



(12) 发明专利申请

(10) 申请公布号 CN 103733190 A

(43) 申请公布日 2014. 04. 16

(21) 申请号 201280038701. 3

(74) 专利代理机构 北京市中咨律师事务所
11247

(22) 申请日 2012. 07. 30

代理人 于静 张亚非

(30) 优先权数据

13/204, 831 2011. 08. 08 US

(51) Int. Cl.

G06F 17/00(2006. 01)

(85) PCT国际申请进入国家阶段日

2014. 02. 07

(86) PCT国际申请的申请数据

PCT/US2012/048771 2012. 07. 30

(87) PCT国际申请的公布数据

W02013/022631 EN 2013. 02. 14

(71) 申请人 国际商业机器公司

地址 美国纽约

(72) 发明人 D·卡涅夫斯基 J·R·克泽罗斯基

C·A·皮茨克维尔 T·N·赛纳斯

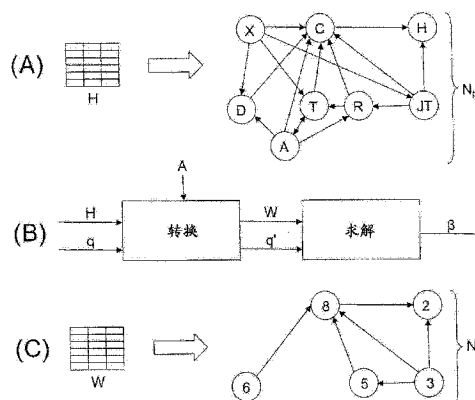
权利要求书3页 说明书17页 附图4页

(54) 发明名称

在保留网络属性的同时保护网络实体数据

(57) 摘要

在一个示例性实施例中, 提供一种包含用于执行操作的指令程序的存储介质, 所述操作包括: 针对主网络中的主节点的主属性存储原始信息; 响应于接收查询, 使用密钥并基于所述查询, 将所述原始信息转换成转换后的信息, 所述查询涉及被查询的属性, 所述转换后的信息是所述被查询的属性的转换后的数据, 所述转换后的信息表示具有代理节点的代理网络, 所述代理节点对应于所述主节点的一部分, 所述转换后的信息使能执行操作, 而无需完整的原始信息的特定知识并且无需泄露完整的原始信息, 所述转换后的信息还使得具有所述密钥的人员能够将所述操作的输出与所述原始信息相关; 以及生成将转换后的查询与所述转换后的信息相关的解, 所述转换后的查询是通过使用所述密钥获得的所述查询的转换后的表示。



1. 一种有形地包含可由机器执行以便执行操作的指令程序的计算机可读存储介质(108),所述操作包括:

针对主网络中的至少一个主节点的至少一个主属性存储(501)原始信息,其中所述主网络包括互连的多个主节点,其中所述主网络中的所述多个主节点之间的连接基于以下项中的至少一个:所述至少一个主属性,以及所述主节点之间的一种或多种类型的关系和相互依赖性;

响应于接收对所存储的原始信息的查询,使用至少一个密钥并基于所接收的查询,将所存储的原始信息转换(502)成转换后的信息,其中所接收的查询涉及所述至少一个主属性的至少一个被查询的属性,其中所述转换后的信息包括所述至少一个主节点的所述至少一个被查询的属性的转换后的数据,其中所述转换后的信息表示包括多个代理节点的代理网络,所述多个代理节点对应于来自所述主网络的所述多个主节点的至少一部分,其中所述转换后的信息并不对应于完整的所存储的原始信息,其中所述转换后的信息被配置为使能针对所述转换后的信息执行至少一个操作,而无需完整的所存储的原始信息的特定知识并且无需泄露完整的所存储的原始信息,其中所述转换后的信息还被配置为使得拥有所述至少一个密钥的人员能够将所述至少一个操作的输出与所存储的原始信息相关;以及

生成(503)将转换后的查询与所述转换后的信息相关的解,其中所述转换后的查询包括通过使用所述至少一个密钥获得的所接收的查询的转换后的表示。

2. 根据权利要求1的计算机可读存储介质(108),其中所述解包括稀疏表示。

3. 根据权利要求1或2中的任一权利要求的计算机可读存储介质(108),其中所述网络包括社交网络或在线社交网络。

4. 根据权利要求1至3中的任一权利要求的计算机可读存储介质(108),其中由无权访问所述至少一个密钥的第三方执行所述至少一个操作。

5. 根据权利要求1至4中的任一权利要求的计算机可读存储介质(108),其中由无权访问所述至少一个密钥的不可信第三方执行所述至少一个操作。

6. 根据权利要求1至5中的任一权利要求的计算机可读存储介质(108),其中所述代理网络针对所述代理节点之间的连接使用至少一个不同于所述主网络的基础。

7. 根据权利要求1至6中的任一权利要求的计算机可读存储介质(108),其中每个主节点对应于多个实体中的一个实体。

8. 根据权利要求1至7中的任一权利要求的计算机可读存储介质(108),其中每个主节点对应于多个实体中的一个不同实体。

9. 根据权利要求1至7中的任一权利要求的计算机可读存储介质(108),其中每个主节点对应于多个实体中的一个实体,并且其中所述多个实体中的每个实体包括以下之一:一个或多个人员、一个或多个组织、一个或多个公司、一个或多个医疗保健专业人员、一个或多个营销人员、一个或多个人口统计学家、一个或多个求职者,或者一个或多个招聘者。

10. 根据权利要求1至9中的任一权利要求的计算机可读存储介质(108),其中由监管所述主网络的网络运营商存储所述原始信息和至少一个密钥。

11. 根据权利要求1至10中的任一权利要求的计算机可读存储介质(108),其中仅所述网络运营商有权访问所述至少一个密钥。

12. 根据权利要求1至11中的任一权利要求的计算机可读存储介质(108),其中所述

转换后的信息还被配置为使能针对所述转换后的信息执行所述至少一个操作,而无需访问完整的所存储的原始信息。

13. 根据权利要求 1 至 12 中的任一权利要求的计算机可读存储介质(108),其中所述转换运行以便删除、替换或掩蔽所述至少一个主属性中与所述至少一个被查询的属性不相关的其它属性。

14. 一种方法(图 5),包括:

针对主网络中的至少一个主节点的至少一个主属性存储(501)原始信息,其中所述主网络包括互连的多个主节点,其中所述主网络中的所述多个主节点之间的连接基于以下项中的至少一个:所述至少一个主属性,以及所述主节点之间的一种或多种类型的关系和相互依赖性;

响应于接收对所存储的原始信息的查询,使用至少一个密钥并基于所接收的查询,将所存储的原始信息转换(502)成转换后的信息,其中所接收的查询涉及所述至少一个主属性的至少一个被查询的属性,其中所述转换后的信息包括所述至少一个主节点的所述至少一个被查询的属性的转换后的数据,其中所述转换后的信息表示包括多个代理节点的代理网络,所述多个代理节点对应于来自所述主网络的所述多个主节点的至少一部分,其中所述转换后的信息并不对应于完整的所存储的原始信息,其中所述转换后的信息被配置为使能针对所述转换后的信息执行至少一个操作,而无需完整的所存储的原始信息的特定知识并且无需泄露完整的所存储的原始信息,其中所述转换后的信息还被配置为使得拥有所述至少一个密钥的人员能够将所述至少一个操作的输出与所存储的原始信息相关;以及

生成(503)将转换后的查询与所述转换后的信息相关的解,其中所述转换后的查询包括通过使用所述至少一个密钥获得的所接收的查询的转换后的表示。

15. 根据权利要求 14 的方法,其中所述解包括稀疏表示。

16. 根据权利要求 14 或 15 中的任一权利要求的方法,其中所述网络包括社交网络或在线社交网络。

17. 根据权利要求 14 至 16 中的任一权利要求的方法,其中每个主节点对应于多个实体中的一个实体,并且其中所述多个实体中的每个实体包括以下之一:一个或多个人员、一个或多个组织、一个或多个公司、一个或多个医疗保健专业人员、一个或多个营销人员、一个或多个人口统计学家、一个或多个求职者,或者一个或多个招聘者。

18. 根据权利要求 14 至 17 中的任一权利要求的方法,其中所述转换运行以便删除、替换或掩蔽所述至少一个主属性中与所述至少一个被查询的属性不相关的其它属性。

19. 一种装置(100),包括:

至少一个存储器(106),其被配置为针对主网络中的至少一个主节点的至少一个主属性存储原始信息,其中所述主网络包括互连的多个主节点,其中所述主网络中的所述多个主节点之间的连接基于以下项中的至少一个:所述至少一个主属性,以及所述主节点之间的一种或多种类型的关系和相互依赖性;以及

至少一个处理器(104),其被配置为响应于接收对所存储的原始信息的查询,使用至少一个密钥并基于所接收的查询,将所存储的原始信息转换成转换后的信息,其中所接收的查询涉及所述至少一个主属性的至少一个被查询的属性,其中所述转换后的信息包括所述至少一个主节点的所述至少一个被查询的属性的转换后的数据,其中所述转换后的信息表

示包括多个代理节点的代理网络,所述多个代理节点对应于来自所述主网络的所述多个主节点的至少一部分,其中所述转换后的信息并不对应于完整的所存储的原始信息,其中所述转换后的信息被配置为使能针对所述转换后的信息执行至少一个操作,而无需完整的所存储的原始信息的特定知识并且无需泄露完整的所存储的原始信息,其中所述转换后的信息还被配置为使得拥有所述至少一个密钥的人员能够将所述至少一个操作的输出与所存储的原始信息相关,

其中所述至少一个处理器(104)还被配置为生成将转换后的查询与所述转换后的信息相关的解,其中所述转换后的查询包括通过使用所述至少一个密钥获得的所接收的查询的转换后的表示。

20. 根据权利要求 19 的装置(100),其中所述解包括稀疏表示。

21. 根据权利要求 19 或 20 中的任一权利要求的装置(100),其中所述网络包括社交网络或在线社交网络。

22. 根据权利要求 19 至 21 中的任一权利要求的装置(100),其中每个主节点对应于多个实体中的一个实体,并且其中所述多个实体中的每个实体包括以下之一:一个或多个人员、一个或多个组织、一个或多个公司、一个或多个医疗保健专业人员、一个或多个营销人员、一个或多个人口统计学家、一个或多个求职者,或者一个或多个招聘者。

23. 一种装置(100),包括:

