



(12)发明专利

(10)授权公告号 CN 102576431 B

(45)授权公告日 2017.05.03

(21)申请号 201080044651.0

(22)申请日 2010.09.29

(65)同一申请的已公布的文献号  
申请公布号 CN 102576431 A

(43)申请公布日 2012.07.11

(30)优先权数据  
09172309.8 2009.10.06 EP

(85)PCT国际申请进入国家阶段日  
2012.04.05

(86)PCT国际申请的申请数据  
PCT/IB2010/054390 2010.09.29

(87)PCT国际申请的公布数据  
W02011/042838 EN 2011.04.14

(73)专利权人 皇家飞利浦电子股份有限公司  
地址 荷兰艾恩德霍芬

(72)发明人 A·I·D·布库尔 R·弗多夫亚克

(74)专利代理机构 永新专利商标代理有限公司  
72002

代理人 张伟 王英

(51)Int.Cl.  
G06Q 10/06(2012.01)

(56)对比文件  
JP 特开2002-55857 A,2002.02.20,全文.  
JP 特开2003-44660 A,2003.02.14,全文.  
US 2006/0287890 A1,2006.12.21,摘要,说明书第14,84,94,102-136段、附图8.

审查员 邹予婷

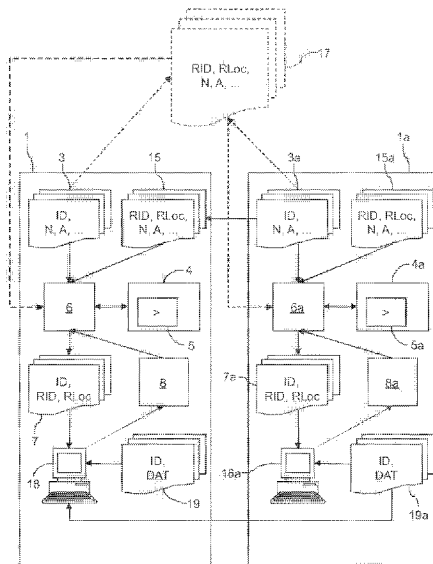
权利要求书2页 说明书11页 附图2页

(54)发明名称

存储于不同实体处的患者信息记录的自主链接

(57)摘要

公开了一种用于链接对应患者信息记录的系  
统。多个实体(1、1a)具有各自的包括患者信息  
记录(3、3a)的患者数据库。每个实体(1、1a)具有  
与之关联的患者识别算法(4、4a),该患者识别算  
法(4、4a)用于匹配同一患者在不同实体(1、1a)  
处的对应患者信息记录(3、3a)。链接子系统(6)  
维护多个实体(1、1a)中的第一实体(1)的一组链  
接(7)。链接子系统(6)被配置为将第一实体(1)  
的患者信息记录(3)与其他实体(1a)的对应患者  
信息记录(3a)进行链接。基于第一实体(1)的患  
者识别算法(4),当第一实体(1)的给定患者信息  
记录(3)与另一个实体(1a)的对应患者信息记录  
(3a)匹配时,建立链接(ID、RID、RLoc)。该链接提  
供了同一患者在不同实体(1、1a)处的本地分配  
的患者识别符(ID、RID)之间的关联。



1. 一种用于将存储在不同实体处的对应患者信息记录进行链接的系统,包括:

-多个实体(1、1a),其具有各自的患者数据库,所述患者数据库包括患者信息记录(3、3a),每个实体(1、1a)具有与之关联的患者识别算法(4、4a),所述患者识别算法用于匹配同一患者在不同实体(1、1a)处的对应患者信息记录(3、3a),其中所关联的所述患者识别算法(4、4a)由每个实体(1、1a)选择;以及

-链接子系统(6),其用于维护所述多个实体(1、1a)中的第一实体(1)的一组链接(7),所述链接子系统(6)被配置为基于所述第一实体(1)的所述患者识别算法(4)将所述第一实体(1)的患者信息记录(3)与其他实体(1a)的对应患者信息记录(3a)进行链接,其中当所述第一实体(1)的给定患者信息记录(3)与另一实体(1a)的对应患者信息记录(3a)匹配时,建立链接。

2. 根据权利要求1所述的系统,其中,所述链接提供同一患者在不同实体(1,1a)处的本地分配的患者识别符之间的关联。

3. 根据权利要求1所述的系统,包括另一链接子系统(6a),所述另一链接子系统(6a)基于所述多个实体(1、1a)中的第二实体(1a)的所述患者识别算法(4a)来维护所述第二实体(1a)的另一组链接(7a)。

4. 根据权利要求1所述的系统,其中,所述多个实体(1、1a)中的至少一个实体(1)被配置为存储其他实体(1)的信息(3a)的本地副本(15),所述本地副本(15)由该实体(1)的所述患者识别算法(4)使用以匹配患者信息记录(3、3a)。

5. 根据权利要求1所述的系统,还包括中央储存库(17),所述中央储存库(17)用于存储由所述多个实体(1,1a)的所述患者识别算法(4,4a)使用以匹配患者信息记录(3、3a)的所述信息(3、3a)的副本。

6. 根据权利要求1所述的系统,包括初始化子系统(8),所述初始化子系统(8)用于最初建立用于所述多个实体(1、1a)中的至少一个实体(1)的所述一组链接(7)、处理所述多个实体(1、1a)中的该实体(1)的所有可获得的患者信息记录(3),至少到以下程度:所有可获得的患者信息记录(3)潜在地在其他实体(1a)处具有对应的患者信息记录(3a)。

7. 根据权利要求1所述的系统,其中,所述链接子系统(6)被配置成周期性地更新所述一组链接(7)。

8. 根据权利要求1所述的系统,其中,所述链接子系统(6)被配置为在需要时或每次在所述多个实体(1,1a)之一处生成新患者识别符时更新所述一组链接(7)。

9. 根据权利要求1所述的系统,其中,所述患者识别算法(4)是基于匹配的患者人口统计学数据。

10. 根据权利要求9所述的系统,其中,所述人口统计学数据包括如下中的至少一项:姓名、地址、社会保险号、个人身份证号、税号、年龄、出生日期、性别。

11. 根据权利要求1所述的系统,其中,所述患者识别算法(4)包括阈值(5),所述阈值(5)对于与所述患者识别算法(4)关联的所述实体(1)而言是特定的。

12. 一种医疗工作站(18),其是根据权利要求1所述的系统的多个实体(1,1a)中的实体(1)的一部分,所述医疗工作站(18)包括显示器,所述显示器用于显示来自另一实体(1a)的患者信息记录(3a、19a)的患者信息,所述另一实体(1a)的所述患者信息记录(3a、19a)与包括所述工作站(18)的所述实体(1)的患者信息记录(3、19)相链接,根据权利要求1所述的系

