通过对特征维度的合理定义,可以尽可能全面地从历史用户通话行为数据中提取特征,以尽可能丰富、细致地描述相应的用户关系。同时,通过对原始特征进行压缩变换,实现了以数据驱动的特征自动重整,有助于获得更为有效的压缩特征。

[0120] 3) 确定训练样本

[0121] 对于关系召回模型

[0122] 在一实施例中,可以采用半监督算法来训练关系召回模型,因而训练样本1可以仅包含部分标注有用户关系的训练样本,而无需对所有训练样本均进行标注,可以极大地简化对历史用户通话行为数据的数据标注操作。

[0123] 用户关系存在很多类型。如果希望训练出的关系召回模型可以识别某一类型的用户关系,那么可以确保训练样本1包含至少一部分属于该类型的训练样本,且这些训练样本的至少一部分被标注了相应的用户关系。而如果希望训练出的关系召回模型可以识别多种类型的用户关系,那么可以确保训练样本1包含分别对应于这些类型的训练样本,且对应于每种类型的训练样本中的至少一部分被标注相应的用户关系。

[0124] 在一实施例中,根据训练样本1中包含的对应于每种用户关系的训练样本,可以分别通过半监督算法训练得到相应的分类器,并由这些分类器组成上述的关系召回模型;例如,关系召回模型可以分别包括亲属关系分类器、同事关系分类器等,以分别用于识别亲属、同事等类型的用户关系。

[0125] 在一实施例中,根据训练样本1中包含的对应于多种用户关系的训练样本,可以通过半监督算法训练得到相应的一个或多个分类器,且每个分类器可以用于识别一种或多种用户关系,并由该一个或多个分类器组成上述的关系召回模型。

[0126] 可见,关系召回模型易于扩展,以用于识别更多类型的用户关系。当存在新类型的用户关系的识别需求时,通过提供标注有该类型的用户关系的训练样本,可以基于上述的半监督模型训练得到针对该新类型的用户关系的分类器,并通过将该分类器添加至原有的关系召回模型中,可以实现对该关系召回模型的扩展。

[0127] 对于重要性识别模型

[0128] 在一实施例中,可以通过训练样本2训练得到重要性识别模型,该重要性识别模型用于识别用户关系的重要程度,或者说识别用户关系是否为重要关系。

[0129] 在一实施例中,关系召回模型所需识别的用户关系均属于重要的用户关系,因而可以将关系召回模型所需识别的用户关系确定为重要性识别模型需要识别的重要关系。

[0130] 在一实施例中,重要性识别模型所需识别的重要关系可以包括预定义的任意用户关系,而与关系召回模型所需识别的用户关系无关。

[0131] 在一实施例中,当关系召回模型采用半监督算法进行训练、重要性识别模型采用

监督算法进行训练时,假定关系召回模型所需识别的用户关系均为重要关系,那么可以提取训练样本1中所有的标注数据(即标注有相应的用户关系的训练样本),以作为训练样本2,使得重要性识别模型可以识别出这些用户关系。

[0132] 在③模型训练的过程中:

[0133] 对于关系召回模型

[0134] 在一实施例中,相关技术中的任意半监督算法均可以应用于对训练样本1进行训练,以得到关系召回模型。例如,该半监督算法可以包括Diederik P.Kingma在论文《Semi-supervised Learning with Deep Generative Models(基于深度生成模型的半监督学习)》中提及的M2半监督算法;那么,该M2半监督算法通过采用深度生成模型对训练样本1进行训练,相比于未采用深度生成模型的半监督算法,该M2半监督算法训练得到的分类器具有更优的分类效果。

[0135] 对于重要性识别模型

[0136] 在一实施例中,可以采用监督算法对训练样本2进行训练,以得到重要性识别模型。例如,该监督算法可以包括GBDT算法;或者,监督算法可以包括GBDT算法的衍生算法,比如Xgboost算法等,该Xgboost算法作为GBDT算法的一种C++实现,其表现优于GBDT算法。当然,还可以采用其他类型的监督算法,本申请并不对此进行限制。