用于针对主网络中的至少一个主节点的至少一个主属性存储原始信息的部件,其中所述主网络包括互连的多个主节点,其中所述主网络中的所述多个主节点之间的连接基于以下项中的至少一个:所述至少一个主属性,以及所述主节点之间的一种或多种类型的关系和相互依赖性;

用于响应于接收对所存储的原始信息的查询,使用至少一个密钥并基于所接收的查询,将所存储的原始信息转换成转换后的信息的部件,其中所接收的查询涉及所述至少一个主属性的至少一个被查询的属性,其中所述转换后的信息包括所述至少一个主节点的所述至少一个被查询的属性的转换后的数据,其中所述转换后的信息表示包括多个代理节点的代理网络,所述多个代理节点对应于来自所述主网络的所述多个主节点的至少一部分,其中所述转换后的信息并不对应于完整的所存储的原始信息,其中所述转换后的信息被配置为使能针对所述转换后的信息执行至少一个操作,而无需完整的所存储的原始信息的特定知识并且无需泄露完整的所存储的原始信息,其中所述转换后的信息还被配置为使得拥有所述至少一个密钥的人员能够将所述至少一个操作的输出与所存储的原始信息相关;以及

用于生成将转换后的查询与所述转换后的信息相关的解的部件,其中所述转换后的查询包括通过使用所述至少一个密钥获得的所接收的查询的转换后的表示。

24. 根据权利要求 23 的装置,其中所述解包括稀疏表示。

25. 根据权利要求 23 或 24 中的任一权利要求的装置,其中用于存储的部件包括至少一个存储器(106)或至少一个计算机可读存储介质(108),并且其中用于转换的部件和用于生成的部件包括至少一个处理器(104)或至少一个集成电路(102)。

在保留网络属性的同时保护网络实体数据

技术领域

[0001] 本发明的示例性实施例一般地涉及网络和社交网络,更具体地说,涉及在对网络和社交网络的数据实现操作和分析时保护数据隐私性。

背景技术

[0002] 本部分旨在为权利要求中详述的本发明的各种示例性实施例提供上下文或背景。此处的内容可以包括能够使用的主题,但不一定包括先前使用、描述或考虑的主题。除非另外指明,否则在此描述的内容并不被视为现有技术,并且不应因为包括在本部分中而被视为公认的现有技术。

[0003] 随着在线访问和通信的可用性增加以及成本降低,对在线社交网络的研究和分析变得更加突出。通常,社交网络可以被视为社交结构,其包括个体、个体组和/或组织(在此统称为“实体”),这些实体表示为通过一种或多种类型的关系或相互依赖性(例如,朋友、亲戚、知识、就业、爱好、兴趣)连接到彼此的“节点”。在线社交网络和结构(例如**Facebook®**)成为有价值的工具,不仅用于个人通信目的,而且还用于例如信息、娱乐和广告功能。

[0004] 随着在线社交网络使用的增加,参与者还会更易于遭受个人数据和私有信息的盗用。当维护与网络中的节点关联的个体的隐私性和匿名性时,难以访问来自社交网络的数据(例如,家庭信息、病史、与用户关联的设备网络、用于收集数据的传感器(例如摄像机)网络)。作为一个实例,可以分析社交网络(例如**Facebook®**)以便获得有关用户和参与者的数据,例如基于例如对社交网络中的个体“朋友”的分析的健康风险。此外,应该认识到,社交网络数据和相关分析可以提供有关个体的味觉、健康和可能行为的大量潜在信息。因为预测能力和信息性质,来自社交网络的数据立即有价值并且极其敏感。因此,每当收集、分析或以其它方式使用该数据时,应该维护隐私性。

发明内容

[0005] 在本发明的一个示例性实施例中,提供一种有形地包含可由机器执行以便执行操作的指令程序的计算机可读存储介质,所述操作包括:针对主网络中的至少一个主节点的至少一个主属性存储原始信息,其中所述主网络包括互连的多个主节点,其中所述主网络中的所述多个主节点之间的连接基于以下项中的至少一个:所述至少一个主属性,以及所述主节点之间的一种或多种类型的关系和相互依赖性;响应于接收对所存储的原始信息的查询,使用至少一个密钥并基于所接收的查询,将所存储的原始信息转换成转换后的信息,其中所接收的查询涉及所述至少一个主属性的至少一个被查询的属性,其中所述转换后的信息包括所述至少一个主节点的所述至少一个被查询的属性的转换后的数据,其中所述转换后的信息表示包括多个代理节点的代理网络,所述多个代理节点对应于来自所述主网络的所述多个主节点的至少一部分,其中所述转换后的信息并不对应于完整的所存储的原始信息,其中所述转换后的信息被配置为使能针对所述转换后的信息执行至少一个操作,而无需完整的所存储的原始信息的特定知识并且无需泄露完整的所存储的原始信息,其中所

述转换后的信息还被配置为使得拥有所述至少一个密钥的人员能够将所述至少一个操作的输出与所存储的原始信息相关；以及生成将转换后的查询与所述转换后的信息相关的解，其中所述转换后的查询包括通过使用所述至少一个密钥获得的所接收的查询的转换后的表示。

[0006] 在本发明的另一个示例性实施例中，一种方法包括：针对主网络中的至少一个主节点的至少一个主属性，在至少一个存储器中存储原始信息，其中所述主网络包括互连的多个主节点，其中所述主网络中的所述多个主节点之间的连接基于以下项中的至少一个：所述至少一个主属性，以及所述主节点之间的一种或多种类型的关系和相互依赖性；响应于接收对所存储的原始信息的查询，第一装置使用至少一个密钥并基于所接收的查询，将所存储的原始信息转换为转换后的信息，其中所接收的查询涉及所述至少一个主属性的至少一个被查询的属性，其中所述转换后的信息包括所述至少一个主节点的所述至少一个被查询的属性的转换后的数据，其中所述转换后的信息表示包括多个代理节点的代理网络，所述多个代理节点对应于来自所述主网络的所述多个主节点的至少一部分，其中所述转换后的信息并不对应于完整的所存储的原始信息，其中所述转换后的信息被配置为使能针对所述转换后的信息执行至少一个操作，而无需完整的所存储的原始信息的特定知识并且无需泄露完整的所存储的原始信息，其中所述转换后的信息还被配置为使得拥有所述至少一个密钥的人员能够将所述至少一个操作的输出与所存储的原始信息相关；以及第二装置生成将转换后的查询与所述转换后的信息相关的解，其中所述转换后的查询包括通过使用所述至少一个密钥获得的所接收的查询的转换后的表示。

[0007] 在本发明的另一示例性实施例中，一种装置包括：至少一个存储器，其被配置为针对主网络中的至少一个主节点的至少一个主属性存储原始信息，其中所述主网络包括互连的多个主节点，其中所述主网络中的所述多个主节点之间的连接基于以下项中的至少一个：所述至少一个主属性，以及所述主节点之间的一种或多种类型的关系和相互依赖性；以及至少一个处理器，其被配置为响应于接收对所存储的原始信息的查询，使用至少一个密钥并基于所接收的查询，将所存储的原始信息转换为转换后的信息，其中所接收的查询涉及所述至少一个主属性的至少一个被查询的属性，其中所述转换后的信息包括所述至少一个主节点的所述至少一个被查询的属性的转换后的数据，其中所述转换后的信息表示包括多个代理节点的代理网络，所述多个代理节点对应于来自所述主网络的所述多个主节点的至少一部分，其中所述转换后的信息并不对应于完整的所存储的原始信息，其中所述转换后的信息被配置为使能针对所述转换后的信息执行至少一个操作，而无需完整的所存储的原始信息的特定知识并且无需泄露完整的所存储的原始信息，其中所述转换后的信息还被配置为使得拥有所述至少一个密钥的人员能够将所述至少一个操作的输出与所存储的原始信息相关，其中所述至少一个处理器还被配置为生成将转换后的查询与所述转换后的信息相关的解，其中所述转换后的查询包括通过使用所述至少一个密钥获得的所接收的查询的转换后的表示。

[0008] 在本发明的另一个示例性实施例中，一种装置包括：用于针对主网络中的至少一个主节点的至少一个主属性存储原始信息的部件，其中所述主网络包括互连的多个主节点，其中所述主网络中的所述多个主节点之间的连接基于以下项中的至少一个：所述至少一个主属性，以及所述主节点之间的一种或多种类型的关系和相互依赖性；用于响应于接

收对所存储的原始信息的查询,使用至少一个密钥并基于所接收的查询,将所存储的原始信息转换成转换后的信息的部件,其中所接收的查询涉及所述至少一个主属性的至少一个被查询的属性,其中所述转换后的信息包括所述至少一个主节点的所述至少一个被查询的属性的转换后的数据,其中所述转换后的信息表示包括多个代理节点的代理网络,所述多个代理节点对应于来自所述主网络的所述多个主节点的至少一部分,其中所述转换后的信息并不对应于完整的所存储的原始信息,其中所述转换后的信息被配置为使能针对所述转换后的信息执行至少一个操作,而无需完整的所存储的原始信息的特定知识并且无需泄露完整的所存储的原始信息,其中所述转换后的信息还被配置为使得拥有所述至少一个密钥的人员能够将所述至少一个操作的输出与所存储的原始信息相关;以及用于生成将转换后的查询与所述转换后的信息相关的解的部件,其中所述转换后的查询包括通过使用所述至少一个密钥获得的所接收的查询的转换后的表示。

附图说明

[0009] 当结合附图阅读时,在以下具体实施方式中,本发明实施例的上述和其它方面变得更加显而易见,这些附图是:

[0010] 图 1 示出其中可以实现本发明的各种示例性实施例的示例性系统的框图;

[0011] 图 2 示出具有实例社交网络的数据的表;

[0012] 图 3A 示出图 2 的表中的数据如何表示主网络;

[0013] 图 3B 示出图 2 的表的转换和求解操作;

[0014] 图 3C 示出可如何使用图 4 的表中的转换后的数据生成代理网络;

[0015] 图 4 是示出实例社交网络的转换后的数据的表;以及

[0016] 图 5 是示出用于实现本发明的示例性实施例的方法和计算机程序执行的一个非限制性实例的流程图。

具体实施方式

[0017] 大规模部署服务、应用和系统以及关联的模板存储(例如,个人数据的存储)更加需要保护存储在系统(例如,社交网络、在线社交网络)中或者用于系统的个人用户数据。该数据的盗窃或盗用可以危及用户的隐私性。此外,被盗的用户数据可以用于危及其它用户数据系统,这些用户数据系统具有针对用户注册的相同特征。对于密码认证,可以通过撤销密码并使用新密码替换密码来管理密码丢失。在人际网络数据标识系统的情况下,不能直接采用这种安全机制,因为人类个人特征的数量较小。