统中的链接子系统(6)已经执行了该链接。

13. 一种链接不同实体(1、1a)处的对应患者信息记录(3、3a)的计算机实施的方法,每个实体(1、1a)具有各自的包括患者信息记录(3、3a)的患者数据库,每个实体(1、1a)具有与之关联的患者识别算法(4、4a),所述患者识别算法(4、4a)用于匹配同一患者在不同实体(1、1a)处的对应患者信息记录(3、3a),其中所关联的所述患者识别算法(4、4a)由每个实体(1、1a)选择,所述方法包括基于所述不同实体(1、1a)中的第一实体(1)的所述患者识别算法(4)来维护所述第一实体(1)的一组链接(7),所述一组链接(7)将所述第一实体(1)的患者信息记录(3)与其他实体(1a)的对应患者信息记录(3a)进行链接,其中当所述第一实体(1)的给定患者信息记录(3)与另一实体(1a)的对应患者信息记录(3a)匹配(204)时,建立(206)链接。

14. 一种链接不同实体(1、1a)处的对应患者信息记录(3、3a)的装置,每个实体(1、1a)具有各自的包括患者信息记录(3、3a)的患者数据库,每个实体(1、1a)具有与之关联的患者识别算法(4、4a),所述患者识别算法(4、4a)用于匹配同一患者在不同实体(1、1a)处的对应患者信息记录(3、3a),其中所关联的所述患者识别算法(4、4a)由每个实体(1、1a)选择,所述装置包括基于所述不同实体(1、1a)中的第一实体(1)的所述患者识别算法(4)来维护所述第一实体(1)的一组链接(7)的模块,所述一组链接(7)将所述第一实体(1)的患者信息记录(3)与其他实体(1a)的对应患者信息记录(3a)进行链接,其中当所述第一实体(1)的给定患者信息记录(3)与另一实体(1a)的对应患者信息记录(3a)匹配(204)时,建立(206)链接。

## 存储于不同实体处的患者信息记录的自主链接

### 技术领域

[0001] 本发明涉及链接在不同实体处存储的对应患者信息记录。

### 背景技术

[0002] 通常,患者能够从在地理上分散于多个站点的多个健康护理提供者接受护理。在每个站点,患者通常被给予不同的患者识别符。该患者识别符可以在健康护理提供者处本地使用。此外,单个患者的患者数据(例如医疗图像和其他相关医疗信息)在多个站点间传播并且被标记不同的本地患者识别符。为了能够检索到存储于别处的患者数据,这些患者识别符被协调(reconcile)并且将相应的患者记录链接在一起。

[0003] 随着时间推移,在诸如癌症之类的复杂疾病的背景下针对患者收集的平均数据量已经惊人地增大。例如,针对复发的癌症患者,大部分的病历可能与临床医生相关。在复发的癌症患者的情况下,相关的与癌症有关的健康事件(health episode)可能追溯到很多年前。共病(co-morbidity)常常也是相关的,这是因为在选择疗法时它们是非常有约束性的因素。例如,很多化疗药剂对心脏是有毒性的,并且为了选择正确的疗法,关于心脏病的先前信息可能是重要的。关于所有这些与健康相关的事件的信息存在于单个机构的系统中是非常不可能的。然而,诊疗患者的处置临床医生应当能够从与癌症相关的和与癌症不相关的患者记录中提取所有相关的先前健康事件,该患者记录可以包括很多事件并且跨越了数十年。

[0004] 一般地,通过患者主索引(MPI)将信息流传送到患者记录中和将信息流从患者记录中传出,当存在单元记录时,患者主索引向实体中的每个患者分配唯一的医疗记录编号(MRN)。在此,该单元记录包括由该实体维护(maintain)的实际患者数据。它可以是患者的电子健康记录、放射科信息系统(RIS)中的患者记录和由该实体维护的所有其他患者数据(例如研究、图像、实验室数据)。单元记录中的所有数据项都可以通过本地分配的患者识别符(例如MRN)进行链接。为了跨越分布式的数据源得到患者概况,各个机构中的本地识别符被协调。目前这是通过构建全行业的患者主索引(EMPI)来实现的,该全行业的患者主索引将作为行业一部分的医院中的所有识别符进行相互关联。通过集成这些源中的各个MPI来开发EMPI。通常,通过比较人口统计学属性来实现该集成以创建行业级的识别符,该人口统计学属性例如是姓/名、性别、出生日期、地址等。该集成很少基于在行业中的不同组织之间共享的单个识别符。大多数现有系统部署概率算法,该概率算法通常将固定的记录与用于匹配的若干候选者进行比较,针对每个候选者计算似然比(加权分数),其中将该似然比与所选择的接受阈值和拒绝阈值进行比较。使用结果来决定是否链接该记录。当不能自动做出决策时(所计算的似然落入两个阈值之间),有资格的人员需要在它们被接受(或被拒绝)之前复核或标记可能的匹配(失配(mismatch))。不确定的匹配的人工复核有助于使链接误差最小化,这是因为它们可能具有深远的影响,最终危及患者的健康。然而,提交大量记录来进行人工复核的成本非常高,并且可能使得整个方案不切实际。

[0005] Mohamed Yakout等人的论文“Efficient Private Record Linkage”(IEEE

International Conference on Data Engineering, 2009, 第1283-1286页) 公开了一种未利用第三方的用于私人记录链接的协议。该协议由两个阶段构成。在阶段1, 产生用于匹配的候选记录对。在阶段2, 完成计算在每个候选对之间的欧几里德距离 (Euclidean distance) 的任务。双方都参与欧几里德距离的计算, 而不泄露它们的各自记录的原始表达。

## 发明内容

[0006] 具有一种用于将不同实体处的对应患者信息记录进行链接的改进系统将是有益的。为了更好地解决这个问题, 本发明的第一方面提供了一种系统, 其包括:

[0007] -具有各自的患者数据库的多个实体, 该患者数据库包括患者信息记录, 每个实体具有与之关联的患者识别算法, 该患者识别算法用于匹配同一患者在不同实体处的对应患者信息记录; 以及

[0008] -用于维护 (maintain) 该多个实体中的第一实体的一组链接的链接子系统, 该链接子系统被配置为基于第一实体的患者识别算法来将该第一实体中的患者信息记录与其他实体中的对应患者信息记录进行链接, 当第一实体的给定患者信息记录与另一实体的对应患者信息记录匹配时, 建立链接。

[0009] 通过将单独的患者识别算法与不同实体关联, 这些实体不必对单个患者识别策略达成一致。因此, 改进了链接的灵活性。不同患者识别算法允许这些实体应用关于患者记录的链接的不同策略。这可以用于通过让这些实体选择它们自己的患者识别算法来保持这些实体的记录链接的自主性 (autonomous)。