[0137] 需要指出的是:在上述实施例中,当采用M2等半监督算法训练关系召回模型、采用Xgboost等监督算法训练重要性识别模型时,需要向M2等半监督算法提供至少部分标注的训练样本1、向Xgboost等监督模型提供完全标注的训练样本2;但是,在一些实施例中,也可以采用Xgboost等监督算法对关系召回模型进行训练、采用M2等半监督算法对关系召回模型进行训练,那么相应的模型训练所需的训练样本也会随之发生变化。

[0138] 基于上述的准备阶段的处理操作,可以在图4所示的处理阶段中,通过关系召回模型、重要性识别模型对用户X的历史通话记录数据进行分析,以识别出与该用户X存在某种用户关系的其他用户。相应地,图5是本申请一示例性实施例提供的一种识别用户关系的示意图,如图5所示,假定希望识别出用户X的家人和同事,可以采用下述处理过程:

[0139] 获取用户X的历史通话记录数据,即该用户X参与过的历史通话行为对应的记录数据。在一实施例中,该历史通话记录数据可以由用户X通过其手机等电子设备上传;在另一实施例中,该历史通话记录数据可以从运营商处取得。

[0140] 从用户X的历史通话记录数据中抽取待识别特征。在一实施例中,可以参考图4中对于历史用户通话行为数据的特征抽取方式,从用户X的历史通话记录数据中抽取待识别特征。在一实施例中,可以对抽取的待识别特征进行压缩变换,以得到相应的待识别压缩特征,该压缩变换过程可以参考图4中的压缩变换过程,此处不再赘述。通过在图5所示的实施例中,采用与图4所示实施例相同的特征抽取方式,使得上述的待识别特征或待识别压缩特征贴适于图4所示的训练样本1和训练样本2的数据结构,以便于重要性识别模型和关系召回模型可以更好地予以识别。

[0141] 重要性识别模型与关系召回模型之间可以采用图5所示的串联关系。

[0142] 在一实施例中,待识别压缩特征被首先输入重要性识别模型中,由重要性识别模型识别出重要关系号码和非重要关系号码,以及每一重要关系号码对应的用户与用户X之间属于重要关系的概率 $P(\text{sig})$ 。然后,重要关系号码被输入关系召回模型,由关系召回模型

中对应于各个类型的用户关系的分类器,分别确定重要关系号码对应的用户与用户X之间是否具有相应类型的用户关系,这样关系召回模型就无需对非重要关系号码进行识别,可以降低相应的数据处理量。

[0143] 在一实施例中,关系召回模型中可以包括家人关系分类器、同事关系分类器等,以分别用于识别家人关系、同事关系等用户关系。其中,这些分类器之间可以采用图5所示的串联关系。比如,重要性识别模型输出的重要关系号码可以被首先输入家人关系分类器,并由该家人关系分类器确定每一重要关系号码对应的用户与用户X之间是否具有家人关系,并输出家人号码和非家人关系号码,以及每一家人号码对应的用户与该用户X之间属于家人关系的预测概率 $P(\text{fam})$ 。然后,非家人关系号码被输入同事关系分类器,由该同事关系分类器确定每一非家人关系号码对应的用户与用户X之间是否具有同事关系,并输出同事号码和非同事关系号码,以及每一同事号码对应的用户与该用户X之间属于同事关系的预测概率 $P(\text{work})$ 。

[0144] 根据重要性识别模型输出的重要关系号码及其预测概率 $P(\text{sig})$,以及关系召回模型输出的家人号码及其预测概率 $P(\text{fam})$ 、同事号码及其预测概率 $P(\text{work})$,可以对用户X的家人、同事等进行概率排序,以确定出诸如“最重要的3位家人”、“最重要的2位同事”等。

[0145] 例如,假定第 i 个家人号码对应于预测概率 $P_i(\text{sig})$ 与预测概率 $P_i(\text{fam})$,则可以计算相应的概率得分为: $S_i = P_i(\text{sig}) \times P_i(\text{fam})$ 。那么,通过对所有家人号码对应的概率得分进行计算和排序,可以将概率得分最高的3个家人号码确定为“最重要的3位家人”使用的号码。

