



(12)发明专利

(10)授权公告号 CN 106133716 B

(45)授权公告日 2019.03.29

(21)申请号 201580013972.7

(22)申请日 2015.01.14

(65)同一申请的已公布的文献号
申请公布号 CN 106133716 A

(43)申请公布日 2016.11.16

(30)优先权数据
61/964,800 2014.01.14 US

(85)PCT国际申请进入国家阶段日
2016.09.13

(86)PCT国际申请的申请数据
PCT/US2015/011477 2015.01.14

(87)PCT国际申请的公布数据
W02015/109030 EN 2015.07.23

(73)专利权人 阿雅斯迪公司
地址 美国加利福尼亚州

(72)发明人 P·Y·鲁美 E·卡达格
J·格瑞韩恩 J·刘易斯
H·塞克斯顿

(74)专利代理机构 北京纪凯知识产权代理有限公司 11245

代理人 徐东升 王爽

(51)Int.Cl.
G16H 10/60(2018.01)
G16H 50/70(2018.01)
G16H 50/20(2018.01)

(56)对比文件
CN 102687138 A,2012.09.19,
CN 103384887 A,2013.11.06,
US 6526405 B1,2003.02.25,
US 20120232930 A1,2012.09.13,

审查员 李宝

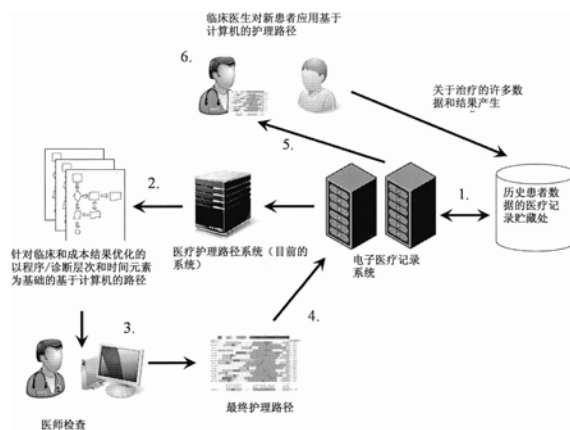
权利要求书3页 说明书26页 附图18页

(54)发明名称

共有序列识别

(57)摘要

示例方法包括接收片段的历史信息;从历史信息构造事件集;以一般标签和综合标签将每个事件归类;通过使用一般标签和综合标签获知对事件的事件度量以执行维数降低从而将向量与每个事件相关联并确定每两个向量之间的角度;使用在每对事件集之间的距离确定事件集度量;导出对片段的序列度量,该序列度量获得两个片段之间的优选匹配;导出对片段的子序列度量,该子序列度量是对每个片段的子序列的事件集度量的函数;基于距离将片段群集成子群;对于至少一个子群通过寻找优选事件序列和子群的片段来生成共有序列;以及生成指示共有序列的报告。



1. 一种用于共有序列识别的方法,其包括:
 - 接收片段的历史信息,每个片段包括在一时段内取得的至少一个事件序列;
 - 从所述历史信息中构造事件集,所述事件集中的每个包括至少一个事件序列;
 - 用一般事件类别标签和综合事件类别标签将来自所述历史信息的每个事件归类;
 - 通过使用所述一般事件类别标签和综合事件类别标签获知对事件的事件度量以执行维数降低,从而将向量与每个事件相关联并且确定每两个向量之间的角度;
 - 使用在每对事件集之间的距离利用所述事件度量来确定事件集度量;
 - 导出对所述片段的序列度量以计算片段之间的距离,所述序列度量相对于描述所述事件集度量的加权的表格MinCost(,)获得两个片段之间的优选匹配;
 - 导出对所述片段的子序列度量以计算片段之间的距离,所述子序列度量是对每个片段的子序列的所述事件集度量的函数;
 - 基于使用所述序列度量和所述子序列度量而获得的距离将片段群集成子群;
 - 对于至少一个子群,通过关于所述序列度量以及事件的优选序列和所述子群的片段之间的所述子序列度量的函数寻找所述优选序列来生成共有序列;以及
 - 生成指示所述共有序列的报告。
2. 根据权利要求1所述的方法,其中用一般事件类别标签将每个事件归类包括检索在所述历史信息中的本体,并且使用所述本体来确定所述一般事件类别标签。
3. 根据权利要求1所述的方法,其中在两个片段之间的所述优选匹配为最优匹配。
4. 根据权利要求1所述的方法,其中所述序列度量为护理路径度量即CP度量。
5. 根据权利要求1所述的方法,其中所述子序列度量为事件集护理路径度量即ESCP度量。
6. 根据权利要求1所述的方法,其中所述事件集度量的函数为加权和。
7. 根据权利要求1所述的方法,其中每个子序列相对于相关片段中的一个或多个锚点被定义。
8. 根据权利要求1所述的方法,其中每个事件集包括多个事件。
9. 根据权利要求8所述的方法,其中所述多个事件中的至少一个的多个行为的顺序不是可区别的。
10. 根据权利要求1所述的方法,其中根据所述历史信息构造事件集包括构造以不超过预定时段分离的事件集。
11. 根据权利要求1所述的方法,进一步包括过滤所述事件以移除不经常发生的事件。
12. 一种用于共有序列识别的系统,其包括:
 - 事件集构造模块,其被配置为接收片段的历史信息,每个片段包括在一时段内取得的至少一个事件序列,并且被配置为根据所述历史信息构造事件集,所述事件集中的每个包括至少一个事件序列;
 - 归类模块,其被配置为用一般事件类别标签和综合事件类别标签根据所述历史信息将每个事件归类;
 - 度量构造模块,其被配置为通过使用所述一般事件类别标签和综合事件类别标签获知对所述事件的事件度量以执行维数降低,从而将向量与每个事件相关联并确定每两个向量之间的角度;

距离模块,其被配置为使用在每对事件集之间的距离利用所述事件度量来确定事件集度量;

片段度量组装模块,其被配置为导出对所述片段的序列度量以计算片段之间的距离,所述序列度量相对于描述所述事件集度量的加权的表格MinCost(,)获得两个片段之间的优选匹配,并且导出对所述片段的子序列度量以计算片段之间的距离,所述子序列度量是对每个片段的子序列的所述事件集度量的函数;

自动分组模块,其被配置为基于使用所述序列度量和所述子序列度量而获得的距离将片段分组成子群;以及

共有模块,其被配置为对于至少一个子群,通过关于所述序列度量以及事件的优选序列和所述子群的片段之间的所述子序列度量的函数寻找所述优选序列来生成共有序列,并且生成指示所述共有序列的报告。

13. 根据权利要求12所述的系统,其中所述归类模块被配置为用一般事件类别标签将每个事件归类包括所述归类模块被配置为检索所述历史信息中的本体,并使用所述本体来确定所述一般事件类别标签。

14. 根据权利要求12所述的系统,其中在两个片段之间的所述优选匹配为最优匹配。

15. 根据权利要求12所述的系统,其中所述序列度量为护理路径度量即CP度量。

16. 根据权利要求12所述的系统,其中所述子序列度量为事件集护理路径度量即ESCP度量。

17. 根据权利要求12所述的系统,其中所述事件集度量的函数为加权和。

18. 根据权利要求12所述的系统,其中每个子序列相对于相关片段中的一个或多个锚点被定义。

19. 根据权利要求12所述的系统,其中每个事件集包括多个事件。

20. 根据权利要求19所述的系统,其中所述多个事件中的至少一个的多个行为的顺序不是可区别的。

21. 根据权利要求12所述的系统,其中所述事件集构造模块被配置为根据所述历史信息构造事件集包括构造以不超过预定时段分离的多个事件集。

22. 根据权利要求12所述的系统,进一步包括被配置为过滤所述事件以移除不经常发生的事件的过滤模块。

23. 一种包括可执行指令的计算机可读媒体,所述可执行指令可由处理器执行以执行方法,所述方法包括:

接收片段的历史信息,每个片段包括在一时段内取得的至少一个事件序列;

根据所述历史信息构造事件集,所述事件集中的每个包括至少一个事件序列;