[0010] 该链接可以被配置为提供同一患者在不同实体处的本地分配的患者识别符之间的关联。可以通过远程站点处的患者信息记录被赋予的患者识别符来唯一地确定在该远程站点处的患者信息记录。此外, 第一实体的患者信息记录也可以具有唯一的患者识别符。因此, 第一实体的任一患者记录与其他实体的任一患者记录之间的链接可以包括这两个患者信息记录的患者识别符。这是指定链接的适当方式。此外, 可以包括对其他实体的识别, 使得该患者识别符属于哪个实体是清楚的。

[0011] 该系统可以包括另一链接子系统, 该另一链接子系统用于基于多个实体中的第二实体的患者识别算法来维护该第二实体的另一组链接。第二实体的该组链接独立于由第一实体生成的那一组链接。关于在将它们的患者信息记录与其他实体的患者信息记录进行匹配时所使用的标准, 两个实体都是自主的。尽管如此, 如果两个实体都决定在同一对患者信息记录之间进行链接, 则这两个实体能够将患者信息记录交叉链接。

[0012] 多个实体中的一个实体 (例如第一实体) 可以被配置为存储其他实体中的信息的本地副本。具体而言, 该实体可以被配置为存储由该实体的患者识别算法使用以进行匹配患者信息记录的信息。这减少了必需的数据通信量, 这是因为可以利用本地存储的信息来应用患者识别算法。

[0013] 该系统可以包括中央储存库 (central repository), 该中央储存库用于存储由多个实体的患者识别算法使用以进行匹配患者信息记录的信息的副本。这减少了必需的存储量, 这是因为该信息的任何本地副本都不是必需的。

[0014] 该系统可以包括初始化子系统, 该初始化子系统用于最初建立用于多个实体中的

至少一个实体的一组链接、处理多个实体中的所述一个实体的基本上所有可获得的患者信息记录,至少到以下程度:该可获得的患者信息记录潜在地具有在其他实体处对应的患者信息记录。通过这一方式,链接是立即可获得的,这样可以避免关于是否存在针对患者的附加数据的混淆。此外,可以避免任何的不一致性。

[0015] 该链接子系统可以被配置为周期性地更新该组链接。通过这一方式,可以在非工作时间内安排更新。而且,保持该组链接是最新的。可替代地或附加地,该链接子系统可以被配置为在需要时或每次在实体之一处生成新患者识别符时更新该组链接。应需更新提供了灵活性。每次生成新患者识别符时的更新保持该链接高度最新。

[0016] 患者识别算法可以基于匹配的患者人口统计学数据。在患者之间区分时,患者人口统计学数据是非常有用的。人口统计学数据可以包括如下至少一项:姓名、地址、社会保险号、个人身份证号、税号、年龄、出生日期、性别。不排除其他数据类型。

[0017] 患者识别算法可以包括一个或多个阈值,该一个或多个阈值对于与该患者识别算法相关联的实体而言是特定的。通过不同地设置阈值,实体可以实施特殊的患者匹配策略。

[0018] 医疗工作站可以是所述系统的多个实体中的实体的一部分。医疗工作站可以包括显示器,该显示器用于显示来自另一实体的患者信息记录的患者信息,该另一实体的患者信息记录与包括该工作站的实体的患者信息记录相链接,链接子系统已经执行了该链接。

[0019] 本发明的另一方面提供了一种将不同实体处的对应患者信息记录进行链接的方法,每个实体具有各自的患者数据库,该患者数据库包括患者信息记录,每个实体具有与之关联的患者识别算法,该患者识别算法用于将同一患者在不同实体处的对应患者信息记录进行匹配,该方法包括基于多个实体中的第一实体的患者识别算法来维护第一实体的一组链接,该组链接将第一实体的患者信息记录与其他实体的对应患者信息记录进行链接,当第一实体中的给定患者信息记录与另一实体中的对应患者信息记录匹配时,建立链接。

[0020] 本发明的另一方面提供了一种计算机程序产品,该计算机程序产品包括用于使处理器系统执行所述方法的指令。

[0021] 本领域的技术人员应当认识到,可以以任何被认为有用的方式对本发明的上述实施例、实施方式和/或方面中的两个或多个进行组合。本领域技术人员可以基于本说明书对图像采集装置、工作站、系统和/或计算机程序产品进行变型和变化,该变型和变化对应于所描述的该系统的变型和变化。

## 附图说明

[0022] 利用下文描述的实施例,本发明的这些和其他方面变得清楚,并且将利用下文描述的实施例阐述本发明的这些和其他方面。在附图中:

[0023] 图1是用于将存储在不同实体处的对应患者信息记录进行链接的系统的图示;以及

[0024] 图2是用于对不同实体处的对应患者信息记录进行链接的方法的流程图。

## 具体实施方式

[0025] 患者从地理上分散在多个站点的多个健康护理提供者接受护理是越来越普遍的。这可能导致这样的状况:患者通常被给予多个患者识别符,原因在于每次患者在不同健康

护理供应者处注册时,他或她都会被分配新的患者识别符。而且,诸如患者的医疗图像和其他相关的医学信息之类的患者数据被收集并且存储在多个站点。然而,为了进行诊断或处置的目的,在一个健康护理提供者处可能需要在另一个健康护理提供者处收集和存储的数据。为了能够检索到存储于别处的患者数据,可以协调各个站点处的患者识别符,并且将各个站点处的患者识别符链接在一起。如果可以执行该协调和/或链接而不在不同的站点采用公共的患者识别符,则这是有益的。

[0026] 目前,通过使用身份匹配算法来将不同系统中属于同一患者的记录链接在一起,可以实现该协调和/或链接。当前使用的身份匹配方法(包括概率性的和基于规则的)都假定该匹配应当在所有站点都相同,并且在匹配结果上联盟范围的单一观点是充分的。然而,实际上,这可能并非如此。自主的健康护理组织可能同意共享患者数据,而未必在患者记录匹配方面达成同一观点。通过在数据一致性与完整性之间进行权衡来调谐匹配过程,并且在现有技术的系统中,不同站点不能选择不同的权衡(trade-off)。

[0027] 可以通过允许不同组织具有不同的身份匹配算法(也是关于属性权重和阈值)来扩展已知的身份协调方法,这可以获得不同的映射,满足源自分布式异构型环境的协调需求,在该分布式异构型环境中提供者可能希望在仍然能够共享患者数据的同时保持自主。