[0146] 类似地,假定第 j 个同事号码对应于预测概率 $P_j(\text{sig})$ 与预测概率 $P_j(\text{work})$,则可以计算相应的概率得分为: $S_j = P_j(\text{sig}) \times P_j(\text{work})$ 。那么,通过对所有同事号码对应的概率得分进行计算和排序,可以将概率得分最高的2个同事号码确定为“最重要的2位同事”使用的号码。

[0147] 图5所示的实施例可以存在多种变化,例如:

[0148] 在一实施例中,重要性识别模型与关系召回模型之间可以采用区别于图5所示实施例的串联顺序,比如关系召回模型可以位于重要性识别模型之前。

[0149] 在一实施例中,重要性识别模型与关系召回模型之间可以采用并联关系。

[0150] 在一实施例中,关系召回模型中的多个分类器之间可以采用区别于图5所示实施例的串联顺序,比如同事关系分类器可以位于家人关系分类器之前。

[0151] 在一实施例中,关系召回模型中的多个分类器之间可以采用并联关系。

[0152] 综上所述,本申请提出了基于通讯行为数据对用户关系进行识别的方案,无需获取用户的通讯录数据或该通讯录数据中标注的用户关系信息,而可以根据不同用户关系在通讯行为中表征出的特征,对用户关系进行准确识别。

[0153] 图6示出了根据本申请的一示例性实施例的电子设备的示意结构图。请参考图6,在硬件层面,该电子设备包括处理器602、内部总线604、网络接口606、内存608以及非易失性存储器610,当然还可能包括其他业务所需要的硬件。处理器602从非易失性存储器610中读取对应的计算机程序到内存608中然后运行,在逻辑层面上形成用户关系的识别装置。当然,除了软件实现方式之外,本申请并不排除其他实现方式,比如逻辑器件抑或软硬件结合的方式等等,也就是说以下处理流程的执行主体并不限定于各个逻辑单元,也可以是硬件

或逻辑器件。

[0154] 请参考图7,在软件实施方式中,该用户关系的识别装置可以包括:

[0155] 获取单元71,获取第一用户与第二用户之间的通讯行为数据;

[0156] 提取单元72,提取所述通讯行为数据在预设维度上的特征信息;

[0157] 确定单元73,根据所述特征信息确定所述第一用户与所述第二用户之间的用户关系。

[0158] 图8示出了根据本申请的一示例性实施例的电子设备的示意结构图。请参考图8,在硬件层面,该电子设备包括处理器802、内部总线804、网络接口806、内存808以及非易失性存储器810,当然还可能包括其他业务所需要的硬件。处理器802从非易失性存储器810中读取对应的计算机程序到内存808中然后运行,在逻辑层面上形成对象关系的识别装置。当然,除了软件实现方式之外,本申请并不排除其他实现方式,比如逻辑器件抑或软硬件结合的方式等等,也就是说以下处理流程的执行主体并不限于各个逻辑单元,也可以是硬件或逻辑器件。

[0159] 请参考图9,在软件实施方式中,该对象关系的识别装置可以包括:

[0160] 获取单元91,获取第一对象与第二对象之间的通讯行为数据;

[0161] 提取单元92,提取所述通讯行为数据在预设维度上的特征信息;

[0162] 确定单元93,根据所述特征信息确定所述第一对象与所述第二对象之间的对象关系。

[0163] 可选的,所述通讯行为数据包括以下至少之一:

[0164] 语音通话记录数据、视频通话记录数据、通讯消息收发记录数据。

[0165] 可选的,所述通讯行为数据包括所述第一对象与所述第二对象在预设时间段内的通讯行为的描述信息。

[0166] 可选的,所述确定单元93具体用于:

[0167] 通过关系召回模型确定所述第一对象与所述第二对象之间的对象关系;其中,所述关系召回模型由基于历史对象通讯行为数据生成的第一训练样本进行训练得到。

[0168] 可选的,所述第一训练样本包括所述历史对象通讯行为数据在所述预设维度上的特征信息。

[0169] 可选的,所述第一训练样本中的特征信息包括:

[0170] 所述历史对象通讯行为数据在所述预设维度上的原始特征;

[0171] 或者,由所述原始特征经过压缩转换得到的压缩特征。

[0172] 可选的,所述原始特征通过变分自编码器被压缩转换为相应的压缩特征。

[0173] 可选的,所述预设维度与下述信息中至少之一相关:

[0174] 通讯发起时刻、通讯时长、通讯角色、通讯类型。

[0175] 可选的,所述关系召回模型包括分类器,所述分类器用于确定所述第一对象与所述第二对象是否属于特定类型的对象关系;其中,所述第一训练样本中的部分数据包含标注的所述特定类型的对象关系信息,以通过半监督算法训练得到所述分类器。

[0176] 可选的,所述半监督算法基于深度生成模型对所述第一训练样本进行训练,以得到所述分类器。

[0177] 可选的,所述关系召回模型包括对应于不同类型的多个分类器。

[0178] 可选的,所述确定单元93具体用于:

[0179] 根据所述特征信息确定所述第一对象与所述第二对象之间的关系重要程度;

[0180] 当所述关系重要程度大于预设重要程度时,根据所述特征信息确定所述第一对象与所述第二对象之间的对象关系。

[0181] 可选的,所述确定单元93通过下述方式,实现根据所述特征信息确定所述第一对象与所述第二对象之间的关系重要程度:

[0182] 通过重要性识别模型确定所述第一对象与所述第二对象之间的关系重要程度;其中,所述重要性识别模型由基于历史对象通讯行为数据生成的第二训练样本进行训练得到。

[0183] 可选的,所述第二训练样本包含标注的对象关系信息,以通过监督算法训练得到所述重要性识别模型。

[0184] 可选的,所述监督算法包括梯度提升决策树算法或相关衍生算法。

[0185] 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,计算机的具体形式可以是个人计算机、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、电子邮件收发设备、游戏控制台、平板计算机、可穿戴设备或者这些设备中的任意几种设备的组合。

[0186] 在一个典型的配置中,计算机包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。

[0187] 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。

[0188] 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。

[0189] 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。

[0190] 这里将详细地对示例性实施例进行说明,其示例表示在附图中。下面的描述涉及附图时,除非另有表示,不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本申请相一致的所有实施方式。相反,它们仅是与如所附权利要求书中所详述的、本申请的一些方面相一致的装置和方法的例子。

[0191] 在本申请使用的术语是仅仅出于描述特定实施例的目的,而非旨在限制本申请。在本申请和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式,除非上下文清楚地表示其他含义。还应当理解,本文中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

[0192] 应当理解,尽管在本申请可能采用术语第一、第二、第三等来描述各种信息,但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如,在不脱离本申请范围的情况下,第一信息也可以被称为第二信息,类似地,第二信息也可以被称为第一信息。取决于语境,如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

[0193] 以上所述仅为本申请的较佳实施例而已,并不用以限制本申请,凡在本申请的精神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本申请保护的范围之内。