用一般事件类别标签和综合事件类别标签根据所述历史信息将每个事件归类;

通过使用所述一般事件类别标签和综合事件类别标签获知对所述事件的事件度量以执行维数降低,从而将向量与每个事件相关联并且确定每两个向量之间的角度;

使用在每对事件集之间的距离利用所述事件度量来确定事件集度量;

导出对所述片段的序列度量以计算片段之间的距离,所述序列度量相对于描述所述事件集度量的加权的表格MinCost(,)获得两个片段之间的优选匹配;

导出对所述片段的子序列度量以计算片段之间的距离,所述子序列度量是对每个片段

的子序列的所述事件集度量的函数,其中所述子序列相对于锚点被定义;

基于使用所述序列度量和所述子序列度量而获得的距离将片段分组成子群;

对于至少一个子群,通过关于所述序列度量以及事件的优选序列和所述子群的片段之间的所述子序列度量的函数寻找所述优选序列来生成共有序列;以及

生成指示所述共有序列的报告。

共有序列识别

技术领域

[0001] 本公开总体涉及从历史信息中确定共有序列 (consensus sequence), 并且更具体地涉及用于在应用程序被配置成根据一些所关注的的质量度量从过去的行为序列记录中识别共有时间模式的不同应用程序状态时, 从应用程序复制内容的技术。

背景技术

[0002] 随着计算机的存储容量和处理能力两者的增长, 数据收集已蓬勃发展。不幸的是, 随着所存储数据的量和复杂性增长, 从所存储数据中得出有意义信息的能力受到限制。

[0003] 数据科学家传统上收集先前存储的数据并试图通过基于查询的途径 (由此查询大量数据) 来得出有意义的信息。不幸的是, 基于查询的途径需要数据科学家猜测所存储数据中的关系并然后手动做出有意义的查询。此途径具有有限价值, 特别是在数据的量和复杂性扩展时。进一步地, 查询形成中的错误可导致误导性的结果。

发明内容

[0004] 示例方法包括接收片段的历史信息、从历史信息构造事件集、以一般标签和综合标签将每个事件归类、通过使用一般标签和综合标签获知对事件的事件度量以执行维数降低从而将向量与每个事件相关联并确定每两个向量之间的角度; 使用在每对事件集之间的距离确定事件集度量; 导出对片段的序列度量, 该序列度量获得两个片段之间的优选匹配; 导出对片段的子序列度量, 该子序列度量是对每个片段的子序列的事件集度量的函数; 基于距离将片段群集成子群; 对于至少一个子群通过寻找优选事件序列和子群的片段来生成共有序列; 以及生成指示共有序列的报告。

[0005] 用一般事件类别标签将每个事件归类可包括检索历史信息中的本体, 并使用该本体以确定一般事件类别标签。在一些实施例中, 在两个片段之间的优选匹配为最优匹配。序列度量可为CP度量。子序列度量可为ESCP度量。

[0006] 在各种实施例中, 事件集度量的函数为加权和。每个子序列可相对于相关片段中的一个或多个锚点定义。在一些实施例中, 每个事件包括多个事件。事件中的至少一个事件的多个行为的顺序是不可区分的。从历史信息中构造事件集可包括构造以不超过预定时段分离的多个事件集。该方法可进一步包括过滤事件以移除不经常发生的事件。

[0007] 示例系统可包括事件集构造模块、归类模块、归类模块、度量构造模块、距离模块、片段度量组装模块、自动分组模块, 以及共有模块。事件集构造模块可被配置为接收片段的历史信息 (每个片段包括在一时段内取得的至少一个事件序列), 并从历史信息中构造事件集, 事件集中的每个包括至少一个事件序列。归类模块可被配置为用一般事件类别标签和综合事件类别标签将来自历史信息的每个事件归类。度量构造模块可被配置为通过使用一般标签和综合标签获知对事件的事件度量以执行维数降低, 从而将向量与每个事件相关联并确定每两个向量之间的角度。距离模块可被配置为使用在每对事件集之间的距离使用事件度量来确定事件集度量。片段度量组装模块可被配置为导出对片段的序列度量以计算片

段之间的距离,该序列度量获得相对于描述事件集度量的加权的成本函数的两个片段之间的优选匹配,并导出对片段的子序列度量以计算片段之间的距离,该子序列度量是对每个片段的子序列的事件集度量的函数。自动分组模块可被配置为基于使用序列度量和子序列度量而获得的距离将片段群集成子群。共有模块被配置为对于至少一个子群,通过寻找相对于在优选序列之间的序列度量和子序列度量的函数的优选事件序列与子群的片段来生成共有序列,以及生成指示共有序列的报告。

[0008] 示例计算机可读媒体可包括可执行指令。可执行指令可由处理器执行以执行方法。该方法可包括接收片段的历史信息、从历史信息构造事件集、以一般标签和综合标签将每个事件归类、通过使用一般标签和综合标签获知对事件的事件度量以执行维数降低从而将向量与每个事件相关联并确定每两个向量之间的角度;使用在每对事件集之间的距离确定事件集度量;导出对片段的序列度量,该序列度量获得两个片段之间的优选匹配;导出对片段的子序列度量,该子序列度量是对每个片段的子序列的事件集度量的函数;基于距离将片段群集成子群;对于至少一个子群通过寻找优选事件序列和子群的片段来生成共有序列,以及生成指示共有序列的报告。

附图说明

[0009] 图1是其中可实践实施例的示例环境。

[0010] 图2是在一些实施例中的示例共有系统的框图。

[0011] 图3是在一些实施例中的用于生成共有时间模式的方法的流程图。

[0012] 图4是在一些实施例中的在接收的数据中的事件的示例图解。

[0013] 图5是在一些实施例中的中等维数的欧几里得空间中的单位向量的示例图解。

[0014] 图6是一些实施例中的与一般事件配对的不匹配对的示例图解。

[0015] 图7a和图7b是一些实施例中的在网格中的配对和路径的示例图解。

[0016] 图8a是一些实施例中的带有簇的邻域图的示例图解。

[0017] 图8b是一些实施例中的在簇核(cluster core)中的已比对事件集的示例图解。

[0018] 图8c是一些实施例中的簇核的共有片段的示例图解。

[0019] 图9是在一些实施例中的示例事件集,其中一个锚点事件示出编号的群。

[0020] 图10是在一些实施例中的医疗护理示例的流程图。

[0021] 图11是在一些实施例中的医疗护理的一些实施例的应用的另一示例。

[0022] 图12是在一些实施例中的示例护理路径共有界面。

[0023] 图13是在一些实施例中的示例护理共有编辑界面。

[0024] 图14a-图14d描绘了在一些实施例中的基于评分来确定分区以用于自动分组的示例。

[0025] 图15描绘了在一些实施例中的示例自动分组模块。

[0026] 图16是在一些实施例中的用于自动分组的示例流程图。

[0027] 图17是用于描述在一些实施例中的关于图16的步骤的示例林状图。

[0028] 图18是示例性数字装置的框图。

[0029] 在各图中相似参考符号指示相似要素。

具体实施方式

[0030] 在本文中描述从过去的行为序列记录中识别有益的共有时间模式的示例技术。通过根据所关注的一些质量度量选择多组大体成功的过去的行为序列,所产生的共有序列可形成用于有益未来行为的模板。

[0031] 在各种实施例中,在有限时段(“片段”)取得的行为序列包括一个或多个“事件”的多个事件集的序列。事件包括一些种类的复杂行为,该种类包括但不限于给出患者药物、钻油井或提出买股票的命令。通过建立片段之间相似性的测量,我们可将它们划分成簇并为每个簇导出事件的共有(例如平均)序列,这将形成共有序列。

[0032] 各种实施例包括构造适当事件集序列、对事件的一个或多个度量、对事件集的一个或多个度量,以及对片段的一个或多个度量。随后可将用于从片段的簇中导出共有序列的程序公式化。本文中描述了用于预测先前没有观察的片段的结果的示例系统和方法。