[0028] 图1示出了用于对存储在不同实体1、1a处的对应患者信息记录19、19a进行链接的系统。该系统包括多个实体1、1a。该实体可以包括多个经由网络连接的计算机。实体还可以包括服务器系统,在该服务器系统上已经安装了患者数据存储系统。因此,该系统可以包括存储模块、用于控制该系统的用户接口、诸如网络端口之类的数据通信端口、和/或通往广域网的连接。可以利用广域网来互连不同的实体。该广域网可以是因特网。在图中,显示了两个实体1、1a以示出可行的示例。然而,通过简单扩展本文所述的技术,任何数量的实体都是可以的。而且,在此通过示例的方式针对实体之一1或1a所描述的内容也可以被应用于系统中的其他实体,包括实体1和/或实体1a。

[0029] 实体1、1a可以具有各自的患者数据库。这些数据库包括患者信息记录3、3a。例如,实体1的数据库中的患者信息记录3可以包括患者识别符(ID)、患者姓名(N)、患者地址(A)和/或其他人口统计学信息(.....)。而且,可以在单独的记录19中存储其他医疗数据(DAT),该其它医疗数据(DAT)例如是图像和/或实验室结果和/或用药信息,利用患者识别符(ID)将该单独的记录19链接到人口统计学信息(N、A、.....)。可替代地,这些附加数据(DAT)也可以与该人口统计学信息(N、A、.....)一起存储在患者记录3中。

[0030] 实体1具有与之关联的患者识别算法4。可以在对同一患者在不同实体处的对应患者信息记录进行匹配的过程中使用该患者识别算法4。为此,患者识别算法4可以包括若干比较操作。而且,患者识别算法4可以包括一个或多个概率估计或似然估计,该概率估计或似然估计可以与两个患者记录匹配的概率相关。可以与一个或多个阈值5相关联地使用这些比较和/或估计的结果。这些阈值5例如可以确定用于创建两个患者记录3和3a之间的匹配所必需的最小概率或最小相似度。

[0031] 该系统还可以包括链接子系统6、6a。如图1所示,可以将单独的链接子系统6、6a实施为不同实体1、1a中的独立单元。然而,这不是限制性的。在可替代的示例中,可以将链接子系统实施为集中式的服务(未示出)。该集中式的服务可以被配置为考虑各自的患者识别算法4、4a来维护实体1、1a的一组链接7、7a。

[0032] 在图1的示例性配置中,实体1的链接子系统6被配置为维护实体1的一组链接7。该链接子系统6可以将实体1的患者信息记录3与实体1a的对应患者信息记录3a进行链接。同样地,链接子系统6可以将实体1的患者信息记录与任何其他实体(图中未示出)的对应患者信息记录进行链接。基于实体1的患者识别算法4,当实体1的给定患者信息记录3与另一个实体1a的对应患者信息记录3a匹配时,可以建立链接。例如,可以比较患者信息记录3、3a的对应字段(N、A、.....);如果它们匹配,则可以建立链接。如果因为患者信息记录3、3a稍有不同而使匹配不完美,则患者识别算法4及其可选阈值5可以表示是否建立链接的策略。因此,至少在大多数情况下,可以自动做出是否建立链接的决策。患者识别算法4还可以得出以下决策,即该识别算法4不能做出最终决策;在这样的情况下,人类操作者可以通过例如询问患者来评估数据并且可能获得校正或任何其他相关信息。将患者识别算法的这样的决策转发到链接子系统6,链接子系统6通过根据具体情况生成链接7或者将该患者记录放置在列表上以供人类操作者复核来相应地动作。

[0033] 链接7可以提供同一患者在不同实体处的本地分配的患者识别符之间的关联。例如,链接7可以包括来自自由实体1存储的本地患者信息记录3的识别符(ID)、以及来自其他实体1a的患者信息记录3a的“远程”识别符(RID)。在未保证远程识别符(RID)在多个实体当中是唯一的的情况下,不保证远程识别符(RID)唯一地识别远程患者信息记录3a。因此,链接7还可以包括存储了由远程识别符(RID)识别的相关患者信息记录3a的其他实体1a的实体识别符(RLoc)。该实体识别符(RLoc)可以包括例如其他实体1a的网络位置。然而,可以使用任何种类的识别符。为了方便起见这里使用了词语“远程”;该词语并不暗示实体之间的任何物理距离。即使识别符在多个实体当中是唯一的,实体识别符(RLoc)也可以有助于使患者数据检索更有效率。

[0034] 可替代地,链接7可以包括指向对应患者信息记录的超链接或统一资源定位符(URL)。在该情况下,可以在链接7中省略由远程实体1a分配的患者识别符。

[0035] 如上所述,可以提供实体1a的另一链接子系统6a以用于基于实体1a的患者识别算法4a来维护实体1a的一组链接7a。对于图1中未示出的任何其他实体,同样也是如此。还可以由中央服务器(未示出)提供该组链接7、7a,该服务器能够基于各自的患者识别算法4、4a来维护实体1、1a的该组链接7、7a。

[0036] 实体1可以包括用于存储至少一个其他实体1a中的患者信息记录3a的至少一部分的本地副本15的存储器。然后链接子系统6和/或患者识别算法4可以使用该本地副本,而不必在需要时从其他实体1a传输数据。因此,为此目的仅需要存储由患者识别算法4使用以匹配患者信息记录3、3a的患者信息字段。

[0037] 代替本地副本15、15a,该系统可以包括中央储存库17,该中央储存库17用于存储由多个实体的患者识别算法使用以匹配患者信息记录的信息的副本。已经由虚线符号表示该替代。也可以使用本地副本15、15a与中央储存库17的组合。

[0038] 可以提供初始化子系统8来用于最初建立用于实体1的一组链接。可以提供类似的初始化子系统8a来用于其他实体1a。初始化子系统8可以最初处理实体1的基本上所有可获得的患者信息记录3。然而,初始化子系统8可以跳过事先已知的在其他实体处没有对应患者信息记录的患者信息记录3。患者信息记录3的初始处理可以包括由链接子系统6结合患者识别算法4进行评估,以便建立起在本地患者信息记录3与存储于其他实体1a处的患者信



息记录3a之间的一组链接7。

[0039] 链接子系统6可以被配置为周期性地更新该组链接7,该周期例如是每天晚上或每小时(例如在工作时间内)或每周或每月。可替代地,或附加地,链接子系统6可以被配置为在需要时或每次在实体之一处生成新患者识别符时更新该组链接。为此,可以提供用户接口来提交该需要或创建新的患者信息记录,该需要或创建新的患者信息记录会触发新的患者识别符的生成。可以通过工作站18提供这样的用户接口。