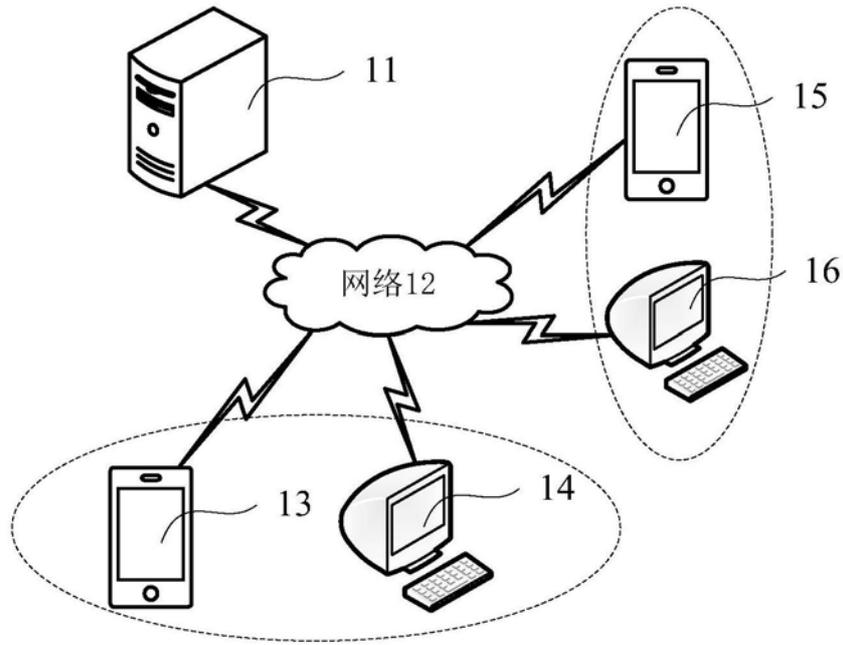


图1

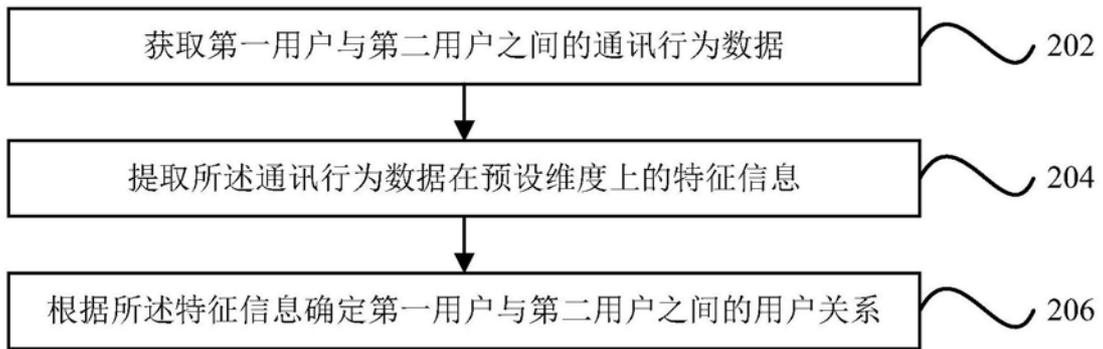


图2