[0033] 尽管本文所述的系统和方法表明了医疗护理中的应用(例如,构造护理路径,其为护理提供者和患者之间交互的序列),以及在同一域中实践的减少,但应认识到相同和/或相似技术可应用于任何数量的领域(例如,油和天然气、金融、生物技术等)。

[0034] 图1是其中可实践实施例的示例环境100。在各种实施例中,用于共有序列生成和/或结果预测的数据分析可(例如,通过在本地数字装置上的软件和/或硬件)被本地执行、(例如,经由云计算)跨网络执行或两者的组合。在本地执行全部或一些活动和在网络上执行全部或一些活动之间具有许多优点。尽管关于医疗系统和装置描述图1,但如本文所述,应认识到本文所述的实施例可用于任何数量的领域。

[0035] 环境100包括通过通信网络110通信的医疗记录贮藏处102、电子医疗记录系统104、共有系统106和医疗专业人员装置108。环境100描绘了其中跨通信网络110执行功能的实施例。在该示例中,医师或患者可通过经由通信网络110在数据存储服务器中存储数据来利用云计算。共有系统106可执行分析和生成共有序列报告和/或基于共有序列进行预测。

[0036] 医疗记录贮藏处102、电子医疗记录系统104、共有系统106和医疗专业人员装置108可以是或包括任何数量的数字装置。数字装置是包括存储器和处理器的任何装置。数字装置在图18中进一步被描述。系统可以是任何数量的数字装置。

[0037] 在各种实施例中,医疗记录贮藏处102可包括任何量的历史信息(例如,历史患者数据)。医疗记录贮藏处102可包括例如电子医疗记录(EMR)数据库。在一个示例中,医疗记录贮藏处102可从任何数量的医疗专业人员和相关专业人员收集关于任何数量患者的信息(例如,来自保险公司的信息)。例如,医疗记录贮藏处102可包括指示关于任何数量患者的治疗、实验室、测试、操作、药物等的记录。

[0038] 在各种实施例中,医疗记录贮藏处102可包括关于在多个医疗设施的患者和/或与任何数量的医疗工作人员相关联的任何量的信息。在一些实施例中,医疗记录贮藏处102的历史信息可包括关于任何数量患者的历史信息。

[0039] 电子医疗记录系统104可包括任何数量患者的任何数量的患者记录(例如,患者数据)。在一个示例中,电子医疗记录系统104可接收并为任何数量的医师提供关于任何数量的患者的医疗信息。在一个示例中,电子医疗记录系统104可包括本地患者信息(例如,医院等的任何数量患者的患者信息)和/或当前信息(例如,将要执行的实验室等)。

[0040] 医疗专业人员装置108是与医疗专业人员相关联的任何装置。在各种实施例中,医

师可利用医疗专业人员装置108。在各种实施例中,医疗专业人员装置108可提供患者信息至医疗记录贮藏处102和或电子医疗记录系统104。医疗专业人员装置108可接收基于患者历史数据的共有序列报告(例如,护理路径)和/或基于发现的共有序列与当前患者信息提供预测。医疗专业人员和/或医疗专业人员装置108可考虑到患者评估共有序列报告以确定优选的行为方针。

[0041] 通信网络110可以是允许数字装置通信的任何网络。通信网络110可以是互联网和/或包括LAN和WAN。通信网络110可支持无线和/或有线通信。

[0042] 共有系统106是可被配置为分析数据(例如,来自电子医疗记录系统104的历史患者信息)以生成共有序列报告(例如,指示来自自己执行行为的过去的序列记录的共有时间模式的报告)的数字装置。

[0043] 共有系统106也可从医疗专业人员装置108接收患者信息,并基于接收的信息和发现的共有序列提供行为方针或评估。关于图2描述示例共有系统106。

[0044] 图2是在一些实施例中的示例共有系统106的框图。共有系统106可包括事件集构造模块202、度量构造模块204、归类模块206、过滤模块208、距离模块210、片段度量组装模块212、自动分组模块214、共有模块216和预测模块218。在本文中用针对图3的流程图论述的更详细示例来描述每个模块。

[0045] 在各种实施例中,事件集构造模块202被配置为从事件构造事件集和片段。事件集中的事件是其顺序不可区分(例如,由医生同时命令(order)的测试的集合)的行为(例如,交互)。那么片段为这些事件集的序列。

[0046] 例如,事件集构造模块202可从任何数量的来源(诸如例如,医疗记录贮藏处102和/或电子医疗记录系统104)接收数据。数据可以例如包括关于任何数量患者的医疗信息。在一些实施例中,数据可以包括患者标识符,以及该患者的任何量的信息,该患者的任何量的信息包括例如在分派医疗检查时的医疗检查、当分派和/或执行医疗程序时的医疗程序(例如,X射线、外科手术等、药物的应用)、结果评估、预后、症状等。

[0047] 数据可来自任何数量的数据库。数据库可以是包含数据的任何数据结构(例如,多维数据的非常大的数据集)。在一些实施例中,数据库可为关系数据库。在一些实施例中,关系数据库可与MySQL、Oracle、微软SQL服务器、Aster nCluster、Teradata和/或Vertica一起使用。本领域技术人员应认识到数据库可以不是关系数据库。

[0048] 在一些实施例中,用户识别数据结构并且然后识别ID和数据字段。数据S可基于ID和数据字段内的信息。本领域技术人员应认识到,数据S可为有限度量空间或其概括,诸如图表或加权图表。在一些实施例中,数据S可通过公式、算法或通过明确指定每个成对距离的距离矩阵来指定。

[0049] 在各种实施例中,通过(例如,最优地)配对事件集,并然后使用对多个(例如,一对)事件集的度量,关于允许至少两个片段之间的差异(或相等性、相似性)的程度的定量归因的片段来构造度量。在一些实施例中,护理路径度量(CP)(例如,序列度量)使用这两个步骤,并且事件集护理路径度量(ESCP)(例如,子序列度量)仅利用后面的步骤。将在下文中详述这两个度量的构造。

[0050] 在下文中,我们示出了如何构造适当的事件集序列、对事件的度量、对事件集的度量,以及最终对片段的度量。我们然后将用于从片段的簇推导共有序列的过程公式化。我们

还描述用于预测先前未观察的片段的的结果的方法。最终,我们说明在健康护理中的应用(构造护理路径,即,在护理提供者和患者之间的交互的序列)和在同一域中实践的减少。

[0051] 度量构造模块204被配置为从数据获知对事件的度量。在一些实施例中,度量构造模块204如本文所述从接收的数据构造可用片段(例如,事件集的序列)。度量构造模块204可利用归类模块206。归类模块206可将接收的数据中的事件归类。在一些实施例中,归类模块206可从另一来源接收标记或其他类别。例如,通过接收的数据识别的事件中的全部或一些可被分配类别(例如,接收的数据中的事件中的全部或一些可与可被用于归类事件的元数据相关联)。在一个示例中,归类模块206可接收可被用于归类事件的本体。在一个示例中,在医院中执行的各种事件可被分配类别(例如LABS)。接收的数据可包括例如事件和类别的描述。

[0052] 在各种实施例中,归类模块206可生成或接收事件的类别(例如,“一般类别(general category)”)。归类模块206可以例如基于元数据或与接收的数据相关联的其他信息生成类别。例如,归类模块206可通过使用与事件相关联的一些数据以生成类别,来生成事件类别(例如,基于自然语言处理、语义分析、与其他事件相关的事件的时间、描述事件的可用元数据、事件中涉及的患者的类型、事件中涉及的医生或医疗专业人员的类型等)。

[0053] 归类模块206可生成用于事件的综合类别,以及事件的一般类别。综合类别可以是例如比一般类别更一般或更宽泛的类别(例如,综合类别可以比作为本体的一部分提供的类别更宽泛或更抽象)。应认识到,事件可以被分配两种类别或更多类别(例如一般类别和综合类别)。

[0054] 过滤模块208可以可选地过滤(例如,删减)过于不经常发生以至于不可区别的事件。在一些实施例中,过滤模块208利用一个或多个集合函数来识别和/或过滤事件(例如,消除不经常的事件)。例如,过滤模块208可跨多个患者集合事件的频率以确定任何数量事件的相对频率。利用集合,过滤模块208可生成频率阈值(例如,最低5%)以滤除降至低于所生成阈值的事件。