[0040] 患者识别算法4可以基于将存储于本地患者信息记录3和远程患者信息记录3a中的患者人口统计学数据(N、A、.....)进行比较。例如,该算法可以包括姓名、地址、社会保险号、个人身份证号、税号、年龄、出生日期和/或性别的比较。该算法可以向不同类型的人口统计学信息分配不同的权重(即重要程度)。而且,该算法可以针对不同种类的字段不同地处理字段之间的失配。可以针对不同实体1、1a不同地设置患者识别算法的所有这些元素。例如,患者识别算法4、4a可以包括对于与患者识别算法4、4a相关联的实体1、1a而言是特定的阈值。

[0041] 如上所述,实体1可以包括医疗工作站18。此外,实体1可以包括具有不同功能的不同工作站。该医疗工作站18中的一个或多个可以包括显示器,该显示器用于显示来自与实体1的患者信息记录3相链接的另一实体的患者信息记录3a、19a的患者信息。该链接可以由链接子系统6生成的该组链接7中的链接之一来体现。例如,一旦已经调用了实体1的患者信息记录3,工作站18就可以显示存储在患者信息记录3中的信息的至少一部分。而且,工作站可以针对所显示的患者信息记录3的患者识别符(ID)查找任何链接7。如果找到了这样的链接7,则显示器可以指示:针对该患者可以从另一实体获得其他信息。可选地,显示器可以基于远程实体识别符或地址(RLoc)来指示该信息位于哪个实体。可以使用户能够请求附加的信息。在接收到该请求时,工作站18针对匹配远程患者识别符(RID)的患者信息记录3a、19a,触发向远程实体1a的查询。然后,工作站18可以显示如在对应远程患者识别符(RID)的患者信息记录19a中列出的、在远程实体1a处可获得的数据元(DAT)的列表。例如,患者信息记录19a可以包括先前患者检查(例如CT扫描)的一个或多个结果或药物处方。根据请求或自动地,工作站可以检索存储在远程患者记录19a中的一个或多个附加数据元(DAT)。由此检索的信息可以在工作站18上显示,或者可以接受进一步处理。

[0042] 应当理解的是,该系统中的某些元件的功能也可以被视为过程中的步骤。图2示出了用于将不同实体1、1a处的对应患者信息记录3、3a进行链接的过程。每个实体1、1a可以具有各自的包括患者信息记录3、3a的患者数据库。而且,每个实体1、1a可以具有与之关联的患者识别算法4、4a,该患者识别算法4、4a用于将同一患者在不同实体1、1a处的对应患者信息记录3、3a、19、19a进行匹配。该方法可以包括维护实体1的一组链接7的步骤。该组链接7可以将实体1的患者信息记录3、19与另一实体1a的患者信息记录3a、19a链接。该方法可以包括如下步骤:基于实体1的患者识别算法4,当实体1的给定患者信息记录3与实体1a的对应患者信息记录3a匹配时,建立链接。类似地,该方法可以包括维护用于实体1a和用于系统中的任何其他实体的这样一组链接。可以将该方法实施为适于在服务器系统上执行的计算机程序。

[0043] 图2示出了将如上所述的系统中的不同实体1、1a处的对应信息记录3、3a进行链接的过程的示例性流程。该过程可以在步骤201被触发,该过程例如由特定时间段的流逝或由

人工请求或由系统中新患者识别符的增加来触发。在该过程已被触发之后,可以逐个地处理患者信息记录3。在步骤202,选择实体1的这些患者信息记录3中的一个。在触发是由在系统中的新患者识别符的增加引起的情况下,仅需要选择和与该新患者识别符对应的患者信息记录3。在步骤203,可以选择系统中的其他实体1a中的一个。在步骤204,可以核实所选择的实体1a是否具有如下的患者信息记录3a:其信息字段与所选择的患者信息记录3中的信息相匹配。在步骤205,验证在步骤204是否已经发现匹配的患者信息记录3a。如果在步骤204中未发现匹配的患者信息记录3a,则该过程进行到步骤207。如果在步骤204中发现匹配的患者信息记录3a,则在步骤206,在所选择的患者信息记录3与匹配的患者信息记录3a之间建立链接。这可以通过存储包括在所选择的患者信息记录3中发现的患者识别符和在匹配的患者信息记录3a中发现的患者识别符的对来执行。也可以将用于识别所选择的其他实体1a的实体识别符与该患者识别符一起存储。在步骤207,可以验证针对当前选择的患者信息记录3是否已经选择了所有实体。如果没有选择所有实体,则可以通过进行到步骤203来选择下一个实体。如果针对当前选择的患者信息记录3已经选择了所有实体,则在步骤208,可以验证是否已经选择了所有的本地患者信息记录3。如果没有选择所有的本地患者信息记录3,则可以通过进行到步骤202来选择下一个患者信息记录3。如果选择了所有的本地患者信息记录3,则该过程可以在步骤209终止。

[0044] 本领域的技术人员应当认识到,可以将该过程和系统应用于由各种采集模态采集的多维图像数据,该多维图像数据例如是二维(2D)图像、三维(3D)图像或四维(4D)图像,该各种采集模态例如是,但不限于,标准X射线成像、计算机断层摄影(CT)、磁共振成像(MRI)、超声波(US)、正电子发射断层摄影(PET)、单光子发射计算机断层摄影(SPECT)和核医学(NM)。可以利用适当的患者信息记录19或19a将该图像数据作为患者数据(DAT)链接到本地实体1处或远程实体1a处的患者识别符(ID)。该患者数据(DAT)还可以包括例如实验室结果或药物处方或诊断。

[0045] 自动拒绝、接受或提交用于人工复核的匹配的数量可以同时取决于与比较期间的不同属性相关联的权重以及所选择的拒绝阈值和接受阈值,基于所述比较可以计算似然比。由于这个原因,可以执行以下评估和选择,即评估与所使用的属性对相关联的适当匹配权重以及选择用于自动匹配的适合的接受阈值和拒绝阈值,以改进身份匹配的实施。

[0046] 更高的拒绝阈值可能引起更大数量的潜在匹配被自动拒绝,这意味着被拒绝的真正匹配的百分比更高。作为结果,可能遗漏相关的患者信息。当该信息是重要的时候,例如,对特定物质的过敏症或接受化疗的癌症患者的先前心脏事件,遗漏该信息可能严重危及患者的健康。更低的接受阈值意味着更多的潜在匹配被自动接受。在实际生活中,这意味着可能将患者信息匹配到错误的患者,并且可能基于错误数据做出医疗决策。该第二类型的错误也可能危及患者生命。

[0047] 医院在处理数据时可以选择不同的权衡。当它们决定进行(重新进行)它们认为必要的测试时,或者当它们具有准确的现场患者问诊过程时,它们可以选择冒着遗漏一部分病史的风险。在该情况下,数据的一致性更重要。另一方面,其他健康护理组织可以决定访问完整的数据是极重要的,并且依赖临床医生筛选出不相关的数据或被认为是错误的