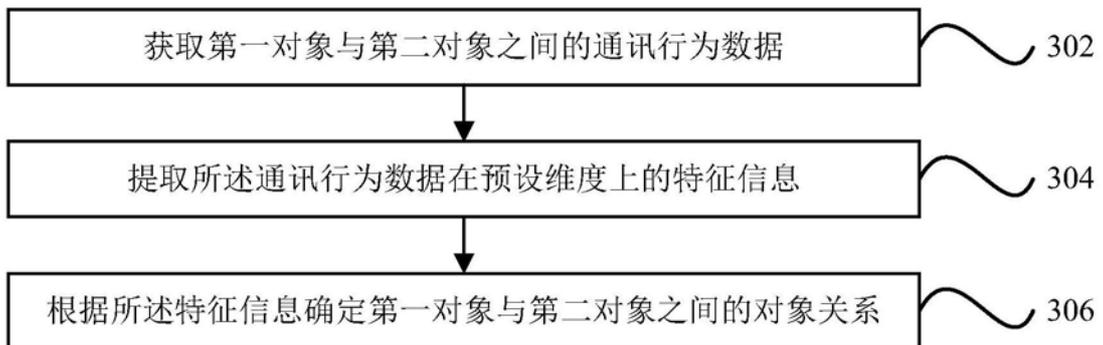


图3

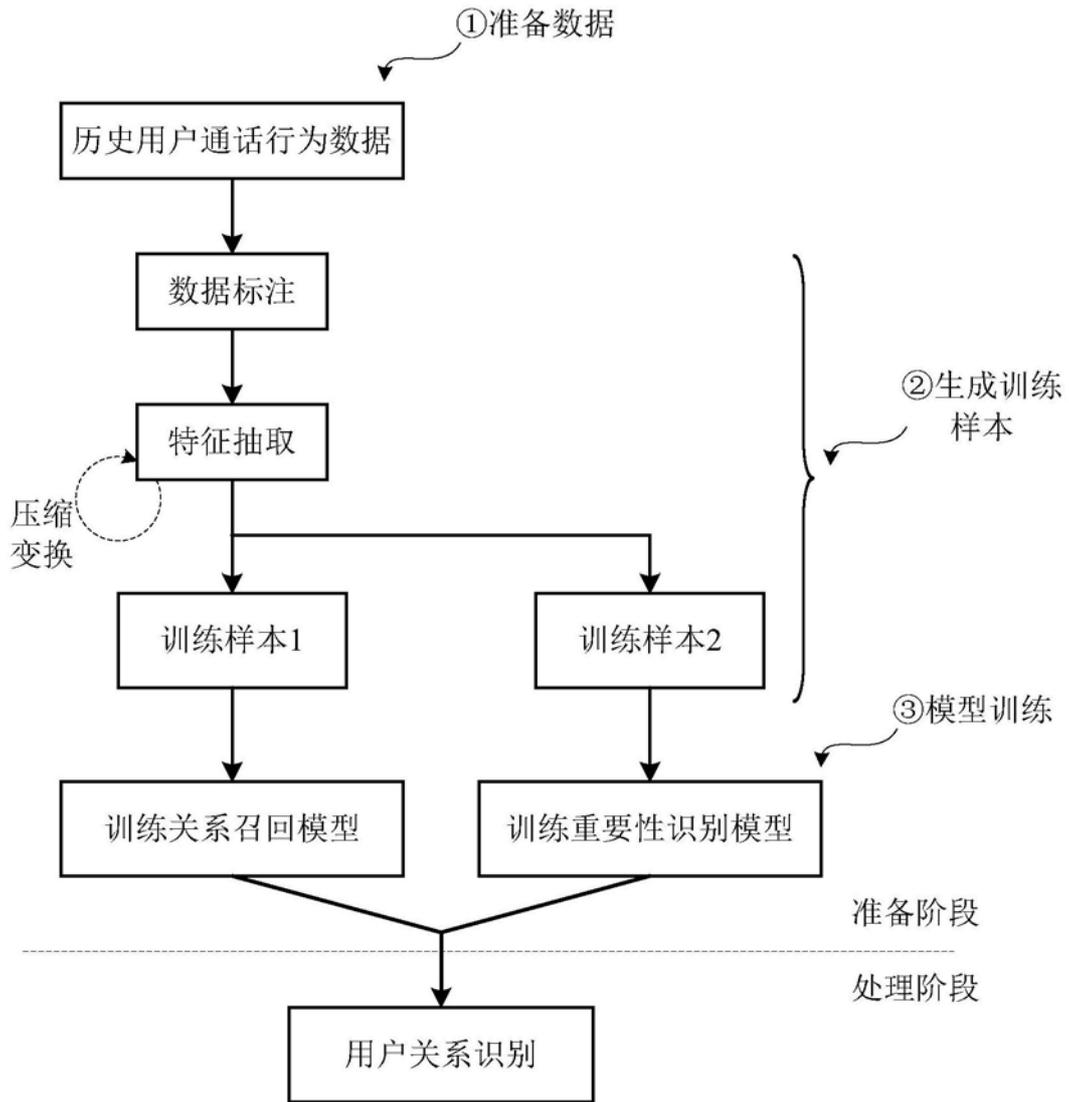


图4