[0055] 在各种实施例中,过滤模块208通过利用频率阈值识别将要过滤(例如,移除)的事件,并且随后用与一般类别相关联的综合类别的实例替代该事件。在一些实施例中,过滤模块208用综合事件(例如,与综合类别相关联的另一事件)的实例替代将要过滤的事件。应认识到用综合事件替代事件或维持事件以及改变分配的类别(例如,改变为综合类别)可以降低噪声。

[0056] 返回度量构造模块204,度量构造模块204可使用(多个)类别获知对事件的度量。以下论述在高水平获知的度量。可在关于流程图3的论述中发现更详细的描述。

[0057] 在各种实施例中,度量构造模块204利用维数降低来形成对事件的度量。“情境”可从每个事件集中定义。情境可以是在事件集上存在的类别的集合,该事件集可通过给每个类别分配维度来向量化,并且进一步标准化为标准值(例如,等于一的欧几里得标准值)。

[0058] 对于每个事件,度量构造模块204可将事件所属的所有事件集的情境求和。结果,每个事件可与单位向量相关联,并且在此向量之间的角度可用作对事件的度量的基础。

[0059] 距离模块210可计算事件集之间的距离(例如,对于每个群的要素对使用贪婪算法)。例如,距离模块210可确定距两个事件集的所有对的距离,从而移除精确匹配。

[0060] 片段度量组装模块212可被配置为生成对片段的两个度量,包括例如护理路径

(CP) 度量和事件集护理路径 (EPSC) 度量。这两者均依靠于事件“锚定”，在此之后的概念是在一对片段中的事件和事件集由于它们的相关片段 (不仅由于它们自身的准确事件) 因此具有一些固有关系。例如，在外科手术之前发生的事件和外科手术之后发生的事件之间可存在差异，并使得该外显可以是显著的。在一些意义上这可称为最优化 (在分布解决语句中)，但改进不仅在速度上：使用锚定也可大大降低噪音。

[0061] 通过一个锚点事件，可在外科手术前用减小的负数并在外科手术后用增大的正数将群编号。通过多个锚点，在最早锚点之前的事件-群可用减小的负值编号，并且在每个锚点后用多个指数编号 (一个用于锚定指数并且一个用于事件-群序数)。可为不同锚定群之间的配对建立规则 (包括惩罚) ——例如，对于外科手术治疗，示例规则为在外科手术前群和外科手术后群之间没有配对。

[0062] 图9是示例事件集，其中一个锚点事件示出在一些实施例中的编号的群。时间902描绘用外科手术前用减小的负数和在外科手术后用增大的正数编号的事件。父类别904可以是综合类别，并且类别906可以是在接收的数据中提供的一般类别。代码908可以是与事件相关联的代码，而间隔时间910可以是事件的持续时间。描述912可提供关于事件的额外信息。

[0063] 在各种实施例中，片段度量组装模块212可使用动态时间规整 (DTW) 的修改版本来构造CP度量。DTW是用于测量可在时间或速度上变化的两个时间序列之间的相似度的众所周知的算法。片段度量组装模块212可使用DTW来匹配一对片段中的事件群从而定义距离。

[0064] 在计算ESCP度量时，片段度量组装模块212可通过锚定值 (例如，无条件地) 匹配事件。该匹配可利用事件中的重要重叠识别 (多个) 子集来辅助构造共有。

[0065] 自动分组模块214可使用CP度量将子群自动分组。在各种实施例中，数据集的数据点或图表中的节点可自动地被分组 (“自动分组”)。分组可以是对原目标 (即，数据点的集合或图表的节点的集合) 的可能的分区进行评分的给定评分函数 (例如，CP度量) 的可能的最大量 (例如，最优的最大量) 的近似。

[0066] 自动分组可被用来自动寻找共享一个或多个给定属性的一些集合Y的子集的集体 (collection)。在一个示例中，自动分组可被用来寻找是Y的分区的子集的集合，其中Y是有限度量空间X或图表中的节点的子集。然而，应认识到基于本公开，本文中描述的方法无此要求。

[0067] 在各种实施例中，数据集 (例如，原数据集或在可视化中的节点) 的可能的分区的选择可被识别和评分。分区是给定集合的不相交子集的集体。每个分区的子集的联合等于整个原始集。层次聚类方法可被用在原始对象Y上以创建Y的一族分区。

[0068] 自动分组是其中识别最高评分的分区的过程。最高评分的分区可以是一个或多个给定评分函数的最大值。在一些实施例中，可生成可能的分区的有限数量的分区。实际上，在一些情况下，如果评分方法为不优选的，则结果可以是更优的，因为至少一些层次聚类算法通常避免实际上可为此评分函数的全局极值的带有大量混杂的单独事件或其他低劣集的分区。应认识到层次聚类过程可用来调节数据从而仅呈现“良好的备选者”，并因此可改进一些评分器的效率。

[0069] 由于数据集的分区的数量为高 (例如， $(N/\log(N))^N$)，因此生成每个可能分区可能是不实际的。不幸的是，大多数局部改进方法可能容易卡住。生成分区子集的一些技术涉

及通过做出初始分区并且然后做出局部改变(例如,将节点从一个分区移动到另一个)来尝试将对图表分区上的模块性评分最大化。模块性是落入给定群内的边缘的分数(fraction)减去在边缘随机分布的情况下的预期的此分数。不幸的是,模块性测量Q评分(score)通常表现出极度退化,因为准许指数数量的不同的高评分解决方案并通常缺少明显的全局最大值。通过局部方法对分区上的函数最大化的另一途径是使用概率性技术,诸如模拟退火。本文所述的至少一些实施例提供可应用于宽范围的评分函数的确定性的可替换方案。

[0070] 在所生成的那些分区的一个或多个不同分区中的子集可至少部分基于CP度量值来选择。可生成包括选择的子集的新分区,或如果所有的选择的子集已经为生成的分区的一部分,那么可选择先前存在的分区。

[0071] 使用评分函数自动分组的示例关于图14-图17论述。应认识到可使用任何评分函数诸如例如CP度量值执行自动分组。

[0072] 共有模块216可被配置为寻找自动分组子集(例如,来自自动分组模块214的选择的分区)的评分。例如,给定片段S的子集,共有模块216可计算S中的点x,使得 $\sum(S\text{中的}y) CP(x, y)$ 最小:我们将这种点称为在CP下在S点中的“最大值向心性”的那些点。使用CP给定该最中心的子集(称此为M),共有模块216然后使用ESCP寻找M的最中心子集,并且其为共有模块216表示为S的核心C的子集。

[0073] 在各种实施例中,在已计算评分时,共有构造为最优化问题:我们寻找事件-集c的候选序列,使得 $Q(c, S) = \sum(C\text{中的}y) CP(c, y)$ 被最小化,受到“可信度”限制:c中的事件不能是不实际的。具体地,在一个示例中,这意味着共有模块216可以以实际片段开始,并然后谨慎编辑该实际片段,保持编辑使得 $Q(c, S)$ 改进。共有模块216可使用任何最优化技术(带有贪婪算法的一级回溯)。

[0074] 预测模块218被配置为使用本文中描述的距离测量来预测新片段的的结果(即,提出的行为方针)。使用一个或多个此类距离矩阵的线性组合、以及结果因变量(dependent outcome variable)的值,预测模块218可构造预测因子,其可在给定新实体状态、片段或该两者的组合的输入的情况下预测结果因变量的值。

[0075] 包括在共有系统106和说明书中的别处中的模块、引擎和数据存储表示特征部。在本文中描述的模块和数据存储可通过电子硬件(例如ASIC)、软件、固件或其任何组合来实施。作为单独的模块和数据存储的不同特征部的描绘不一定暗示模块和数据存储是否通过共同的或单独的电子硬件或软件部件实施。在一些实施方式中,与在本文中描绘的一个或多个模块和数据存储区相关联的特征部可通过共同的电子硬件和软件部件体现。在一些实施方式中,与在本文中描绘的一个或多个模块和数据存储区相关联的特征可通过单独的电子硬件和软件部件来实现。