[0048] 在参与机构仅松散地联系并且在它们的管理上保持自主的环境中,参与机构可能

希望对它们的数据和与处理该数据相关联的质量过程保持完全控制。这例如意味着,如果医院A中的复核员工将两个患者记录链接在一起,另一个机构(比如说医院B)应当仍然能够对这两个记录进行其自己的复核并且可能(本地地)否决在A处做出的链接。如果医院B不能这样做,那么其会失去对其数据的自主性,并且原则上不能保证数据的一致性。作为自主的组织,B不需要考虑并且应用在A处采用的决策。对于自动匹配,这应当同样适用:当实体是不同机构或医院的计算机系统时,每个机构应当能够应用其自己的匹配过程,如对应图1的系统和图2的过程的情况。当一对机构,A和B,已经决定同样的匹配过程时,当在机构C中对数据进行匹配时不强制执行相同的匹配过程,这是因为C可以具有其自己的匹配过程。

[0049] 在下文中,将描述非限制性的用例。考虑两个机构A和B。如上所述,它们中的每个都可以应用其自己的匹配算法来对患者记录进行链接,这导致截然不同的链接组7、7a。可以将这样的一组链接称为患者登记簿(PR)。机构A和B使用相同的人口统计学属性来匹配记录,但是它们分别为这些属性分配不同的权重 $WA_1, \dots, WA_n$ 和 $WB_1, \dots, WB_n$ 。它们还使用不同的阈值来自动接受或拒绝匹配。这可能导致在它们的PR内容上的差异。例如,医院A已经将具有本地识别符115的记录匹配到站点B处的具有远程识别符371的记录,尽管站点B未考虑这一链接。站点B可能已经拒绝这一链接或将其放到人工复核例外列表上。类似地,当将医院C添加到联盟中时,A和B可以使用相同的算法或不同的算法来将它们的记录与C中的记录进行匹配。

[0050] 不同实体或机构可以使用不同的属性权重。例如,某一地区中常见的姓名应当接受与该地区中不常见姓名的匹配权重不同的匹配权重,并且跨地区的匹配可能意味着使用不同的权重。93岁的年龄在两个老年诊所之间具有的匹配权重应当比在两所大型大学医院间匹配数据时的匹配权重低。在下文中,描述非限制性的示例性用例,包含比较由两所医院A和B中的两个记录捕捉的人口统计学信息。医院A中存储的记录具有患者姓名“Jansen, Peter”、年龄“53”、地址“Addr1”和性别“男”,而医院B中存储的记录具有患者姓名“Janssen, Peter”(具有两个s)、年龄53、地址“Addr2”和性别“男”。两个姓名(Peter Jansen和Peter Janssen)在荷兰非常频繁地出现。取决于A和B,是否匹配这两个记录的决策可能是大不相同的。例如,荷兰的两所医院将使该姓名具有低权重,这是因为两个姓名都是现有的、相当常见的姓名。这可能引起匹配的自动拒绝,因为地址是不同的。如果在两个患者记录中年龄都是93,那么对于两家通用医院来说匹配的权重会增大,这是因为它们的患者不会有那么多是93岁的,但对于两家老年诊所来说就并非如此,这是因为它们会有更多年长的患者。如果A和B是法国的两所医院,不常见姓名的权重会高得多,并且该记录很可能会提交以进行人工复核。

[0051] 也可以是如下的情况:两家健康护理组织相对于彼此采取不同的方法,从而导致不同的匹配。例如,急救医院通常不太感兴趣于从患者的社区医院检索较老的无关数据,这是因为它们处理的是急救病例,并且重做大多数测试;此外,当患者被转送到急救医院时,他很可能与所有相关数据一起到达。它们的匹配过程将非常谨慎,目的在于小的例外列表(需要人工复核的),并且具有高的拒绝阈值。另一方面,社区医院将很可能需要访问关于患者的急救事件的信息,这是因为它们将要在将来诊疗该患者,所以它们倾向拥有更多,而非更少的,来自远程站点的信息。

[0052] 本文描述的系统可以不构建包含被认为对整个联盟有效的所有链接的单个全局

患者登记簿 (PR), 而是构建自己的PR, 并且针对每个机构来维护它们, 所述PR从而用作相对于该机构的本地真正事实 (local ground truth)。每个机构可以在本地或在公共联盟服务 (Common Federation Service) 处维护其自己的PR。本地维护PR的优点在于不需要广域通信来搜索PR并且即使当该站点变得与联盟暂时断开连接时该PR中的信息也是可获得的。当然, 可以实施机制以在诊治新患者时更新PR中的信息 (这可以每天、每周等等来实施)。

[0053] 本地站点可以选择针对联盟中的每个其他站点应用不同的匹配过程 (算法、权重和阈值)。匹配的结果可以包括将本地系统中的患者记录连接到另一站点处的患者记录的一组链接。这样的站点处的计算机系统可以是上述系统的实体。

[0054] 可以不构建公共的联盟范围的PR, 而是每个本地PR可以针对每个本地已知的患者的识别符将如在其他站点处的属于该患者的匹配记录的所有识别符与保存这些记录的站点的识别符一起进行存储。由于该匹配仅针对本地站点有效, 所以通过不同匹配过程得到链接这一事实不会影响数据的一致性。本地系统具有自主性来选择从其他站点检索什么数据, 而无需考虑那些站点处的匹配决策。站点A可以决定将PID=123的自有记录与站点B处的PID=456的记录进行链接, 但不与站点C处的PID=789的记录进行链接。站点B可能已经将PID=456与PID=789链接的事实对A处的决策没有影响。

[0055] 本地PR中存储的数据总量可能大于将已经存储于单个全局PR中的数据量, 但每个本地PR可以显著小于该全局PR, 这是因为它仅需要存储与本地站点处已知的患者相关的数据。目前, 平均低于20%的患者在多于一个站点处具有数据, 因此存储在多于一个站点处的PR中的记录的百分比可能也低于20%。该百分比仅仅是一个示例。它可能根据情况而有所不同, 并且可能随着时间而波动。

[0056] 此外, 每个站点可以维护这样的数据库, 该数据库包含供联盟中所有远程站点 (或至少一些远程站点, 例如最重要的那些远程站点) 进行匹配所使用的人口统计学数据。如果不是在每个站点构建它们, 而是在公共站点 (其可以称为联盟服务站点) 维护所有这些不同的PR, 那么可以准备单个人口统计学数据库, 该单个人口统计学数据库包含所有站点的所有相关人口统计学资料 (即用于匹配的那些)。可替代地, 实体可以独立地查询其他实体的数据库。患者登记簿 (PR) 可以记录将本地患者识别符与远程患者识别符相链接所必需的数据。