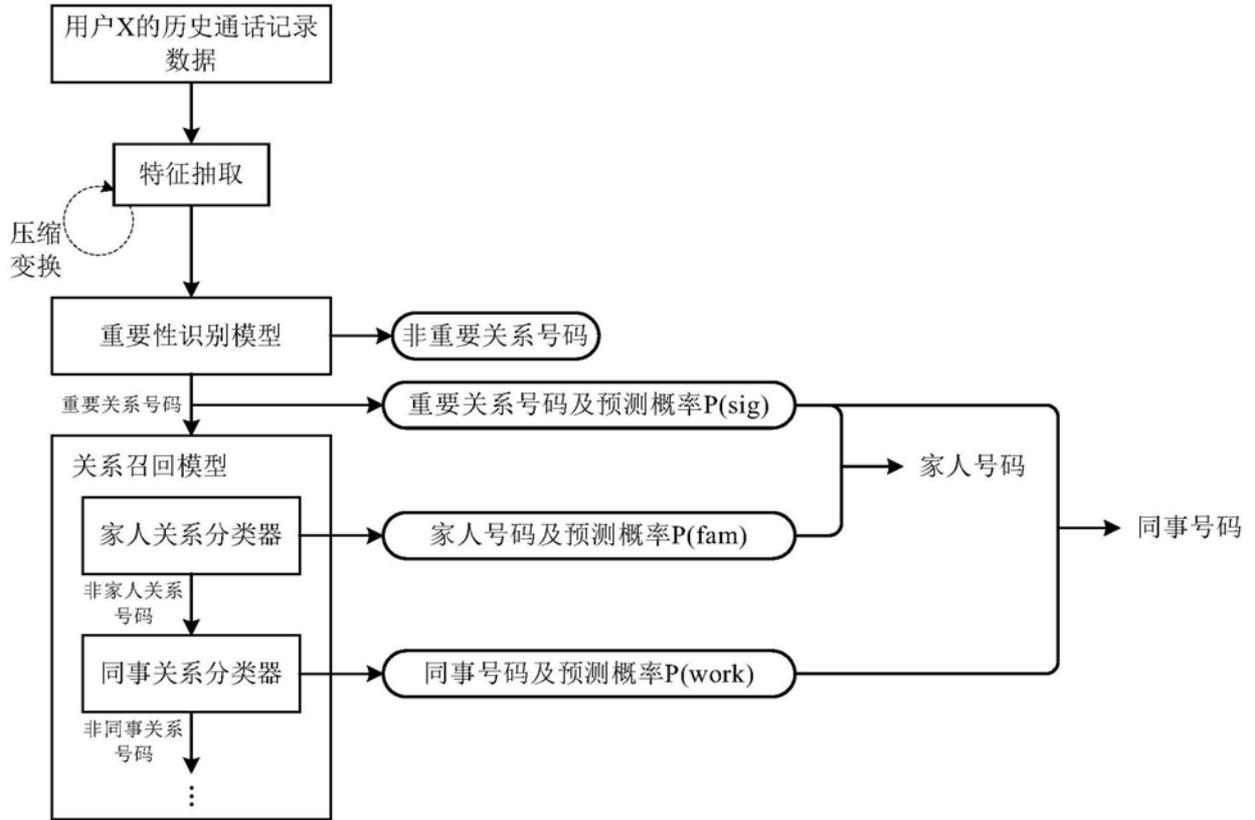


图5

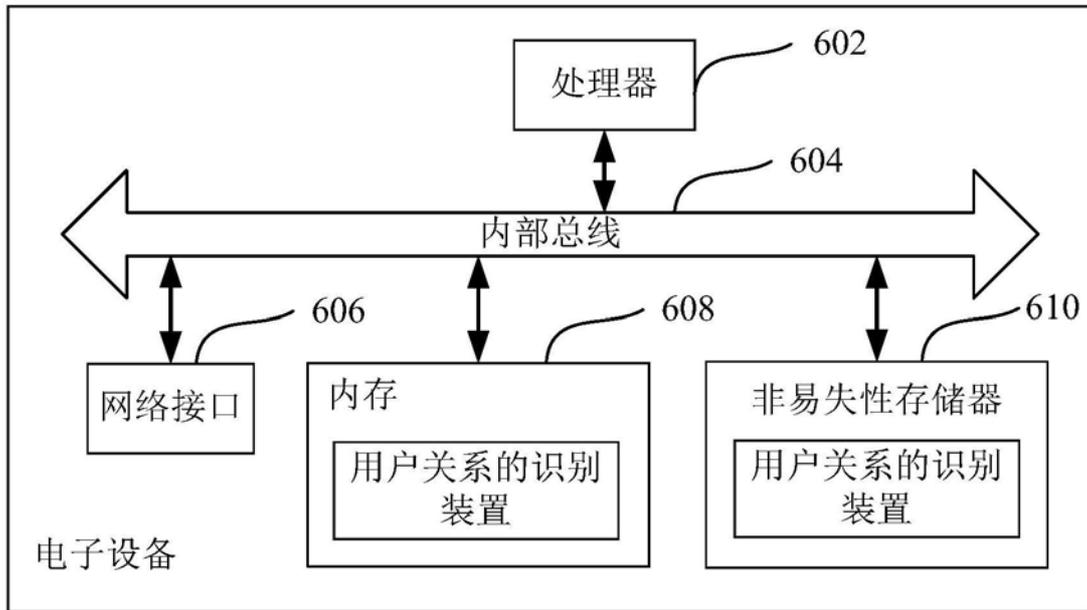


图6