[0018] 一种用于为人类特征给予可撤销性的方法是避免在系统中显式存储人类特征模板,从而消除泄露原始人类特征的任何可能性。可以针对该目的设计多种模板保护技术。这些技术可以被分类为:i)基于特性转换的技术,以及ii)加密系统。在特性转换中,使用用户特定的密钥转换模板,并且仅将转换的模板存储在系统中。在认证期间,同样转换输入特征,并且将转换后的特征与存储的模板相匹配。在加密系统方法中,将外部密钥与模板关联,使得不能从存储的模板获得模板或关联的密钥。作为一个实例,(仅)当为系统提供真正的生物特征识别时,才可以恢复密钥。但是,可能不希望使用真正的生物特征识别,因为它可能更大规模地危及用户隐私性。

[0019] 可以根据两个主要准则评估特性转换技术的安全性:(i)不可逆性,以及(ii)多样性。不可逆性指在给出安全模板的情况下,恢复原始人类特征的难度。多样性指在给出从相同人类特征生成的另一个安全模板的情况下,猜测一个安全模板的难度。这两个准则可以假设也可以不假设对手具有密码知识。但是,这两个度量具有某些局限性。在合理的系统阈值处,评估多样性的错误接受率(FAR)通常为零,因为在评估中使用了有限的数据库。另一方面,用于不可逆性的度量没有考虑人类特征特性的分布,这可以导致安全性的明显过度估计。

[0020] 这种安全问题的其它解决方案包括对人际网络中的数据进行平均,以便将个体数据合并为平均度量,然后使用这些度量进行进一步的分析或模拟。这样,不会泄露原始的个体数据。这种方法无法解决问题,体现在两个重要的方面。第一,它不允许将网络及其组成节点理解为系统的个体且唯一的元素,因此,它使数据中的重要信息和预测能力降级(即,因为将可用的数据仅限于平均数据)。第二,它不允许数据用户在匿名化之后,根据对数据的任意分析来与网络中的个体节点通信。

[0021] 另一种用于实现隐私性的技术是用户使用并非其实际姓名的别名。尽管用户可以使用别名,并且因此破坏他或她的身份与社交网络中表示他们的节点之间的关联,但该节点和其它节点(即使其它节点同样别名化)之间的关系(例如,链接、拓扑图等)可足以标识用户身份。

[0022] 在旨在收集、分析和提供通信链路的系统中,任何根据从人际网络中收集的数据维护隐私性的尝试都应该满足某些要求。

[0023] (1) 隐私系统应该保留与网络中的节点关联的个体的匿名性。这意味着不仅个体的身份不应该与节点数据显式关联,而且节点和其它网络数据(例如拓扑关系、网络统计等)之间的关系也不应该允许网络数据分析领域的技术人员确定个体的身份和网络节点或网络数据之间的关联。

[0024] (2) 隐私系统不应该毁坏通过数据描述的网络的功能属性。该要求说明隐私系统和网络数据的用户应该能够使用数据执行分析,这些分析是有关特定网络和网络中的特定节点的功能的信息性分析。不满足该要求的数据转换的实例说明如下:对来自网络中特定类别的节点的数据进行平均必定使得平均后的数据无用,因为它不能向用户通知有关网络及其节点的功能的数据。

[0025] (3) 隐私系统应该允许网络数据的用户根据对数据的分析,与实际网络及其节点交互,并且从实际网络接收可相对于交互解释的反馈。例如,网络数据的用户应能够向网络中通过分析标识的特定节点发送消息,并且接收回复,其中系统从不泄露实际网络中的这些节点的身份及其与个体身份的关系。

[0026] 此时有用的是,指出上面的要求(2)和(3)直接与上面引用的已知为不可逆性的数学属性相关。这意味着很容易地将某些信息映射到另一种表示,但非常难以找到允许所映射数据的用户恢复原始信息的逆映射。

[0027] 在此公开了用于以安全且私有的方式分析和转换从社交网络(和相关网络)收集的数据的示例性系统、方法和技术。示例性技术促进基于网络数据的查看、分析和网络模拟,同时保留网络中的个体的隐私性。在一个示例性实施例中,可以使用隐藏密钥进行可逆的数据转换,从而允许转换后的数据的用户匿名访问实际网络的成员并且与这些成员交互

(例如,根据分析和模拟的结果)。

[0028] 在本发明中的至少某些示例性实施例中:

[0029] (1) 在社交网络中引入隐私性,但仍然允许对网络数据进行稳健且有用的分析。

[0030] (2) 引入稀疏表示以便将用户网络数据与其它网络数据匹配,从而根据请求获得用户特征。这允许系统使用可能的最少数据量运行。

[0031] (3) 引入以下方法:用于扩展和增加新的结构网络数据,而无需重新计算基本的安全组件。

[0032] 在本发明的至少某些示例性实施例中,这些技术提供多个超过标准方法的优点以便匿名化网络数据,包括:

[0033] (1) 仅使用最少数量的节点或每个节点的数据(例如,使用稀疏表示)提供用户需要的信息,因此降低从其它不考虑的节点暴露信息的风险。

[0034] (2) 它允许对社交网络信息进行稳健且可操作的分析,而无需牺牲用户隐私性。

[0035] 在此公开的示例性系统和方法用于以安全且私有的方式分析和转换从社交网络(和相关网络)收集的数据。例如,研究人员可以分析与表示为网络中节点的人员关联的数据,而无需牺牲用户隐私性。示例性技术促进基于网络数据的查看、分析和网络模拟,同时保留网络中的个体的隐私性。在某些示例性实施例中,可以使用隐藏密钥进行可逆的数据转换,从而允许转换后的数据用户根据分析和模拟的结果,匿名访问实际网络的成员并且与这些成员交互。

[0036] 要指出的是,可以使用本发明的示例性实施例提供或协助多种技术,以便根据社交网络以私有的方式生成和/或使用动态实时报告。这种报告的生成在共同受让的 2010 年 6 月 3 日提交的第 12/793,286 号美国专利申请(其全部内容在此通过参考引入)中描述。

[0037] 在本发明的一个示例性实施例中,一种示例性系统(例如,用于分析社交网络、确保隐私性)包括存储社交网络数据结构的存储器以及隐私引擎。社交网络数据结构包括多个属性(例如,与结构中的人员/节点关联的信息);多个节点,每个节点对应于一个实体并且具有与该实体关联的至少一个属性;以及连接至少两个节点的至少一个连接。所述系统还包括至少一个耦合到存储器的处理器。所述至少一个处理器可操作以执行所述隐私引擎。

[0038] 本发明的各种示例性实施例使用稀疏表示(SR)实现对数据的有意义的分析。在某些示例性实施例中,针对数据的一个或多个部分选择性地执行分析,同时隐藏(例如,掩蔽、匿名化)数据的其它未选择的部分(例如,通过使用仅保留数据的某些特性的代理网络)。这样,可以保留用户匿名性(例如,针对提供给不可信的第三方的数据),同时仍然能够分析数据并且提供有用的结果和信息。

[0039] 下面提供对 SR 和相关技术的一般描述。随后,考虑将 SR 具体应用于本发明的示例性实施例(例如,针对创建和使用代理网络;针对不可信的第三方)。

[0040] 考虑 n 个项目的数据量,每个项目的数据表示为 m 维向量 $h_i \in \mathbb{R}^m$ 。假设 H 是 $m \times n$ 矩阵(例如,字典),其包括项目的个体数据向量 h_i (例如, h_i 是特定训练文档的特征向量)的列,使得 $H = [h_1, h_2, \dots, h_n] \in \mathbb{R}^{m \times n}$ 。矩阵 H 是过完备字典,使得项目数量 n (例如,实例)远大于每个向量 h_i 中的表项数量(即, h_i 的维度,其为 m): $m \ll n$ 。如果给出新

向量或目标向量 $y \in \mathbb{R}^m$, 则 SR 使能针对 β 求解以下等式:

$$[0041] \quad y = H\beta$$

[0042] 其中 β 是稀疏向量或稀疏表示。针对 β 实施稀疏条件以便它从 H 中选择少量表项(例如,实例)以描述 y 。这样,可以使用目标向量 y 通过稀疏向量 β 确定 H 中的类似表项。

[0043] 通常,稀疏表示可以被视为一组参数,这些参数根据一个或多个准则充分描述一组项目中的特定子集。作为非限制性实例,项目组可以包括一组事件、模型、特征、结构、个体(例如,个人、人员)和/或个体组(例如,个人/人员组)。作为非限制性实例,所述准则可以包括以下一个或多个:特征、结构、描述和/或属性。作为进一步的非限制性实例,所述准则可以包括一个或多个属性或其它有关相似性、重要性和/或相关性的指示。作为非限制性实例,可以使用以下一个或多个表示所述数据:向量、矩阵、树和/或其它复杂结构。“稀疏”的指定源于具有相对少量的必要或重要表项(例如,与表项的整体数量相比)的参数组。必要或重要的表项被广泛地解释为在一个或多个基数和/或矩阵中具有一个或多个大值的表项。稀疏表示的不必要或不重要的表项如此小,以便可以安全忽略它们。如在上例中描述的,稀疏表示描述矩阵 H 的哪些列 h_i 必要(例如,哪些列与目标向量 y 匹配)。稀疏表示 β 中的小表项对应于矩阵 H 中可以被忽略的列 h_i (例如,因为它们不对应于目标向量 y)。

[0044] 为了将上面的方法扩展到社交网络,假设 $i=1, 2, \dots, n$ 是 n 个用户的索引,每个用户具有一组特征(例如,年龄、身高、体重、位置、爱好等)或其子集,这些特征存储在 m 维向量 $h_i \in \mathbb{R}^m$ 中。假设 H 是 $m \times n$ 矩阵,其包括个体数据向量 h_i 的列,以便 $H = [h_1, h_2, \dots, h_n] \in \mathbb{R}^{m \times n}$, 其中 $m \ll n$ (如上所述)。假设具有新用户或者由请求者构造的新查询,所述新用户/查询由特征向量 $y \in \mathbb{R}^m$ 表示。用户数据的持有者(例如,可信的中央网络实体)想要知道在 H 中表示的哪些用户与特征向量 y 具有最佳匹配。为了确定这一点,尝试查找以下优化问题的最稀疏解:

$$[0045] \quad \min_{\beta} \|y - H\beta\|_2, \text{ 使得 } \|\beta\|_1 < \varepsilon,$$