[0057] 在下文中, 假设每个站点维护其自己的PR。然而, 技术人员可以使下文所述内容适用于中央PR的情况并且适用于逐个地查询其他实体的情况。当在站点A处增加先前未知的患者的新记录时, 可以在本地PR中创建新记录, 一开始仅保存新的本地PID, 并且在需要时还保存该患者的本地已知的人口统计学资料。接下来, 基于该新患者的人口统计学资料, 可以从每个其他站点采集所有可能相关的人口统计学数据, 并且可以运行对应的匹配算法。在该PR中存储本地人口统计学数据时, 当前的站点可以直接将其新患者的人口统计学资料与每个远程PR中的人口统计学数据进行匹配, 并且无需查询远程系统。作为替代, 本地站点可以向远程站点发送人口统计学信息, 并且要求该站点提供与其PR中的人口统计学数据的所有可能匹配。接下来, 本地站点可以对从远程站点返回的数据执行匹配算法, 并且在本地站点处的新患者的PR条目中存储被认为匹配的所有远程识别符和保存该远程识别符的远程站点的识别符。

[0058] 由于匹配的链接对每个站点是特定的, 所以当站点A处引入新患者记录并且执行

匹配时,远程站点将不会自动意识到该新条目。而且,本地创建的链接可能对它们而言是不相关的,并且可能不被采用。然而,站点可能需要知道在联盟中的其他站点处输入的新记录,这是因为那些记录可能与自己的记录匹配并且可能提供相关信息。这可以通过在创建新患者记录并且该新患者记录被分配新的识别符时让每个站点A通知所有其他站点来实现。这可以在站点A执行其与每个远程站点中的数据的身位匹配时有效地完成。当站点A向站点B发出用于新患者的人口统计学资料并且请求潜在匹配的记录的人口统计学资料时,也可以发送新记录的本地分配的识别符。通过该方式,基于接收的人口统计学资料,站点B还能够将站点A处的新记录与自己的PR(在站点B处的)中的记录进行匹配,并且当它发现与自己记录(其PR中的条目)的匹配时,站点B向本地条目增加从A接收的患者识别符和对A的引用。针对输入的每个新记录,在本地站点和联盟中所有其他站点之间可以存在单个通信步骤,在该步骤期间人口统计学数据和/或患者识别符被交换。之后,可以将该新记录与每个站点处的现有记录进行匹配。

[0059] 通过该方式,当新记录被增加到本地站点时,所有PR可以保持最新,即使匹配过程对于每个站点而言是本地的情况也是如此。

[0060] 下文描述实现PR构建和更新的过程。假定存在N个站点的联盟(其中站点对应于如上所述系统的实体)。

[0061] 1. 根据以下步骤,每个站点可以全部一次性(类似大爆炸(big-bang))地或者更多是逐步地构建其自己的初始PR:

[0062] a. 检索并存储在其他N-1个站点处的患者的人口统计学数据(例如姓名、年龄、性别、地址)和ID。该信息也可能在正常操作期间被使用,因此其可以被保存在人口统计学数据库中。

[0063] b. 使用自己的用于匹配的算法、权重和阈值(所有这些可能对于每对站点都是不同的)来将自己站点处的患者的人口统计学资料与远程站点处的患者的人口统计学资料进行匹配。

[0064] c. 在自己的PR中存储被识别为自动匹配的记录对,该记录对例如是(本地患者ID、远程站点ID、远程患者ID)的三元组(triplet)。也可以存储诸如时间戳之类的其他相关信息。

[0065] d. 利用不能被自动匹配或被自动拒绝的记录对(即落入两个阈值之间)构建用于人工复核的例外列表。

[0066] e. 使操作者能够复核该例外列表并且向PR添加被断言为匹配的记录对。

[0067] 2. 每个站点可以在正常操作期间执行以下步骤:

[0068] a. 当在本地站点注册新患者时,进行用于他或她的人口统计学数据资料与人口统计学数据库中的记录的匹配。

[0069] b. 在发现匹配时,向PR添加三元组(本地患者ID、远程站点ID、远程患者ID)。

[0070] c. 当发现潜在的匹配(在两个阈值之间)时,向用于人工复核的例外列表添加三元组。

[0071] d. 周期性地(例如一天一次,一周一次):

[0072] i. 检索在其他站点处登记的患者的相关人口统计学资料。将自己的患者人口统计学资料与新检索的远程患者的人口统计学资料进行匹配。

[0073] ii. 当发现匹配时,向PR添加三元组(本地患者ID、远程站点ID、远程患者ID)。

[0074] iii. 当发现潜在的匹配(在两个阈值之间)时,向用于人工复核的例外列表添加三元组。

[0075] iv. 在人口统计学数据库中存储新检索的人口统计学资料。

[0076] 应当认识到,本发明也应用于被配置为实施本发明的计算机程序,尤其是载体上或载体中的计算机程序。该程序可以是源代码、目标代码、居于源代码和目标代码之间的代码(例如部分编译的形式)的形式,或者可以是适于在实施根据本发明的方法时使用的其他任何形式。还应当认识到,该程序可以具有很多不同的架构设计。例如,实施根据本发明的方法或系统的功能的程序代码可以细分成一个或多个子例程。对于技术人员而言,用于在这些子例程之间分配功能的很多不同方式将是显而易见的。可以将该子例程一起存储在一个可执行文件中以形成独立的程序。该可执行文件可以包括计算机可执行指令,例如,处理器指令和/或解释器指令(例如Java解释器指令)。可替代地,一个或多个或所有子例程可以被存储在至少一个外部库文件中并且可以(例如在运行时)与主程序静态或动态地链接。该主程序包含至少一个对至少一个子例程的调用。该子例程还可以包括对彼此的函数调用。涉及计算机程序产品的实施例包括与本文所述的方法中的至少一个的每个处理步骤相对应的计算机可执行指令。可以将这些指令细分成子例程和/或存储在静态或动态地链接的一个或多个文件中。涉及计算机程序产品的另一实施例包括与本文所述的系统和/或产品中的至少一个的每个模块相对应的计算机可执行指令。可以将这些指令细分成子例程和/或存储在静态或动态地链接的一个或多个文件中。

[0077] 计算机程序的载体可以是能够承载程序的任何实体或装置。例如,该载体可以包括诸如ROM(例如是CD ROM或半导体ROM)之类的存储介质或者诸如软盘或硬盘之类的磁记录介质。此外,该载体可以是诸如电信号或光信号之类的可传输载体,该电信号或光信号可以由这样的电缆或其他装置或模块构成。可替代地,该载体可以是其中嵌入了程序的集成电路,该集成电路被配置为执行相关方法或在执行相关方法时被使用。