[0076] 模块和数据存储可通过电子硬件和软件部件实施,包括但不限于一个或多个处理单元、一个或多个存储器部件、一个或多个输入/输出(I/O)部件以及互连部件。互连部件可被配置为在一个或多个处理单元、一个或多个存储器部件和一个或多个I/O部件之间提供通信。例如,互连部件可包括被配置为在电子部件之间传递数据的一条或多条总线。互连部件也可包括被配置为控制电子部件之间通信的控制电路(例如,存储器控制器和/或I/O控制器)。

[0077] 图3是在一些实施例中的用于生成共有时间模式的方法的流程图。在步骤302中接

收历史信息。例如，事件集构造模块202可接收关于任何数量患者的历史信息（例如，历史医疗信息）。

[0078] 在各种实施例中，事件集构造模块202被配置为从接收的数据构造事件集并且从事件构造片段。事件集中的事件是其顺序不可分辨（例如，由医生同时命令的测试的集合）的行为（例如交互）。那么片段为这些事件集的序列。

[0079] 在各种实施例中，度量被构造在片段上，以通过（例如，最优地）将至少两个片段的事件集配对并然后使用对于多对事件集的度量来允许至少两个片段之间的差异（或相等性、相似性）的程度的定量归因。在一些实施例中，护理路径度量（CP）使用这两个步骤，并且事件集护理路径度量（ESCP）仅利用后面的步骤。

[0080] 在步骤304中，事件集构造模块202构造事件集的序列。为了构造事件集，事件集构造模块202可利用指定事件属于的集合的一个或多个启发法。这里，可构造以小时间量分开的集合（见图4）。应认识到可使用任何时间量（例如，范围从30秒直到2小时）。在一些测试中，我们发现产生的对事件的度量在5分钟处为最小嘈杂的。假定其他问题的行为的时标非常不同，那么我们相当肯定该数字需要为此类场合而修订。

[0081] 度量构造模块204可被配置为从数据获知对事件的度量。在一些实施例中，度量构造模块204从接收的数据构造可用片段（例如，事件集的序列）中的全部或一些。度量构造模块204可需要对事件的归类和/或过滤。

[0082] 在步骤306中，归类模块206分配事件类别。例如，归类模块206可通过供应的自体将事件分配给较高级别的群（例如，类别）。例如，医院中的各种实验室测试可属于类别LABS。这些类别用于本文所述的度量构造示例。

[0083] 在一些实施例中，归类模块206为任何数量的类别（例如，针对由供应的自体提供的任何数量的类别）创造综合事件类别。如本文所论述，综合类别可以是例如比一般类别更一般或更宽泛的类别（例如，综合类别可以比被提供作为自体的一部分的类别更宽泛或更抽象）。应认识到，事件可以被分配两个或更多个类别（例如，一般类别和综合类别）。

[0084] 例如，用于LABS的综合事件类别可标记为“GENERIC LAB”，并且应理解为一些迄今为止尚未获知的实验室的占位符。注意一般事件的类别可定义为针对其创建一般事件的类别。

[0085] 在一些实施例中，对事件-群的度量取决于具有对事件的度量。如果没有给出对事件-群的度量，则度量构造模块204可使用数据来构造对事件-群的度量。如本文所论述，事件可以类别被群集。类别可用于构造对事件的度量。

[0086] 在步骤308中，过滤模块208可选地过滤过于不经常发生以至于不可区别的事件。在一些实施例中，过滤模块208利用一个或多个聚合函数进行识别和/或过滤（例如，消除不经常的事件）。在一个示例中，过滤模块208使用事件频率的累积分布函数生成累积分布并且过滤（例如“扔掉”）5%尾部。在一些实施例中，过滤模块针对最普通事件标记事件0，针对接下来的最普通事件标记1，等等。过滤模块208然后可过滤或（例如，移除或消除）以一数字开始的每个事件，使得从该数字起的事件的总计数 \leq 事件总数的5%。在各种实施例中，过滤模块208利用同一类别的综合事件的实例来替代事件的过滤过程。应认识到该步骤可在所得的事件矩阵空间中显著降噪。

[0087] 在步骤310中，度量构造模块204使用类别和/或过滤的事件来获知对事件的度量。

在各种实施例中,可存在数量级上比类别更多的事件类型(即,类别空间的维数比事件空间的维数低得多)。我们使用此维数降低来对事件进行度量。我们将来自事件集的“上下文”定义为在该事件集中存在的类别的收集。例如,如果外科医生同时命令三个实验室检验和胸部x光,则这形成类似于以下的集合:

[0088] {"全面代谢检查,"

[0089] "CBC差,"

[0090] "脂肪酶,"

[0091] "XR胸正位(PA)及侧位"}

[0092] 这样做是生成“上下文”{LAB,LAB,LAB,XRAY}。度量构造模块204可通过在欧几里得空间中向每个类别分配“维度”将这些上下文向量化。如果我们假设给予XRAY维度2和LAB维度4,则上面的向量化上下文将类似于{0,0,1,0,3,...},其中所有其他实体为0。我们进一步将这些上下文标准化以使得它们具有欧几里得范数=1,这意味着上下文变为{0,0,1/sqrt(10),0,3/sqrt(10),...}。

[0093] 现在对于每个事件,度量构造模块204可取得事件所属的所有事件集的上下文之和,并且对于综合事件,度量构造模块204可将具有与综合事件相同的类别的任何事件的所有上下文进行总和。在向量标准化之后,这可在中等维数的欧几里得空间中向每个事件(包括综合事件)给予单位向量,并且此类向量之间的角度用作对事件的度量的基础(参见例如图5)。

[0094] 应认识到,可存在是单元素(singleton)的更大数量的事件集。这种缺少“另外信息”(即,无上下文)意味着度量不可有效分辨这些行为。在各种实施例中,假设此事件-集存在,可改变操作为添加一值(例如,.01)乘以(times)包含事件的事件-集之前和/或之后的上下文。该小变化可具有期望的结果而不干扰度量的其余部分。并且其次,因为综合事件就其定义而言是不确定的,所以对于该类别的真实事件,在一般事件和任何其他事件之间的距离可按照“离散度”来定义。

[0095] 类别的离散度可以任何数量的方式定义。例如,类别的离散度可被定义为真实事件的综合向量化和所有向量化之间的角距离的平方的平均值的平方根。这给出如何“抹掉”类别中的事件的向量化的测量,并且可为综合事件和非综合事件之间距离的合理测量。对于一对综合事件,如果它们具有相同类别,则距离为离散度,否则距离为它们的各自离散度之和加上它们的相应向量化之间的角距离。任何真实事件的离散度可被定义为0。

[0096] 作为最优化,由于没有非常大数量的不同事件,因此度量构造模块204可以预先计算并将事件空间的距离矩阵缓存。

[0097] 如果事件没有配备先验以产生对事件的归类的本体,则可使用简单迭代算法从数据推导归类。例如,归类模块206可为所有事件产生全N x N共生矩阵,并使用标准聚类算法来定义该空间中的簇(产生M<N个簇)。度量构造模块204可将那些簇取为上述类别,并且重获知M维空间中的事件度量。度量构造模块204可以在该空间中再次聚类,产生M' < M个簇,并且重获知M' 维空间中的事件度量。度量构造模块204可重复直到该程序收敛并将最终聚类作为正确归类。

[0098] 在步骤312中,距离模块210计算事件集G1和G2之间的距离。在一个示例中,距离模块210通过对来自每个群的元素对使用贪婪算法来计算事件集G1和G2之间的距离。例如,距

离模块210可移除精确匹配(添加它们的距离,其将为0,除非一些综合事件存在),并且然后距离模块210可为两个集合中的所有(剩余)对计算距离。距离模块210可贪婪地移除对(首先移除最短距离,如果两个事件都仍在它们的相应的集合中),并且然后任何留下的未匹配对与一般事件配对(见图6)。应认识到,这可以是典型解决方案直接延伸到动态时间规整(其可在本文中利用)中的配对问题,其中单个类型的“不匹配”事件存在。然而,应认识到,存在距离模块210可应用的其他数量的其他配对(例如,使用匈牙利算法)。