图7

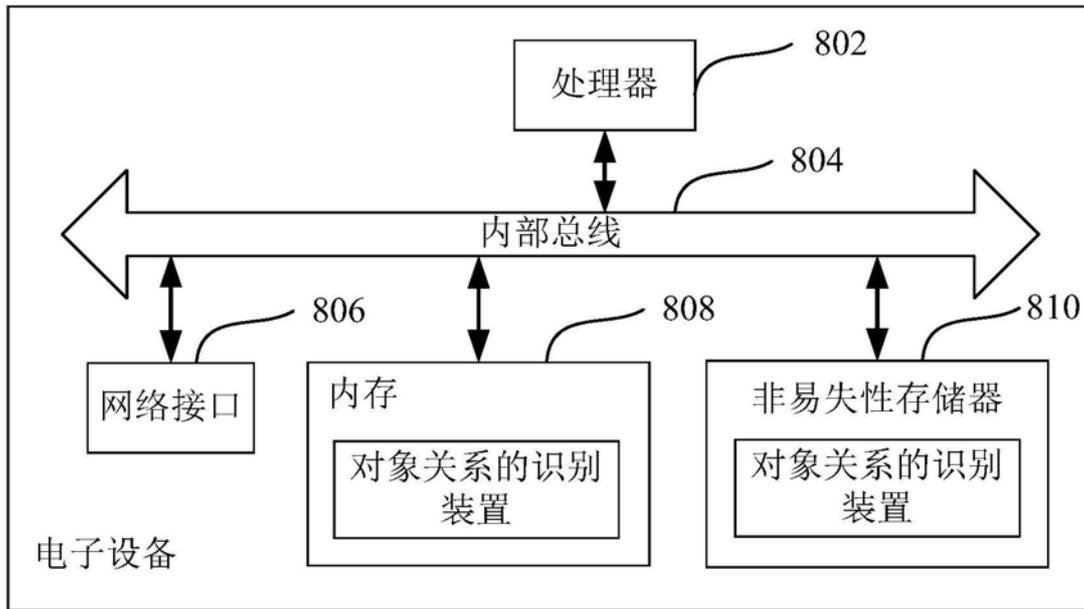


图8

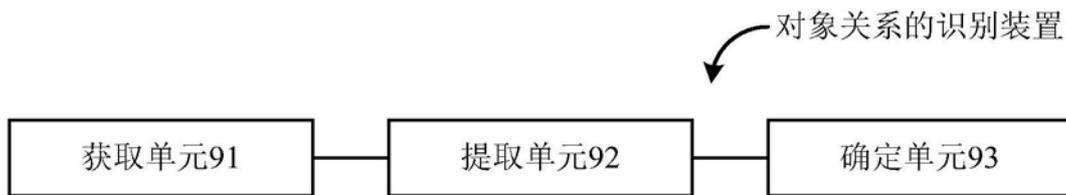


图9