[0078] 应当注意的是,上述实施例是例示而非限制本发明,并且本领域的技术人员将能够设计很多替代实施例而不脱离所附权利要求的范围。在权利要求中,置于括号之间的任何附图标记都不应被解释为限制权利要求。动词“包括”及其词形变化的使用并不排除在权利要求中还存在除了所列举的那些元件或步骤之外的元件或步骤。元件前的冠词“一个”或“一种”不排除存在多个这样的元件。可以利用包括若干不同元件的硬件并且利用适当程序控制的计算机来实施本发明。在列举了若干模块的装置权利要求中,可以由同一件硬件实现这些模块中的若干个。在相互不同的从属权利要求中记载某些手段的事实并不表示这些手段的组合不能被使用以获得优点。

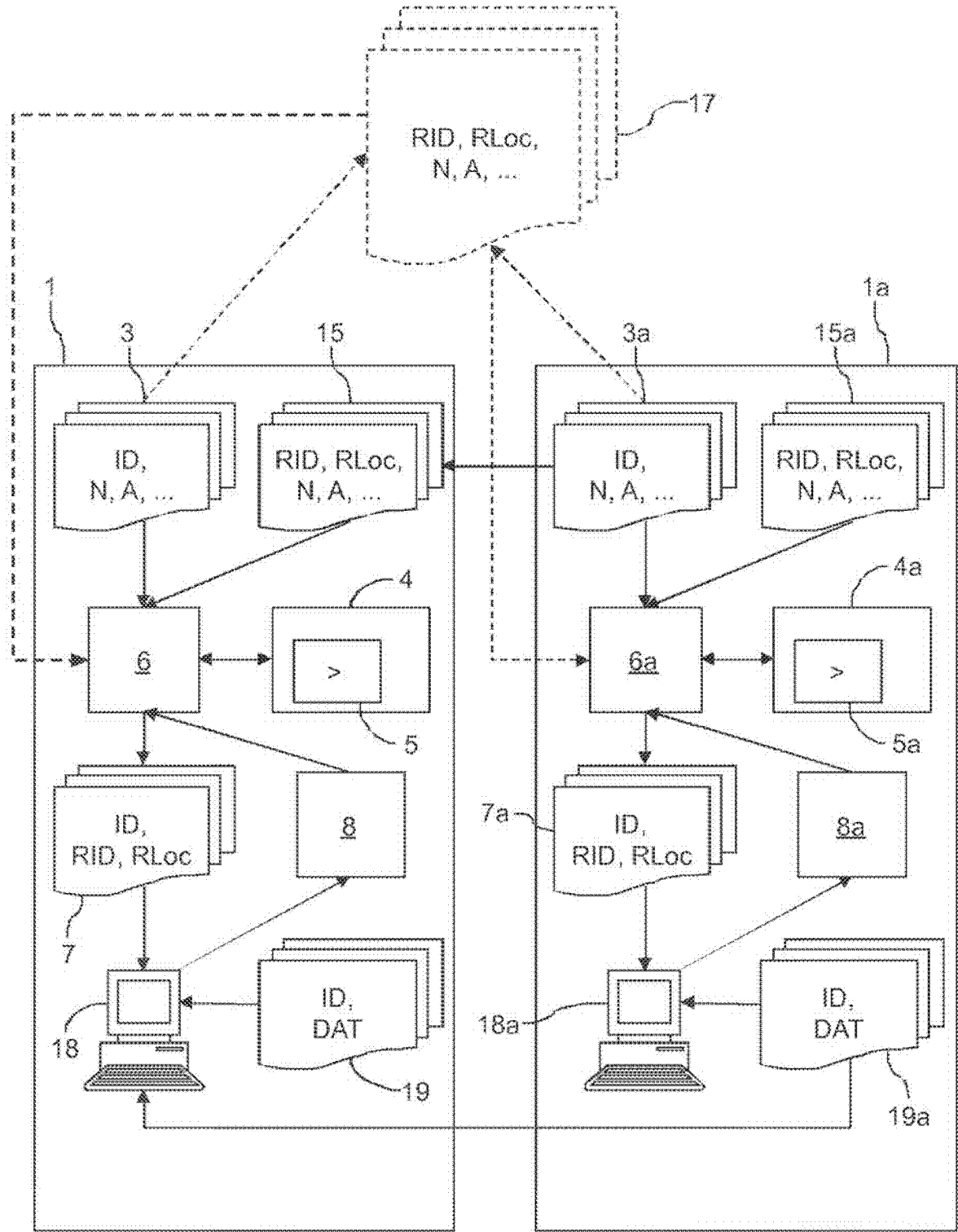


图1

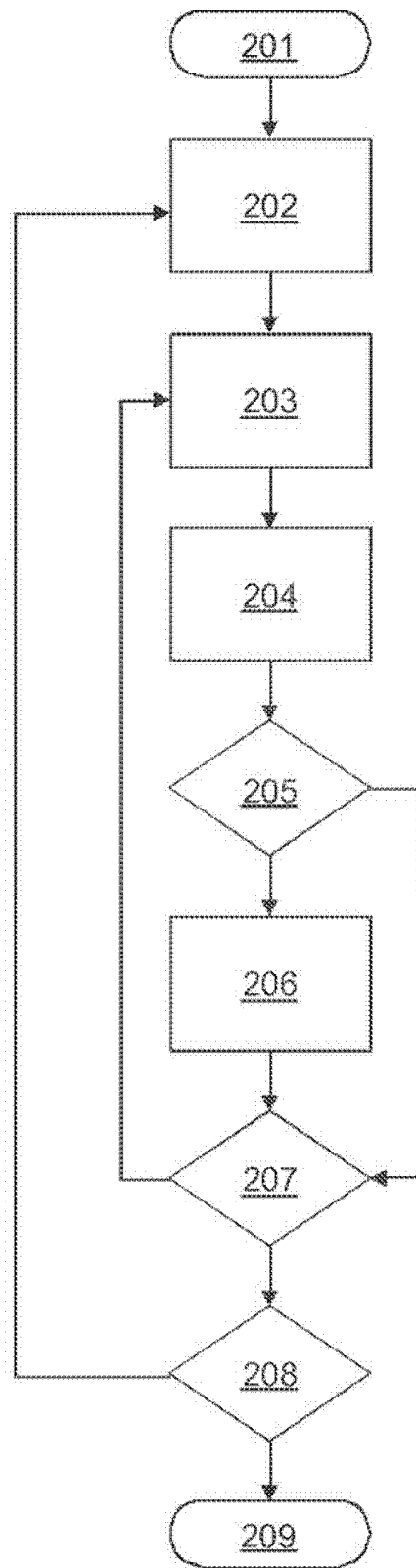


图2