[0099] 当配对事件A和B为两个不同事件-群时,距离模块210可考虑两种情况:在A和B之间的事件距离,以及在A、A和B的综合版本与B的综合版本之间的距离的和。距离模块210可取得这两个值中的较小值作为A和B的配对距离。作为该过程的一部分,距离模块210可有时保存事件群之间的显式匹配。在一个示例中,距离模块210可在布尔标志控制下在距离计算中合并此代码。

[0100] 为了更具体,在一些实施例中,如果我们假设eva代表由“A群”的整数事件代码排序的事件目标的阵列,并且evb代表由“B群”的整数事件代码排序的事件目标的阵列,则我们可使用拉链算法以将事件阵列复制到临时整数堆栈sa和sb中,sa和sb将指数分别包含到eva和evb中(除了任何精确匹配)。我们通过这些精确匹配之间的距离增大返回值,该距离可为零,除非匹配事件是一般的。如果我们保留显式匹配,则我们将在此点处精确匹配的对保存在单独的事件对堆栈中。

[0101] 我们然后将所有未匹配的事件对存储为浮点数/整数/整数三元组f/i/j作为压缩64位Java长整数:这里,浮点数为事件eva[sa(i)]、evb[sb(j)]之间的事件距离、以及事件eva[sa(i)]和evb[sb(j)]和它们的属类之间的事件距离之和的最小值。当f在IEEE 32位格式中并且 $\geq 0.0f$ 时,我们可使用Java中的正则的“长排序(long sorting)”以将这些三元组排序。

[0102] 我们遍历排序的三元组,从而在两个元素仍存在时移除我们遇到的任何对(清除sa/sb中的相应项,因此我们知道它们被移除),并将返回值增大f。当我们已经清空sa/sb中的一者时,我们停止并且处理任何剩余的sb/sa项,以该项及其属类(generic)之间的距离来增大返回值。如果我们保存精确匹配的对,则我们还使用指数来从eva和evb中提取事件对。然后我们返回累积匹配值。

[0103] 在步骤314中,片段度量组装模块212组装对片段的度量。如本文所论述,我们引入两个对片段的度量:护理路径(CP)度量和ESCP(事件集护理路径)度量。

[0104] 在一些实施例中,为了构造CP度量,片段度量组装模块212可使用动态时间规整(DTW,也称为序列比对)的修改版本来匹配一对片段中的事件-群,从而定义距离。这里,事件集为事件序列(可能包括综合事件),并且“综合”集是全部包括综合事件的事件集。如果A为事件集,我们定义synth(A)为通过以其属类形式替换A中的每个代码而形成的集合:对于 $A = \{NUR1, CT5, \text{属类_XP}\}$,那么synth(A)为 $\{\text{属类_NUR}, \text{属类_CT}, \text{属类_XR}\}$ 。

[0105] 假设我们具有定义在事件集对的度量 $d(A, b) \geq 0$,对于*综合事件,该度量也具有 $d(A, *)$ 被 $d(A, \text{synth}(A))$ 最小化的属性。然后,我们可定义对于事件序列对的度量,即使序列具有不同长度。给定相同长度的两个集合的序列 $AB\cdots$ 和 $ab\cdots$,我们可定义“基础度量” $bd(AB\cdots, ab\cdots)$ 为 $d(A, a) + d(B, b) + \dots$ 。我们然后通过用插入“综合集合”来“编辑”序列来得到相同长度的多对序列,并然后取得最小值,从而将其延伸到可能不同长度的两个序列

ABC...和abc...。如果我们以*表示任何综合集合,那么我们可遍及所有可能的等长度的多对编辑来将AB和abc之间的距离定义为 $\min\{bd(AB*,abc),bd(A*B*,a*bc),\dots\}$ 。这是良好定义的,因为具有对应的*的任何可能的编辑对可被编辑,从而移除那些匹配综合集合($bd()$ 之和的值将不增大),并且我们获知每个*可由在其他序列中的匹配的synth()事件来替代。从这点,我们看到我们可遍及有限的编辑集合取得最小值,并且因为不久将变得明显的原因,这些编辑可被认为是形式的“配对”的序列对 $A\langle-\rangle x, Y\langle-\rangle *或*\langle-\rangle z$,我们可分别称为“对角线的”、“水平的”、“垂直的”配对。

[0106] 假设@为以其他方式不能使用的特殊代码。在带有通过成对(x,y)标记的网格中存在此类配对和路径(到右边并且上升)之间的自然对应,其中x在{@,A,B,C,...}范围内,并且y在{@,a,b,...}范围内。对角线地上升对应于对角线配对,而垂直边缘为垂直配对并且水平边缘为水平配对。作为示例,将ABC与A*BC $\langle-\rangle$ ab*c的abc匹配在图7a中被示出;通过AB*C* $\langle-\rangle$ a**b*c将ABC与abc匹配在图7b中被示出。

[0107] 路径直观地具有“在被消耗的序列元素的方向上的分量”。路径必须在左下开始并且上升和/或到右边至右上角。另外,由于我们在寻找关于片段上的共有,因此我们很可能期待不关注带有“充分多的”综合集合的路径,因此如果序列在长度上过于不同或过于远离对角线,则我们很可能返回“巨大值”并退出。这提出在集合水平处的匹配可能在集合数量上是线性的一并且至少低于二次的一因为关注的路径将被受限为在对角线周围。为了寻找最优路径的成本,我们仅需要与所有路径处于其中的网格相同大小的矩阵。我们将到达该点的最小路径成本分配给每个网格点。

[0108] 我们可仅从(B,a)、(C,a)或(B,b)到达点(C,b),并且仅有从这些点中的每个点到(C,b)的一种方式,因此获知这3个值意味着我们可计算第四个——用于动态编程的优选状况(注意,实际上,我们仅需要当前列和先前一个列的值,这减少了所需的中间状态,尽管没有减少计算数量)。我们将通过表格MinCost(,)表示最小成本状态,即适当初始化的MinCost(M,n)将为元素A,...,M与a,...,n的最佳配对的成本,其中MinCost(@,b)意为将*,*与a,b配对(即, $bd(*,*)$,(a,b)),等等。MinCost(@,@)天然为0。现在让我们以g[0]、g[1]等表示事件A,B,C,...,并且以阵列h[]的元素表示a,b,c,...。代替MinCost(),我们将使用定义为MinCost(g[i-1],h[j-1])的矩阵DTW[i][j]——即,DTW[i][j]为将g的第一i项与h的第一j项比对的最优路径的成本。(这为@保留指数0。)DTW为MxN,其中M=g.length+1和N=h.length+1,以及最优路径的成本为DTW[g.length,h.length]。在伪码中,匹配在下面按照DIST()执行,其中synth(set)为事件集的综合版本,并且d(群,其他集合)为事件集之间的非负对称距离:

[0109]

```
DIST(g: 集合[0..N-1], h: 集合[0..M-1]) {
```

```
    // DTW[x][y] 为“消耗”g的第一x元素和h的第一y元素的最小路径的
    //成本——也就是说，在平面中从(0,0)到(x,y)的最佳路径。
```

```
    DTW := double[N+1][M+1]
```

```
    // 对于网格的底部行填充值
```

```
    for i := 0 to N-1
```

```
        DTW[i+1][0] := d(g[i],synth(g[i])) + DTW[i][0] // 水平的
```

```
    //对于网格的左侧列填充值
```

```
    For j := 0 to M-1
```

```
        DTW[0][j] := d(synth(h[j]), h[j]) + DTW[0][j] //垂直的
```

```
    for i := 0 to N-1
```

```
    for j := 0 to M-1
```

```
        DTW[i+1][j+1] :=
```

[0110]

```
            min(d(g[i], h[j]) + DTW[i][j], // 对角线的
```

```
                d(g[i], synth(g[i])) + DTW[i][j+1], //水平的
```

```
                d(synth(h[j]), h[j]) + DTW[i+1][j]) //垂直的
```

```
    return DTW[N, M]
```

```
}
```