[0046] 其中 $\|\beta\|_1$ 指示这是用于界限 β 的 L1 范数(norm)。

[0047] 对于足够小的 ε , 应该具有稀疏解 β , 该解仅具有几个非零表项,这些表项指向 H 中新用户或查询 y 类似的用户/表项。如果矩阵 H 具有受限等距属性(即,“近正交”,RIP), 则解唯一的概率很高,并且 y 可以通过 H 中的元素准确地表示。

[0048] 参考:T. N. Sainath、A. Carmi、D. Kanevsky 和 B. Ramabhadran 的“Bayesian Compressive Sensing for Phonetic Classification(用于语音分类的贝叶斯压缩感知)”(ICASSP 会议记录,德克萨斯州达拉斯,2010年3月),以便获得可如何使用 SR 对特征向量 y 与矩阵 H 中的用户之间的关系进行分类和分析的描述。

[0049] 在上面的描述中,注意因为用户的数据向量 h_i 可以是并且通常是所有可用数据的子集(例如,所有用户的所有数据的子集),并且还因为目标特征向量 y 由请求者指定,所以矩阵 H 将特定于查询 y 。因此,矩阵 H 也可以包括所有可用数据的子集。

[0050] 如上所述,用户数据存在各种安全问题。例如,在两方之间(例如,用户和可信网

络实体,两个用户之间)传输用户数据的全部或部分的情况下,如果没有适当地保护数据传输,则中间代理可获得数据。作为另一个实例,如果要针对数据执行一个或多个分析(例如,以便确定网络中的年龄属性,根据一个或多个属性查找集线器,确定入度和出度,用于营销目的),则可能希望防止执行这些分析的代理(例如,不可信的第三方)能够发现个体用户的身份。

[0051] 考虑到这些安全问题,可以掩蔽 H 以便在不丢失用户特征之间的线性重新分区属性的情况下,防止入侵者或不可信方获得有关用户的私有信息。下面描述用于掩蔽用户数据的一种示例性技术。

[0052] 可以将 y 和 H 乘以矩阵 $A \in \mathbb{R}^{m \times m}$, 以便获得 $z=Ay$ 和 $W=AH$ 。然后问题:

$$[0053] \quad \min_{\beta} \|z-W\beta\|_2$$

[0054] 可以使用与上面讨论相同的优化方法,同时掩蔽 y 和 H 的用户数据。这样,仍然可以分析 y , 但 y 的属性保持隐藏(即,对于没有矩阵 A 的某些人隐藏)。

[0055] 作为一个实例, A 可以由持有其作为私钥的可信方构造。可信方可以根据数据请求者施加的约束来构造 A 。请求者可以选择这些约束以便允许对数据进行有意义的分析,随后使用 z 对查询进行有意义的构造。可信方然后可以检验 A 进行的转换保持网络(现在以 W 表示)中的节点具有某个准则级别的匿名性。因为仅可信方有权访问 A , 所以数据对所有其它方(例如,请求者)保持匿名,但可信方可以恢复节点标识符(例如,用户标识),以便促进数据请求者和由查询向量 z 表示的网络中的实际节点(例如,用户)之间的通信。

[0056] 尽管上面针对稀疏表示进行了描述,但本发明的示例性实施例并不限于这些表示,并且可以与非稀疏表示结合使用。作为一个非限制性实例,当矩阵 H 具有满秩(即,相同数量的列和行,例如 $n \times n$ 矩阵)时,可以创建非稀疏表示。在这种情况下,等式 $y=H\beta$ 具有非稀疏的唯一解。上面针对稀疏表示描述的所有操作(例如,安全问题、矩阵 / 密钥 A)同样适用于非稀疏表示。

[0057] 在本发明的至少某些示例性实施例中,与标准加密方法的差异在于能够基于数据请求者施加的约束,将用户数据映射到转换后的网络(例如,代理网络),同时保持数据的可逆性(例如,针对密钥的持有者)。具体地说,施加这些约束的方式允许保留实体之间的线性关系,以便允许对节点进行有意义的分析。

[0058] 本发明的一个示例性实施例是一种用于匿名推销以便在社交网络中选择个体的系统。

[0059] 考虑从社交网站收集的数据。该数据可以包括有关个体的好恶的信息,以及有关个体与站点上其它个体的关系的信息。该数据可以用于创建节点之间的关系网络,其中每个节点与网络的个体用户关联。网络数据则包括节点数据(例如,个体的偏好)和关系数据(例如,节点之间的连接、个体朋友关系)。使用示例性方法提取该数据并且将其转换成“代理网络”。代理网络与原始网络具有相同数量的节点,但这些代理节点在节点数据以及与它们关联的关系数据方面都进行转换。

[0060] 营销专业人员或其它授权研究人员有权访问转换后的数据,但没有用于执行转换的隐藏密钥。专业人员可以针对代理网络执行多个分析,以便例如在网络中查找其代理可以服务的不同市场。尽管这些市场可以从代理网络中提取(并且因此可以称为“代理市场”),但它们对于专业人员而言具有价值,因为可以使用示例性方法直接将代理网络相关

地返回到实际网络。

[0061] 具体地说,作为非限制性实例,营销专业人员可以设计多种营销策略以便应用于代理网络,包括调查、直接电子邮件营销和定向广告。专业人员部署这些策略,并且示例性系统然后使用隐藏密钥将这些策略转换为类似的策略以便应用于实际社交网络。然后基于该转换将营销策略部署到实际网络,并且例如使用来自调查、电子邮件回复和点击广告数据的聚合数据,在实际网络中测量这些策略的影响。然后示例性系统使用隐藏密钥转换影响测量,以便通过代理网络为营销专业人员提供反馈和回复。然后可以使用该反馈以便例如对营销策略进行重复和进一步发展。

[0062] 这样,可以基于对一个或多个代理网络的分析,向网络(例如,社交网络)中的特定节点执行营销,这些代理网络保留有关网络属性、动态性以及(例如,营销专业人员需要的)输入的响应性(和响应度)的信息,而同时保持与网络节点关联的个体的隐私性(例如,保持与节点关联的数据(例如个人数据)的隐私性)。如果了解这种系统,则可以鼓励个体共享更多的个人数据,确信数据将保持私有并且仅在代理(匿名化)网络的上下文中使用,因此为个体提供更相关的营销,同时保持他或她的隐私性。

[0063] 本发明的一个或多个示例性实施例或其元素可以以计算机程序产品的形式实现,计算机程序产品例如包含在计算机可读存储介质中,计算机可读存储介质具有用于执行指示的步骤(例如,示例性方法)的计算机程序代码。此外,本发明的一个或多个示例性实施例或其元素可以以装置的形式实现,所述装置包括至少一个存储器和至少一个处理器(例如,其耦合到存储器并且可操作以便执行示例性方法步骤)。更进一步,在另一个方面,本发明的一个或多个示例性实施例或其元素可以以部件的形式实现以便执行在此描述的一个或多个方法步骤;所述部件可以包括(i)一个或多个硬件模块,(ii)在一个或多个硬件处理器上执行的一个或多个软件模块,或者(iii)硬件和软件模块的组合。(i)-(iii)的任何一个可以单独或组合地用于实现在此给出的各种示例性实施例。在某些示例性实施例中,一个或多个软件模块存储在至少一个计算机可读介质中。

[0064] 在某些示例性实施例中,分析和/或数据收集可以使用云计算范例,其中按需将共享资源、软件和/或信息远程提供给计算机和其它设备。云计算可能需要使用因特网和/或远程服务器以便维护数据和软件应用。

[0065] 示例性系统和方法存在多个应用,这些系统和方法使得实体(例如,公司、医疗保健专业人员、个体、营销人员、人口统计学家、求职者、招聘者)分析与人员(表示为网络中的节点)关联的数据(例如,有关人员的信息),而无需牺牲用户隐私性。当实现本发明的示例性实施例时,应该对所有相关法律和道德标准加以应有注意和考虑,这些标准与使用社交网络中的数据等有关。在某些示例性实施例中,如果需要额外隐私性,则可以使用“选择加入(opt-in)”系统,其中用户明确同意对与他们相关的信息的每一次使用。在这种示例性实施例中,在存储器中存储属性之前,可能必须接收“选择加入”同意。

[0066] 其中对实际社交网络的分析是有利的一个示例性应用是关于医疗保健。涉及社交网络的最近研究表明一个人的健康习惯不仅影响他或她自己的健康,而且还对他或她的社交网络的成员具有直接且可衡量的影响。此外,已经发现这种影响扩展到直接社交网络成员的社交网络。例如,最近的研究表明,社交网络对健康行为的影响远大于先前的猜测。作为一个实例,一个人决定戒烟会强烈地受他或她的社交网络中的人员是否戒烟的影响一甚

至受他们本身不认识的人员的影响。实际上,吸烟者的全部社交网络看似几乎同时戒烟,这是哈佛医学院的研究人员、医疗社会学家 Nicholas Christakis 和加利福尼亚大学圣地亚哥分校的政治科学家 James Fowler 研究得出的。作为另一个实例,肥胖症可以在社交群体中遵循一种模式,在一个人身上出现并且以“病毒”方式“传播”给另一个人。有关这些研究的更详细的讨论,参见 Thompson C. 的“Are Your Friends Making You Fat? (您的朋友让您变胖了吗?)”,纽约时报杂志,2009 年 9 月 10 日。

[0067] 此外,研究人员使用最流行的社交网络之一 **Facebook®** 进行各种社交网络研究。**Facebook®** 是一个社交网站,其由 Facebook, Inc. 运营和拥有。作为一个实例,对“幸福”的传播感到好奇的研究人员查找在其资料图片中微笑的用户。研究人员发现“微笑简档的聚集方式与弗雷明汉心脏研究中的快乐聚集方式几乎相同”。有关该研究的更详细的讨论,参见 Landau E. 的“Happiness is Contagious in Social Networks (快乐可在社交网络中传染)”,可从 CNN.com 在线获得,2008 年 12 月 5 日。

[0068] 此外,研究人员已表明,即使一个人从未遇见的人员(例如朋友的朋友)也可以影响这个人(例如,“促使”这个人戒烟、放弃饮食过量、变得更快乐)。这种现象可由个体与其它类似他们的个体关联或联系的趋势来解释。这可以在“实际”和纯电子社交网络两者中发生。物理接近度似乎不是关键因素。例如,关于肥胖症,配偶对彼此的影响似乎没有朋友的影响大。一些研究人员认为行为可以忽略联系人。例如,行为可以传播给朋友的朋友,而不会影响联系他们的人员。在一个或多个示例性实施例中,通过包含两个或更多数据节点以及两个或更多连接的路径(例如,通过社交网络中的一个或多个其它节点),将至少一个数据节点连接到主题节点。