[0111] 对于所有 $j \geq 0$, $DTW[0][j+1] = d(\text{synth}(h[j]), h[j]) + DTW[0][j]$ 因为到 $(0, j)$ 的唯一可能路径具有与综合事件匹配的第一 $j+1$ 项。对于 $DTW[i][0]$ 是类似的。

[0112] 事件群距离可如所示在下面伪码中进行。这种观念用于惩罚指数彼此距离过远的事件集的匹配,因此可拒绝“过于不同”的片段对。

[0113]

```
public double eventSetDistance(片段 pA, 事件集 a,  
                               片段 pB, 事件集 b) {  
    //如果锚定代码不匹配则摆脱——我们可利用此来避免  
    //分开锚定段并且分别在它们上进行 DTW,  
    //但是其复杂性通过使网格的侧面更小的性能增益来超重  
    if (a.anchorCode != b.anchorCode)  
        return HUGE;  
    // 我们将 2/3 的“平均综合误差”作为使误差略弱于平均综合误差的规模。  
    double sdError = (a.distanceToSynthetic + b.distanceToSynthetic)/3.0;  
    double indexError = (a.zbIndex - b.zbIndex)/2.0;  
    double incr = sdError * indexError * indexError;  
    //因为即使大量不同时期的群之间的精确匹配仍是非预期的,  
    //通常包括误差  
    return eventSetPairing(pA, a.events, pB, b. events, null) + incr;  
}
```

方法 eventSetPairing()在来自片段 A 和 B 的相应锚定集合中进行事件对的

[0114] 无条件匹配, 并且返回结果距离。

[0115] 最终, 在计算ESCP度量(用于下面的共有核心计算的第二步骤)时, 片段度量组装模块212可通过锚定值仅执行事件的无条件匹配(例如, 对于外科片段, 我们从每个片段中的事件得出正群和负群, 并且将它们配对且返回总误差)。在一些实施例中, 片段集可被一次减小至作为事件集的“时序”良好匹配的那些片段。该步骤尝试在实际事件中寻找带有大量重叠的子集, 以帮助构造共有。

[0116] 共有模块216可从片段的簇中确定共有序列。在已从数据获知CP度量和ESCP度量后, 我们现在描述在数据源中的片段的聚类/划分, 以及用于根据分区计算共有序列的过程。

[0117] 在步骤316中, 构造事件的子群。在各种实施例中, 自动分组模块214使用护理路径度量CP自动地构造相关的子群。在一个示例中, 自动分组模块214使用5个最近的邻点为每个点构造度量空间的图表, 其中在点p和q之间的添加的边缘强度为 $1/(原始距离(p, q))$, 其中如果q为p的第j个最近邻点, 则原始距离(p, q)为j。在构造图表之后, 自动分组模块214使用95%作为“聚类截断”来自动分组。自动分组模块214然后生成图表中的节点的分区, 并且由于这些节点为片段, 因此该分区引起收集“相关子集”, 在该“相关子集”上搜索共有(参见图8a-图8b)。应认识到, 可利用任何数量的最近邻点(例如, 不仅五个), 并且聚类截断可以是任何阈值(例如, 任何百分比、离散值等)。

[0118] 如本文所论述, 自动分组的过程关于图14-图17被描述。

[0119] 在已降低在给定子集上搜索共有护理路径的问题之后, 共有模块216可寻找子集

的“核心”，并且然后在步骤318中共有模块216处理该核心以产生共有。给定片段S的子集，共有模块216计算S中的点x，使得 $\sum(S \text{ 中的 } y) CP(x, y)$ 最小：我们将此点称为在CP下的在S中的“最大值向心性”的那些点。使用CP给定该最中心的子集（称此为M），共有模块216然后使用ESCP寻找M的最中心子集，并且其为共有模块216表示为S的核心C的子集。我们拒绝长度小于40的任何输入子集，由于20为我们发现可用的最小核心尺寸，并且我们优选带有至少100个点的集合。为了提高满足这些尺寸限制的可能性，在第一步骤中，共有模块216寻找M，M的尺寸为150以及 $(1/\sqrt{2.0})$ 乘以S的尺寸中的最小值。C的尺寸被取为100以及 $(1/\sqrt{2.0})$ 乘以M的尺寸中的最小值。这些值将针对不同的数据类型而被调整（参见图8b）。

[0120] 在已计算核心后，共有构造可以是最优化问题：我们正在寻找事件-集c的候选序列，使得 $Q(c, S) = \sum(C \text{ 中的 } y) CP(c, y)$ 被最小化，受到“可信度”限制：c中的事件不能是不实际的。具体地，在一个示例中，这意味着共有模块216可以以实际片段开始，并且然后谨慎地编辑该实际片段，保持编辑使得Q(c, S) 改进。共有模块216可使用标准最优化技术（带有贪婪算法的一级回溯法）与两个非标准启发法，如下所述。

[0121] 第一非标准最优化步骤为当开始最优化时，无需通过展开时间来推断单独事件的时间，因此事件集中的每个事件都具有精确相同的时间，并且事件集保持以固定增量隔开。在添加或移除事件之后，共有模块216重新运行该过程。仅在结束（例如，一旦我们对构成的事件-群感到合适），共有模块216通过取得核心集中的片段中的匹配事件的中值时间来调整共有中的事件时间（参见图8c）。

[0122] 在一些实施例中，第二非标准最优化步骤为在编辑过程中的规则，使得如果事件的计数降至低于某最小数值，则事件不可被移除，我们将该最小数值视为在核心集中的该事件的计数的第一四分位数和第二四分位数的平均值的下整数（floor）；这防止普通事件消失但允许我们在这改进CP向心性时减少数量。我们还尝试添加整个群，但这极少成功，因为共有的起始点不趋向于是丢失的群——代替地，群丢失在核心中的大多数其他事件-群中发现的偶然的单独事件。

[0123] 在步骤320中，预测模块218可使用本文中论述的距离测量值预测新片段的结果（即，提出的行为方针），并且可选地预测表示在片段开始之前所关注实体的状态的一个或多个另外距离。每个此距离可引起实体之间的距离矩阵。使用一个或多个此类距离矩阵的线性组合，以及结果因变量的值，预测模块218构造预测因子，其可在给定新实体状态、片段或该两者的组合的输入的情况下预测结果因变量的值。通过此类预测，人们可以例如关于所关注的一个或多个结果变量最优化实体状态、片段或该两者；基于所述输入来预测结果；或相似任务。

[0124] 图10为在一些实施例中的健康护理示例的流程图。在健康护理示例中，被解决的问题为围手术期的外科护理。也就是说，被解决的问题为确定在医院内在外科手术前后应提供哪个治疗以最大化良好结果的机会。在此情况下的事件为护理提供者 and 患者之间的各种交互：实验室测试、医嘱和服用的药物。事件关于外科手术的时间（锚点）来考虑。在该示例中，事件以在数据源自身（例如电子医疗记录（EMR）数据库）中存在的本体归类。事件集松散地映射到命令集的概念：护理提供者在与患者交互期间指定的被命令和/或施行的大量行为的一个或多个行为。

[0125] 在该示例中，片段为在前后可操作的一些指定时段（例如，在外科手术前30天以及

外科手术30天)的此类集合的序列。示例共有片段(在医疗情况下,“护理路径”)使用本文所述的(多个)方法导出,被医师检查并修改为最终形式(通过来自预测性能力的支持),并然后构建到EMR系统自身中。当医师与患者交互时,EMR系统可向医师推荐护理路径,该医师治疗患者并记录治疗与结果的。系统然后将回馈其自身并随时间推移持续改进护理路径。

[0126] 参见图10的概括。在步骤1中,医疗记录数据库将历史患者信息转到电子医疗记录(EMR)系统,其与数据驱动临床护理路径系统(DDCC)(例如,共有系统106)共享该数据。在步骤2中,DDCC基于临床结果和成本降低目标生成草案临床护理路径。可使这些草案可用于医师。在步骤3中,医师检查草案护理路径并生成最终版本。在步骤4中,最终护理路径上传到EMR,允许它们在患者登录医疗系统时自动激活。在步骤5中,患者具有新的医疗遭遇,并且临床护理路径根据诊断或命令的程序被激活。