[0069] 术语“实体”(例如,如针对网络或社交网络使用的)可以对应于一个或多个人员、人员组或人员安排。在本发明的某些示例性实施例中,作为非限制性实例,实体包括至少一个业务、至少一个产品、至少一个服务、至少一个头像(例如,人员、业务或其它实体在虚拟世界或虚拟环境中的表示)和 / 或至少一个人员。在其中实体包括至少一个人员的这些示例性实例中,与该实体关联的一个或多个属性可以包括以下项中的至少一个:职业信息(例如,与人员的职业、培训和 / 或工作经验相关的信息)、技能信息、健康信息、偏好信息以及个人信息(例如,地址、年龄、朋友、社交网络、家庭)。

[0070] 作为一个实例,要指出的是,如果分析确定用户在其社交网络的一部分中具有大量图书作者、工程师或文字编辑,则报告模块可以生成有关此发现的报告。在某些示例性实施例中,将该报告发送给用户或用户指定的实体。这可以以安全方式进行,例如,如针对本发明的示例性实施例概述的那样。对于评估连接类型和用户行为的潜在影响的用户而言,这可以很有用。它还可以为用户提供信息,用户可以使用该信息获得帮助,例如编辑图书或找工作方面的帮助。

[0071] 图 1 示出其中可以实现本发明的各种示例性实施例的示例性系统的框图。系统 100 可以包括至少一个电路 102 (例如,电路元件、电路组件、集成电路),其在某些示例性实施例中可以包括至少一个处理器 104。系统 100 还可以包括至少一个存储器 106 (例如,易失性存储器件、非易失性存储器件)和 / 或至少一个存储装置 108。作为非限制性实例,存储装置 108 可以包括非易失性存储器件(例如,EEPROM、ROM、PROM、RAM、DRAM、SRAM、闪存、固件、可编程逻辑等)、磁盘驱动器、光盘驱动器和 / 或磁带驱动器。作为非限制性实例,存

储装置 108 可以包括内部存储设备、附加的存储设备和 / 或网络可访问的存储设备。系统 100 可以包括至少一个程序逻辑 110, 其包括可以加载到存储器 106 并且由处理器 104 和 / 或电路 102 执行的代码 112 (例如, 程序代码)。在某些示例性实施例中, 程序逻辑 110 (包括代码 112) 可以存储在存储装置 108 中。在某些其它示例性实施例中, 程序逻辑 110 可以在电路 102 中实现。因此, 尽管图 1 独立于其它元件而示出程序逻辑 110, 但作为非限制性实例, 程序逻辑 110 可以在存储器 106 和 / 或电路 102 中实现。

[0072] 系统 100 可以包括至少一个通信组件 114, 其实现与至少一个其它组件、系统、设备和 / 或装置通信。作为非限制性实例, 通信组件 114 可以包括被配置为发送和接收信息的收发器、被配置为发送信息的发送器和 / 或被配置为接收信息的接收器。作为一个非限制性实例, 通信组件 114 可以包括调制解调器或网卡。作为非限制性实例, 图 1 的系统 100 可以包含在计算机或计算机系统中, 例如台式计算机、便携式计算机或服务器中。作为非限制性实例, 可以使用一个或多个内部总线、连接、电线和 / 或 (印制) 电路板, 将图 1 中所示的系统 100 的组件连接或耦合在一起。

[0073] 应该指出, 根据本发明的示例性实施例, 电路 102、处理器 (多个) 104、存储器 106、存储装置 108、程序逻辑 110 和 / 或通信组件 114 中的一个或多个可以存储在此讨论的各种项目 (例如, 数据、数据库、表、项目、向量、矩阵、变量、等式、公式、运算、运算逻辑、逻辑) 的一个或多个。作为一个非限制性实例, 上面标识的一个或多个组件可以接收和 / 或存储信息和 / 或转换后的信息。作为进一步的非限制性实例, 上面标识的一个或多个组件可以接收和 / 或存储在此描述的功能 (多个)、操作、功能组件和 / 或操作组件。

[0074] 本发明的示例性实施例可以由处理器 104 实现的计算机软件、或硬件、或硬件和软件的组合来执行。作为一个非限制性实例, 本发明的示例性实施例可以由一个或多个集成电路实现。作为非限制性实例, 存储器 106 可以具有适合于技术环境的任何类型, 并且可以使用任何适当的数据存储技术实现, 例如光存储器件、磁存储器件、基于半导体的存储器件、固定存储器和可移动存储器。作为非限制性实例, 处理器 104 可以具有适合于技术环境的任何类型, 并且可以包含基于多核体系架构的微处理器、通用计算机、专用计算机和处理器中的一个或多个。

[0075] 出于示例目的, 下面提供一个用于使用本发明的示例性实施例的非限制性实例。参考图 2-4 描述该示例性实现。要指出的是, 下面对“数据”、“表” (例如, 表 H) 或“矩阵” (例如, 矩阵 H) 的引用可以在如下假设下互换使用: 所述数据同样可以以表形式或矩阵形式表示, 而无需实质性更改也不会丢失信息。

[0076] 考虑具有多个主节点的主网络 N_p , 其中每个主节点具有至少一个主属性。假设该主网络是包括多个实体 (例如, 人员) 的社交网络, 这些实体基于一个或多个主属性 (例如, 彼此的朋友) 连接。作为非限制性实例, 用于这些主节点和关联的主属性的数据可以以任何合适的形式 (例如矩阵、表或数据库) 表示。

[0077] 图 2 示出具有实例社交网络的数据的表 H。如可以看到的, 其中表示多个主属性, 包括包含标识数据 (例如, 姓名、位置、年龄) 和 / 或可以用于标识所述实体的数据 (例如, 生日、交友) 的主属性。图 3A 示出图 2 的表 H 中的数据如何表示主网络 N_p , 其中主节点之间的连接基于“交友”数据 (假设这些关系是定向的)。

[0078] 对于该实例, 假设广告商想要购买针对社交网络 N_p 的定向广告。此外, 出于示例

目的,假设该广告商是寻求扩展市场的优良美洲驼的供应商。广告商与网络运营商签订合同并且需要保护用户的隐私性,网络运营商同意向广告商公开某些相关信息。鉴于此,对数据的查询将用于那些喜欢美洲驼的用户。为了使能进一步分析潜在市场,广告商还请求提供年龄、关系(交友)和冰淇淋偏好的对应数据。

[0079] 鉴于该查询,网络运营商使用私钥 A 转换表 H 中的数据,以便可以为广告商提供结果(表 W)。图 3B 示出转换和求解操作。更具体地说,使用私钥 A,转换矩阵 H 和查询 q 两者(例如,由网络运营商转换),以便分别获得形式为表 W 的转换后的数据和转换后的查询 q'。使用这些结果(W 和 q') (例如,由网络运营商、广告商、第三方使用)获得上面描述的表示 β (例如,稀疏表示)。要指出的是,不与广告商或第三方共享私钥 A。此外,不可从转换结果(W 和 q') 确定私钥 A。此外,在没有私钥 A 的情况下,转换是不可逆的。

[0080] 图 4 是示出实例社交网络的转换后的数据的表 W。在该实例中,转换执行多个操作,包括:删除(例如,掩蔽、删除、替换)针对喜欢美洲驼指示“否”的那些实体的其它数据;删除感兴趣属性(年龄、交友、喜欢美洲驼、喜欢冰淇淋)之外的所有属性的数据;删除(例如,掩蔽、删除、替换)针对喜欢美洲驼指示“否”的那些实体的交友数据;以及将年龄数据归纳为单独的类别或范围(例如,将年龄“26”的数据更改为“25-29”)。如可以理解的,得到的转换后的数据 W 不能标识对应的实体(例如,具有的数据不足以标识实体)。此外,维护返回到原始数据 H 的对应性,以便网络运营商可以确定实体的身份,并且联系或以其它方式定向该实体(例如,针对定向广告)。

[0081] 图 3C 示出可如何使用转换后的数据 W 生成代理网络 N_s 。作为非限制性实例,代理网络 N_s 可以用于表示转换后的数据 W 和 / 或使能进一步分析转换后的数据 W。作为非限制性实例,广告商可以将转换后的数据 W 和 / 或代理网络 N_s 用于定向广告。

[0082] 针对转换后的查询 q', 在一种解释中,还可以隐式看到原始查询 q 包括元素“and we do not need data for the other primary properties (我们不需要其它主属性的数据)”。在这种情况下,可以看到转换后的查询 q' 具有针对其执行的操作,这些操作类似于为了获得转换后的数据 W 而针对原始数据 H 执行的那些操作。例如,转换后的查询 q' 可以包括对不必要的主属性的“查询”,这些主属性属于空集(\emptyset)、噪声或者以其它方式被隐藏或删除。将相同转换(基于相同私钥 A)应用于原始查询 q 以及原始数据 H,这可确保解(表示 β) 针对原始项目 H 和 q 以及转换的项目 W 和 q' 保持相同。

[0083] 所属技术领域的普通技术人员应该理解可用于转换的各种选项。此外,所属技术领域的普通技术人员应该理解可以使用转换后的数据和 / 或代理网络执行的各种操作和 / 或分析。

[0084] 下面立即描述另一个示例性实施例。应该指出,对应性(例如,符号、标识的项目、圆括号)仅是示例性和非限制性的。在该示例性实施例中,针对主网络(社交网络 N_p) 中的至少一个主节点的至少一个主属性(属性、方面、元素、描述),(在矩阵或表 H 中)存储原始信息(数据)。主网络(N_p) 包括互连的多个主节点(社交网络中的每个实体 / 人员一个节点),其中主网络(N_p) 中的多个主节点之间的连接基于以下项中的至少一个:至少一个主属性,以及主节点之间的一种或多种类型的关系和相互依赖性(实体 / 人员之间的朋友关系)。响应于接收对存储的原始信息(H)的查询(q),使用至少一个密钥(私钥 A)并基于接收的查询(q),将存储的原始信息(H)转换为转换后的信息(矩阵或表 W)。接收的查询(q)涉及至少

一个主属性的至少一个被查询的属性(目标数据、目标属性、目标信息)。转换后的信息(W)包括至少一个主节点的至少一个被查询的属性的转换后的数据(W的数据/表项)。转换后的信息(W)表示包括多个代理节点的代理网络(N_s),多个代理节点对应于来自主网络的多个主节点的至少一部分(代理网络可以包括主节点的子集或全部)。转换后的信息(W)并不对应于完整的存储的原始信息(掩蔽或删除标识信息)。转换后的信息(W)被配置为使能针对转换后的信息(W)执行至少一个操作,而无需完整的存储的原始信息的特定知识并且不会泄露完整的存储的原始信息(可以针对W执行分析而不能标识实体/人员)。转换后的信息(W)还被配置为使得拥有至少一个密钥(A)的人员能够将至少一个操作的输出与存储的原始信息相关(A的拥有者可以标识所述实体)。生成将转换后的查询(q')与转换后的信息(W)相关的解(β)。转换后的查询(q')包括通过使用至少一个密钥(A)获得的接收的查询(q)的转换后的表示。

[0085] 下面是本发明的各种非限制性的示例性实施例的进一步描述。为了清晰起见,分别对下面描述的示例性实施例进行编号。这种编号不应该被解释为完全分离各种示例性实施例,因为一个或多个示例性实施例的各个方面可以与一个或多个其它方面或示例性实施例结合实现。

[0086] (1)在本发明的另一个示例性实施例中,并且如图5中所示,提供一种有形地包含可由机器执行以便执行操作的指令程序的计算机可读存储介质,所述操作包括:针对主网络中的至少一个主节点的至少一个主属性存储原始信息,其中所述主网络包括互连的多个主节点,其中所述主网络中的所述多个主节点之间的连接基于以下项中的至少一个:所述至少一个主属性,以及所述主节点之间的一种或多种类型的关系和相互依赖性(501);响应于接收对所存储的原始信息的查询,使用至少一个密钥并基于所接收的查询,将所存储的原始信息转换成转换后的信息,其中所接收的查询涉及所述至少一个主属性的至少一个被查询的属性,其中所述转换后的信息包括所述至少一个主节点的所述至少一个被查询的属性的转换后的数据,其中所述转换后的信息表示包括多个代理节点的代理网络,所述多个代理节点对应于来自所述主网络的所述多个主节点的至少一部分,其中所述转换后的信息并不对应于完整的所存储的原始信息,其中所述转换后的信息被配置为使能针对所述转换后的信息执行至少一个操作,而无需完整的所存储的原始信息的特定知识并且无需泄露完整的所存储的原始信息,其中所述转换后的信息还被配置为使得拥有所述至少一个密钥的人员能够将所述至少一个操作的输出与所存储的原始信息相关(502);以及生成将转换后的查询与所述转换后的信息相关的解,其中所述转换后的查询包括通过使用所述至少一个密钥获得的所接收的查询的转换后的表示(503)。

[0087] 根据以上任何一项的计算机可读存储介质,其中所述解包括稀疏表示。根据以上任何一项的计算机可读存储介质,其中所述网络包括社交网络或在线社交网络。根据以上任何一项的计算机可读存储介质,其中由无权访问所述至少一个私钥的第三方执行所述至少一个操作。根据以上任何一项的计算机可读存储介质,其中由无权访问所述至少一个私钥的不可信第三方执行所述至少一个操作。根据以上任何一项的计算机可读存储介质,其中所述代理网络针对所述代理节点之间的连接使用至少一个不同于所述主网络的基础。根据以上任何一项的计算机可读存储介质,其中每个主节点对应于多个实体中的一个实体。根据以上任何一项的计算机可读存储介质,其中每个主节点对应于多个实体中的一个不同

实体。根据以上任何一项的计算机可读存储介质,其中多个实体中的每个实体包括以下之一:一个或多个人员、一个或多个组织、一个或多个公司、一个或多个医疗保健专业人员、一个或多个营销人员、一个或多个人口统计学家、一个或多个求职者,或者一个或多个招聘者。

[0088] 根据以上任何一项的计算机可读存储介质,其中由监管所述主网络的网络运营商存储所述原始信息和至少一个密钥。根据以上任何一项的计算机可读存储介质,其中仅所述网络运营商有权访问所述至少一个密钥。根据以上任何一项的计算机可读存储介质,其中所述转换后的信息还被配置为使能针对所述转换后的信息执行所述至少一个操作,而无需访问完整的所存储的原始信息。根据以上任何一项的计算机可读存储介质,其中所述转换运行以便删除、替换或掩蔽所述至少一个主属性中与所述至少一个被查询的属性不相关的其它属性。根据以上任何一项的计算机可读存储介质,所述操作还包括:向查询的提供者提供(例如,传输、发送)解。

[0089] 根据以上任何一项的计算机可读存储介质,其中所述计算机可读存储介质包括至少一个存储器或至少一个程序存储设备。根据以上任何一项的计算机可读存储介质,其中所述机器包括计算机或至少一个被配置为执行指令程序的处理器。根据以上任何一项的计算机可读存储介质,还包括在此进一步描述的本发明的示例性实施例的一个或多个方面。

[0090] (2) 在本发明的一个示例性实施例中,并且如图 5 中所示,一种方法包括:针对主网络中的至少一个主节点的至少一个主属性(例如,在至少一个存储器、至少一个存储设备、至少一个计算机可读存储介质上)存储原始信息,其中所述主网络包括互连的多个主节点,其中所述主网络中的所述多个主节点之间的连接基于以下项中的至少一个:所述至少一个主属性,以及所述主节点之间的一种或多种类型的关系和相互依赖性(501);响应于接收对所存储的原始信息的查询,(例如,由第一装置、至少一个处理器、至少一个集成电路)使用至少一个密钥并基于所接收的查询,将所存储的原始信息转换成转换后的信息,其中所接收的查询涉及所述至少一个主属性的至少一个被查询的属性,其中所述转换后的信息包括所述至少一个主节点的所述至少一个被查询的属性的转换后的数据,其中所述转换后的信息表示包括多个代理节点的代理网络,所述多个代理节点对应于来自所述主网络的所述多个主节点的至少一部分,其中所述转换后的信息并不对应于完整的所存储的原始信息,其中所述转换后的信息被配置为使能针对所述转换后的信息执行至少一个操作,而无需完整的所存储的原始信息的特定知识并且无需泄露完整的所存储的原始信息,其中所述转换后的信息还被配置为使得拥有所述至少一个密钥的人员能够将所述至少一个操作的输出与所存储的原始信息相关(502);以及(例如,由第二装置、至少一个处理器、至少一个集成电路)生成将转换后的查询与所述转换后的信息相关的解,其中所述转换后的查询包括通过使用所述至少一个密钥获得的所接收的查询的转换后的表示(503)。

[0091] 根据以上任何一项的方法,其中所述第一装置包括所述第二装置。根据以上任何一项的方法实现为计算机程序。根据以上任何一项的方法实现为存储(例如,包含)在程序存储设备(例如,至少一个存储器,至少一个计算机可读介质)上并可由计算机(例如,至少一个计算机)执行的指令的程序。根据以上任何一项的方法还包括在此进一步描述的本发明的示例性实施例的一个或多个方面。

[0092] (3) 在本发明的进一步示例性实施例中,一种装置包括:至少一个存储器,其被配

置为针对主网络中的至少一个主节点的至少一个主属性存储原始信息,其中所述主网络包括互连的多个主节点,其中所述主网络中的所述多个主节点之间的连接基于以下项中的至少一个:所述至少一个主属性,以及所述主节点之间的一种或多种类型的关系和相互依赖性;以及至少一个处理器,其被配置为响应于接收对所存储的原始信息的查询,使用至少一个密钥并基于所接收的查询,将所存储的原始信息转换成转换后的信息,其中所接收的查询涉及所述至少一个主属性的至少一个被查询的属性,其中所述转换后的信息包括所述至少一个主节点的所述至少一个被查询的属性的转换后的数据,其中所述转换后的信息表示包括多个代理节点的代理网络,所述多个代理节点对应于来自所述主网络的所述多个主节点的至少一部分,其中所述转换后的信息并不对应于完整的所存储的原始信息,其中所述转换后的信息被配置为使能针对所述转换后的信息执行至少一个操作,而无需完整的所存储的原始信息的特定知识并且无需泄露完整的所存储的原始信息,其中所述转换后的信息还被配置为使得拥有所述至少一个密钥的人员能够将所述至少一个操作的输出与所存储的原始信息相关,其中所述至少一个处理器还被配置为生成将转换后的查询与所述转换后的信息相关的解,其中所述转换后的查询包括通过使用所述至少一个密钥获得的所接收的查询的转换后的表示。

[0093] 根据以上任何一项的装置还包括在此描述的本发明的示例性实施例的一个或多个方面。

[0094] (4)在本发明的另一个示例性实施例中,一种装置包括:用于针对主网络中的至少一个主节点的至少一个主属性存储原始信息的部件,其中所述主网络包括互连的多个主节点,其中所述主网络中的所述多个主节点之间的连接基于以下项中的至少一个:所述至少一个主属性,以及所述主节点之间的一种或多种类型的关系和相互依赖性;用于响应于接收对所存储的原始信息的查询,使用至少一个密钥并基于所接收的查询,将所存储的原始信息转换成转换后的信息的部件,其中所接收的查询涉及所述至少一个主属性的至少一个被查询的属性,其中所述转换后的信息包括所述至少一个主节点的所述至少一个被查询的属性的转换后的数据,其中所述转换后的信息表示包括多个代理节点的代理网络,所述多个代理节点对应于来自所述主网络的所述多个主节点的至少一部分,其中所述转换后的信息并不对应于完整的所存储的原始信息,其中所述转换后的信息被配置为使能针对所述转换后的信息执行至少一个操作,而无需完整的所存储的原始信息的特定知识并且无需泄露完整的所存储的原始信息,其中所述转换后的信息还被配置为使得拥有所述至少一个密钥的人员能够将所述至少一个操作的输出与所存储的原始信息相关;以及用于生成将转换后的查询与所述转换后的信息相关的解的部件,其中所述转换后的查询包括通过使用所述至少一个密钥获得的所接收的查询的转换后的表示。

[0095] 根据以上任何一项的装置还包括在此描述的本发明的示例性实施例的一个或多个方面。

[0096] (5)在本发明的进一步示例性实施例中,一种装置包括:存储电路,其被配置为针对主网络中的至少一个主节点的至少一个主属性存储原始信息,其中所述主网络包括互连的多个主节点,其中所述主网络中的所述多个主节点之间的连接基于以下项中的至少一个:所述至少一个主属性,以及所述主节点之间的一种或多种类型的关系和相互依赖性;转换电路,其被配置为响应于接收对所存储的原始信息的查询,使用至少一个密钥并基于

所接收的查询,将所存储的原始信息转换成转换后的信息,其中所接收的查询涉及所述至少一个主属性的至少一个被查询的属性,其中所述转换后的信息包括所述至少一个主节点的所述至少一个被查询的属性的转换后的数据,其中所述转换后的信息表示包括多个代理节点的代理网络,所述多个代理节点对应于来自所述主网络的所述多个主节点的至少一部分,其中所述转换后的信息并不对应于完整的所存储的原始信息,其中所述转换后的信息被配置为使能针对所述转换后的信息执行至少一个操作,而无需完整的所存储的原始信息的特定知识并且无需泄露完整的所存储的原始信息,其中所述转换后的信息还被配置为使得拥有所述至少一个密钥的人员能够将所述至少一个操作的输出与所存储的原始信息相关;以及求解电路,其被配置为生成将转换后的查询与所述转换后的信息相关的解,其中所述转换后的查询包括通过使用所述至少一个密钥获得的所接收的查询的转换后的表示。

[0097] 根据以上任何一项的装置体现为一个或多个集成电路。根据以上任何一项的装置还包括在此描述的本发明的示例性实施例的一个或多个方面。

[0098] 如在此讨论的并且如针对示例性方法具体描述的,本发明的示例性实施例可以与程序存储设备(例如,至少一个存储器)结合实现,所述程序存储设备可由机器读取并有形地包含可由机器执行以便执行操作的指令程序(例如,程序或计算机程序)。所述操作包括使用示例性实施例的步骤或所述方法的步骤。

[0099] 图5中所示的方框还可以被视为对应于一个或多个功能和/或操作,它们由一个或多个组件、电路、芯片、装置、处理器、计算机程序和/或功能块执行。上面的任何和/或全部可以以任何实现与在此描述的本发明的示例性实施例一致的操作的可实行解决方案或布置来实现。

[0100] 此外,图5中所示的方框布置应被视为只是示例性和非限制性的。应该理解,图5中所示的方框可以对应于一个或多个功能和/或操作,它们可以以任何顺序(例如,任何合适、可实行和/或可行的顺序)和/或同时(例如,合适、可实行和/或可行地)执行,以便实现本发明的一个或多个示例性实施例。此外,可以将一个或多个其它功能、操作和/或步骤与图5中所示的这些功能、操作和/或步骤结合使用,以便实现本发明的一个或多个进一步的示例性实施例。

[0101] 即,图5中所示的本发明的示例性实施例可以以任意组合(例如,合适、可实行和/或可行的任意组合)与一个或多个进一步方面结合使用、实现或实行,并且并不仅限于图5中所示的步骤、方框、操作和/或功能。

[0102] 如在此使用的,“查询”被视为针对数据集的询问。询问可以采取任何合适的形式,包括但不限于:检索、排序、搜索、形成子集,以及针对数据执行一个或多个操作。作为一个实例,可以定向查询以便返回与一个或多个条件或属性匹配的数据子集。作为另一个实例,可以定向查询以便返回针对数据集执行一个或多个操作的结果。作为非限制性实例,针对以下一个或多个,查询结果的形式可以包括任何合适的形式(例如,向量、矩阵、表、数据库):结果的存储、针对结果的进一步操作和/或结果的显示。

[0103] 如在此使用的,“代理网络”被视为辅助网络、或至少部分地与主网络相关的网络的辅助表示、或网络(例如,主要网络、主网络)的主表示。作为一个实例,代理网络可以保留主网络的某些属性,同时丢弃、掩蔽或匿名化主网络的其它属性。作为另一个实例,代理网络可以包括主网络的子集或一部分。在某些情况下,代理网络可以包括代理节点,这些代理

节点对应(例如,唯一对应、一对一地对应)于主网络中的主节点。代理节点可以保留主节点的某些属性,同时丢弃、掩蔽或匿名化主节点的其他属性。代理网络中的代理节点之间的连接(例如,“代理连接”)可以基于至少一个不同于主网络中的主节点之间的连接(例如,“主连接”)的基础(例如,至少一个不同的属性、关系和/或相互依赖性)。在某些情况下,代理连接基于不同于主连接的一个或多个其它基础。在其它情况下,代理连接仅基于一个或多个不同于主连接的基础。在至少某些情况下,代理网络可以被视为来自主网络的主节点的备选表示,该表示保留来自主网络的某些信息,使得保留的信息少于来自主网络的所有信息。

[0104] 对术语“连接”、“耦合”或其变体的任何使用应被解释为指示标识的元素之间的任何此类连接或耦合(直接或间接)。作为一个非限制性实例,可以在“耦合”元素之间存在一个或多个中间元素。作为非限制性实例,根据所述示例性实施例,标识的元素之间的连接或耦合可以是物理、电、磁、逻辑或它们的任何合适的组合。作为非限制性实例,连接或耦合可以包括一个或多个印制电连接、电线、电缆、介质或它们的任何合适的组合。

[0105] 所属技术领域的技术人员知道,本发明的示例性实施例可以实现为系统、方法或计算机程序产品。因此,本发明的示例性实施例可以具体实现为以下形式,即:完全的硬件实施方式、完全的软件实施方式(包括固件、驻留软件、微代码等),或硬件和软件方面结合的实施方式,这里可以统称为“电路”、“模块”或“系统”。此外,本发明的示例性实施例还可以实现为在一个或多个程序存储器件或计算机可读介质中的计算机程序产品的形式,该程序存储器件或计算机可读介质中包含计算机可读的程序代码。

[0106] 可以采用一个或多个程序存储设备或计算机可读介质的任意组合。计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质。作为非限制性实例,计算机可读存储介质可以包括以下一个或多个:电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者上述的任意合适的组合。计算机可读存储介质的更具体的例子(非穷举的列表)包括:具有一个或多个导线的电连接、便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本文件中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件(例如,一个或多个处理器)使用或者与其结合使用。

[0107] 计算机可读的信号介质可以包括例如在基带中或者作为载波一部分传播的数据信号,其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式,包括但不限于—电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质可以是计算机可读存储介质以外的任何计算机可读介质,该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。

[0108] 计算机可读介质上包含的程序代码可以用任何适当的介质传输,包括—但不限于—无线、有线、光缆、RF等等,或者上述的任意合适的组合。

[0109] 可以以一种或多种程序设计语言的任意组合来编写用于执行本发明的各个方面的操作的计算机程序代码,所述程序设计语言包括面向对象的程序设计语言—诸如 Java、Smalltalk、C++等,还包括常规的过程式程序设计语言—诸如“C”语言或类似的设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机

或服务器上执行。在涉及远程计算机的情形中,远程计算机可以通过任意种类的网络—包括局域网(LAN)或广域网(WAN)—连接到用户计算机,或者,可以连接到外部计算机(例如利用因特网服务提供商(ISP)来通过因特网连接)。

[0110] 在此参照根据本发明的示例性实施例的方法、装置(系统)和计算机程序产品的流程图和/或框图描述本发明的示例性实施例。应当理解,流程图和/或框图的每个方框以及流程图和/或框图中各方框的组合,都可以由计算机程序指令实现。这些计算机程序指令可以提供给通用计算机、专用计算机或其它可编程数据处理装置的处理器,从而生产出一种机器,使得这些指令在通过计算机或其它可编程数据处理装置的至少一个处理器执行时,产生了实现流程图和/或框图中的一个或多个方框中规定的功能/动作的装置。

[0111] 也可以把这些计算机程序指令存储在计算机可读介质中,这些指令使得计算机、其它可编程数据处理装置、或其它设备以特定方式工作,从而,存储在计算机可读介质中的指令就产生出包括实现流程图和/或框图中的一个或多个方框中规定的功能/动作的指令的制造品(article of manufacture)。

[0112] 也可以把计算机程序指令加载到计算机、其它可编程数据处理装置、或其它设备上,使得在计算机、其它可编程装置或其它设备上执行一系列操作步骤,以产生计算机实现的过程,从而使得在计算机或其它可编程装置上执行的指令提供实现流程图和/或框图中的一个或多个方框中规定的功能/动作的过程。

[0113] 附图中的流程图和框图显示了根据本发明的不同示例性实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段或代码的一部分,所述模块、程序段或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能可以以不同于附图中所标注的顺序发生。例如,两个连续的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意,框图和/或流程图中的每个方框、以及可能框图和/或流程图中的方框的组合,可以用执行规定的功能或动作的专用的基于硬件的系统来实现,或者可以用专用硬件与软件(例如,计算机指令)的组合来实现。

[0114] 通常,本发明的不同示例性实施例可以在不同介质(例如软件、硬件、逻辑、专用电路或其任意组合)中实现。作为一个非限制性实例,某些方面可以在可在计算设备上运行的软件中实现,而其它方面可以在硬件中实现。

[0115] 上面通过示例性和非限制性实例提供的描述将提供对发明者目前构想的用于执行本发明的最佳方法和装置的全面和信息性描述。但是,当结合附图和所附权利要求阅读时,鉴于上面的描述,对于相关技术领域的技术人员来说各种修改和变化可以变得显而易见。但是,所有这些和类似的修改仍将落入本发明的示例性实施例的教导范围之内。

[0116] 此外,可以使用本发明的优选实施例的某些特性获利而无需相应地使用其它特性。因此,上面的描述应被视为只是例示本发明的原理,而并非限制本发明。

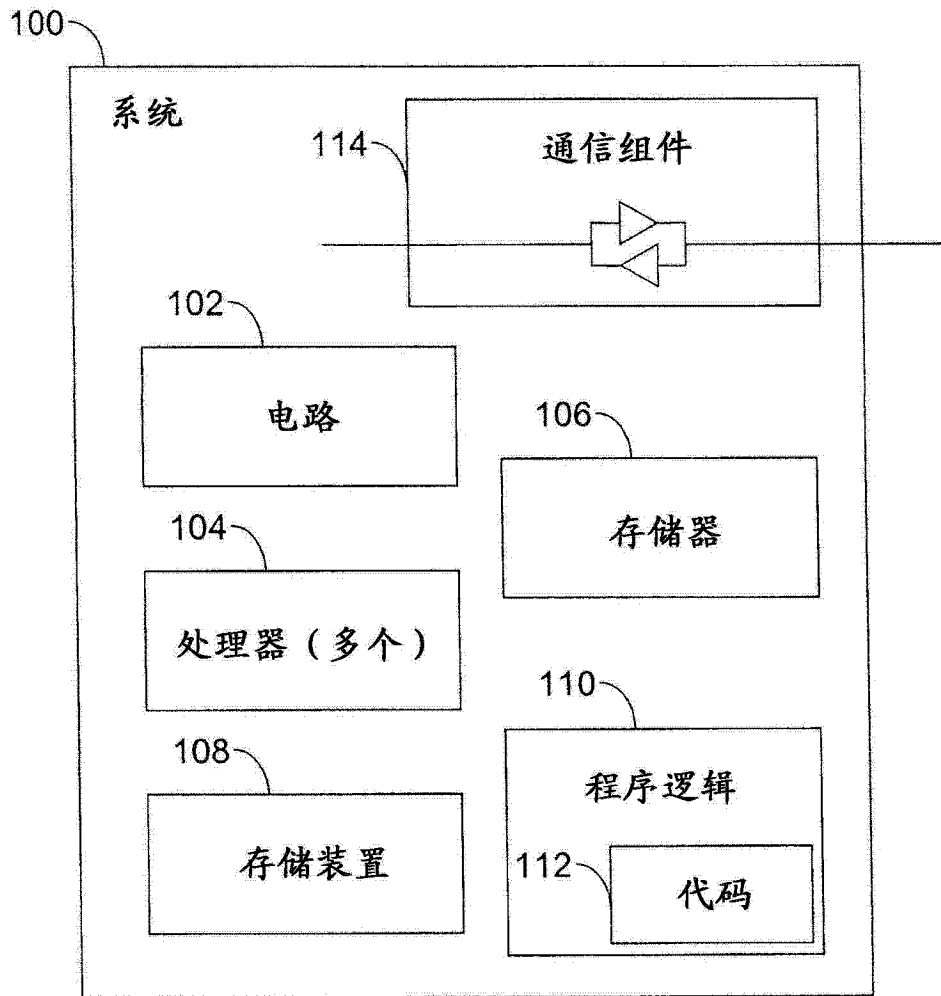


图 1

索引	1	2	3	4	5	6	7	8	n
姓名	Xeni	Howard	JT	Terry	Randall	Doug	Annalee	Corey	...
位置	CA	CT	CT	UK	MA	UK	WY	CA	...
年龄	38	17	15	63	26	42	43	40	...
生日	8/05	1/27	5/25	4/28	10/17	3/11	1/02	7/17	...
交友 (按索引)	3, 4, 6, 8	∅	2, 5, 8	7, 8	4, 8	1, 8	4, 5, 6, 8	2	...
喜欢 美洲鸵 具有博客	否	是	是	否	是	是	否	是	...
戴着护目镜 和披肩	是	否	否	否	是	否	是	是	...
喜欢 热气球	否	是	否	否	否	否	否	是	...
喜欢猛禽	是	否	否	否	是	否	否	否	...
飞萤航空 爱好者	是	否	否	是	否	否	是	否	...
爱好: 地理定位	是	否	否	否	是	否	是	是	...
喜欢 冰淇淋	否	是	否	否	是	否	否	是	...
科幻爱好者	是	是	否	是	是	是	是	是	...
信任海雀	是	否	否	否	是	否	否	否	...

图 2

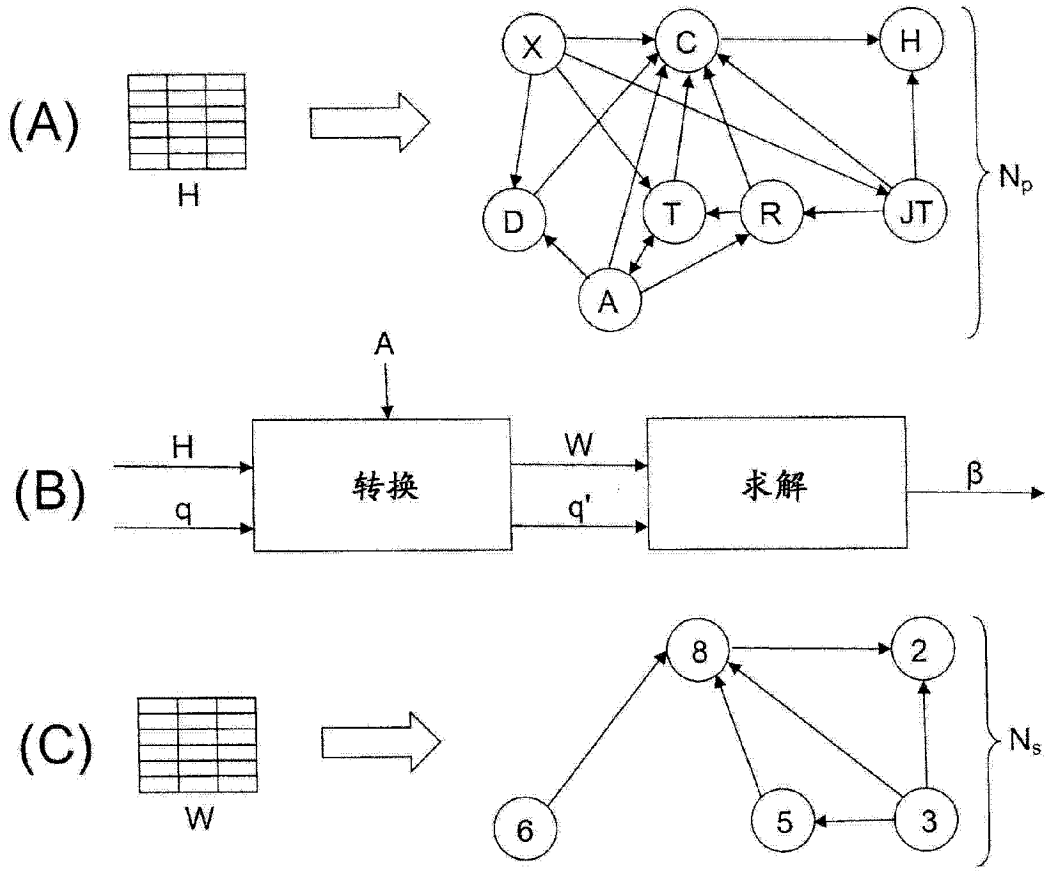


图 3

索引	1	2	3	4	5	6	7	8	n
年龄	\emptyset	15-19	15-19	\emptyset	25-29	40-44	\emptyset	40-44	...
交友 (按索引)	\emptyset	\emptyset	2, 5, 8	\emptyset	8	8	\emptyset	2	...
喜欢 美洲驼	否	是	是	否	是	是	否	是	...
喜欢 冰淇淋	\emptyset	是	否	\emptyset	是	是	\emptyset	是	...

图 4

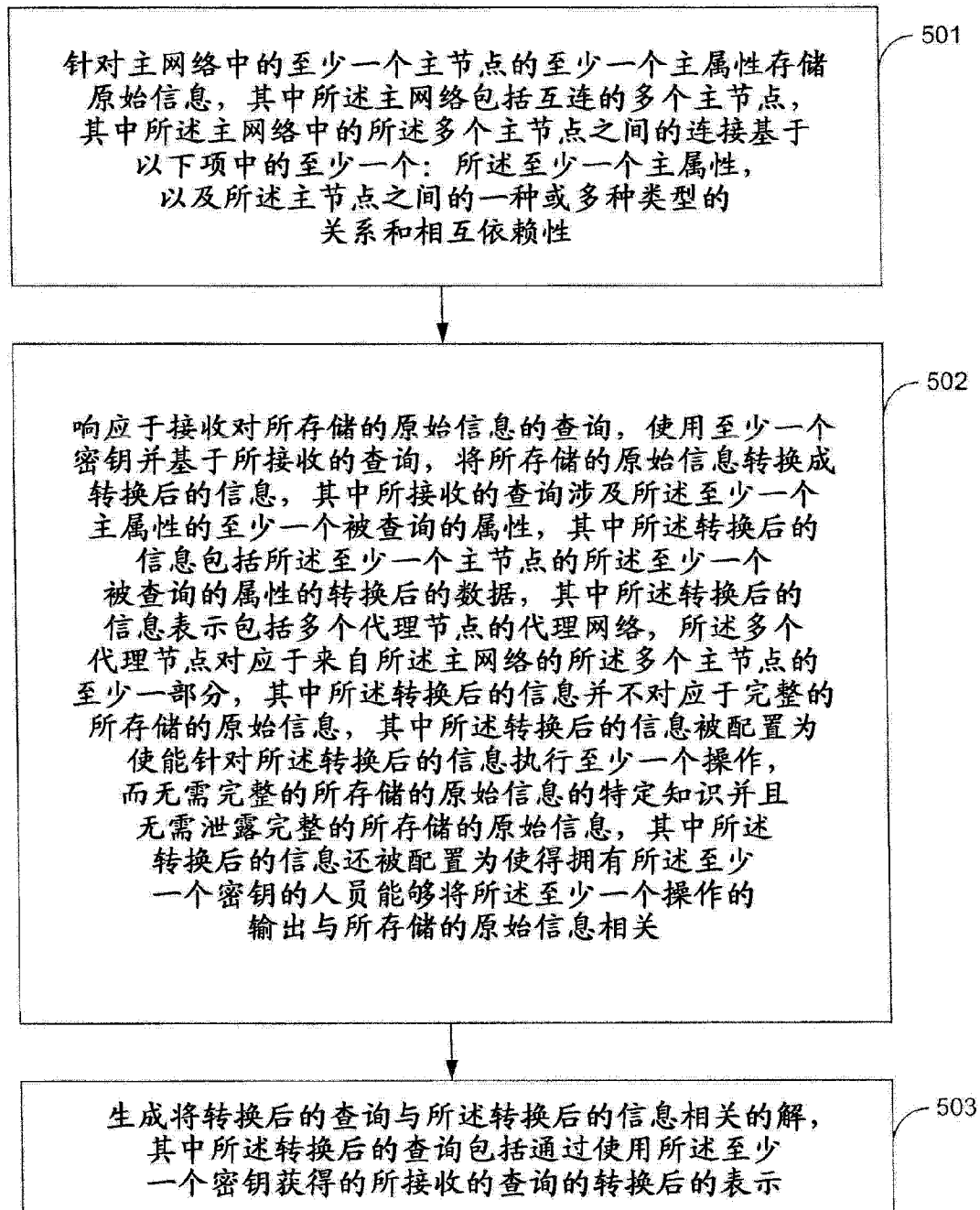


图 5