[0127] 在进一步的示例中,根据历史信息构造事件集的序列。如上所述,事件可得自EMR数据库,并且被聚类为带有5分钟超时的集合。在一些实施例中,预期事件被时戳(timestamped)为至少分钟级的准确度内,并且许多事件可共享时间戳。

[0128] 可随后获知对事件的至少一个度量。数千或数万个事件类型可根据在EMR数据中存在的本体被归类至大约100个高级类别(例如,X射线、止痛药、护嘱等)。给定该归类,可在数据库中存在的治疗(例如,所有外科手术片段)、或在某指定时间的该数据的子集、外科手术操作、医院或相似物中获知度量。可在事件集上导出至少一个度量。可如上所述来处理事件集。

[0129] 随后,可关于片段组装度量。可使用上述修改的DTW算法。在该示例中,存在单个锚点,其为外科手术操作的精确开始时间。在一些实施例中,仅考虑围手术期的事件(即,在外科病房外发生的事件)。在锚点的相对侧上的事件可不被比对(即,引起巨大的比对成本)。当目标是产生被定制成特定护理交互的护理路径时,可仅比较来自相同外科手术操作的片段(例如,全膝关节置换术)。

[0130] 共有序列(护理路径)可从片段簇导出。在簇构造前,如果需要的话,片段群体可被过滤以提供一组更有针对性的治疗。例如,人们可希望仅为患有高血压的患者、在特定区域中的患者针对已知特别熟练的外科医生构造护理路径。在任何此类过滤和聚类之后,可根据由用户提供的评分方案来对簇评分。例如,人们可认为巨大簇为更优的(它们提供更强的统计支持)。也可期望对因子诸如医院成本、停留长度、患者满意度或再入院风险而将簇评分。可使用此类元数据变量的任何线性组合。带有最高评分的簇可被选择用于共有构造,并且结果可以特殊构造的界面的形式报告回用户,用于观察、操纵和输出建议的护理路径(参见图10)。注意该界面还示出构造护理路径的基础片段(包括最优化步骤的开始点)、以及附着到此类片段的任何元数据(诸如患者合并症、生命统计、主治医师等)。

[0131] 该方法还可提供预测针对所关注的结果将要进行建议的护理路径中的什么变化的交互式的方式。对于护理路径自其生成的每个簇核心,可对一些所关注的结果(例如,质量测量,如在医院中停留的长度)参数化预测因子(例如,K个最邻近的回归量或分类因子)。该系统允许用户(例如,医师)交互地更改建议的护理路径,这触发新的护理路径在所关注的结果方面将要做什么的预测。以此方式,用户可将领域知识注入到以其他方式自动化的过程中,并且探索自动生成的开始点中的何种不同的微调可改进下游结果。然后用户可将已编辑的护理路径和/或计算机生成的护理路径保存到系统中,并在实施过程中向前移动。

[0132] 图11是对健康护理的一些实施例的应用的另一示例。在一些实施例中,患者具有新的遭遇(encounter),并且根据诊断或命令的手术来激活临床护理路径。患者在治疗时可经历数个临床遭遇。护理路径可在EMR中数字地跟踪,以在患者经过护理路径时监测患者,使用基于来自临床遭遇的患者信息的预测结果作为警报标记。

[0133] 来自数字护理路径的基于先前患者遭遇而建议低劣结果的预测被标记,并且临床医生可调整患者行为方针以改进预测的结果。可基于护理路径的变化来采取医疗行为。患者结果预测被返回到稳定跟踪。

[0134] 图12是在一些实施例中的示例护理路径共有界面。在该示例中,每个患者标识符1202(为该示例匿名并且虚构的)与在事件序列1204中指示的一系列事件相关联。在该示例中,共有序列为在示例护理路径共有界面的顶部处的1206。共有序列1206可鉴于和每个患者标识符1202相关联的历史数据来指示最优(例如,在益处和/或成本方面最优)。在各种实施例中,示例护理路径共有界面和/或共有序列1206可以是或可以被包括在共有报告中。

[0135] 图13是在一些实施例中的示例护理路径共有编辑界面。在该示例中,每个患者标识符1302(为该示例匿名并且虚构的)与在事件序列1304中指示的一系列事件相关联。在该示例中,类似于图12,共有序列为在示例护理路径共有界面的顶部处的1306。共有序列1306可鉴于和每个患者标识符1302相关联的历史数据来指示最优(例如,在益处和/或成本方面最优)。在各种实施例中,示例护理路径共有界面和/或共有序列1306可以是或可以被包括在共有报告中。

[0136] 示例护理共有编辑界面可允许用户(例如,医疗专业人员)与患者事件信息和/或共有序列1306交互。在一个示例中,用户可点击患者或事件并观察与选择相关联的事件信息1308。应认识到,任何数量的事件可与选择相关联。在该示例中,用户可选择与综合类别“未知LAB”相关联的事件。可与未知实验室相关联的活动1310可描绘为允许用户更好理解过去执行的医疗手术的类型、结果信息、成功可能性、相对成本、手术顺序等。

[0137] 在各种实施例中,示例护理路径共有编辑界面可允许用户编辑护理路径和/或存储的历史信息(例如,添加另外的患者信息)。例如,用户(例如医疗专业人员)可经由编辑界面在编辑界面中添加、移除或编辑事件。可基于变化来生成和/或提供给用户更新的共有序列。在一些实施例中,预测(例如,关于成本和/或停留长度)可部分基于变化来提供。

[0138] 图14a-图14d描绘了在一些实施例中的基于评分来确定分区以用于自动分组的示例。在示例中,存在有限尺寸的固定空间S。空间的性质可仅在存在将空间聚簇并将子集评分的方式上内是相关的。参考图表,S上的图表G指示其节点为子集的集的图表,其中如果且仅如果两个节点具有共同的点,则节点连接到另一节点。分区包括一个或多个子集。该一个或多个子集中的每个包括S的所有元素。例如,分区1402为包括S的所有元素的分区。子集1404a-e包括S的所有元素。所有子集1404a-e的联合为分区1402。

[0139] S上的森林F为S上的图表。如果F中的每个叶子为单元元素(例如,带有一个成员的集合),则森林F为“原子的”。图14a(即,F1)为原子森林,因为如图14a所描绘的F1中的每个叶子为单元元素。应认识到图14b(即,F2)不为原子森林,因为如图14b所描绘的F2中的每个叶子不为单元元素。例如,F2包括叶子{A,B}、{D,E}和{F,G}。

[0140] 存在S的称为根的分区的分区R(在F1中,{a,b,c}、{d,e,f}、{g}),使得F中每个集合可由来自根的唯一路径到达。F中的N为叶子(例如,在原子森林中的单元元素)或其连接到形成N的

分区的节点(例如, F1中的 $\{a, b, c\} \rightarrow \{a, b\}$ 和 $\{c\}$)。对于非叶子节点N, 我们以 $C(N)$ 表示N的子节点。注意叶子的子节点, 即 $C(\text{叶子})$ 为空。如果F和F'具有相同的叶子并且F中的每个节点为F'中的节点, 则我们宣称F'延伸F。如果两个森林不相等, 那么F'包含是F中一个或多个根的联合的节点。示例F3(图14c)延伸F1(图14a)。

[0141] 如果并且仅如果P的每个元素在F1中, 则S上的分区P从属于F1。图14d中描绘的F4的圆形分区P1为F1的从属分区的示例(例如, $\{a, b, c\}$ 、 $\{d, e\}$ 、 $\{f\}$ 和 $\{g\}$)。

[0142] 单元素(S) (Singletons (S)) 表示为通过采取 $\{\{x\} \mid S \text{ 中的 } x\}$ 形成的分区。即, 在图14d中的示例中, 单元素 $(\{a, b, c, d, e, f, g\}) = \{\{a\}, \{b\}, \{c\}, \{d\}, \{e\}, \{f\}, \{g\}\}$ 。这与原子森林的叶子的集合相同。假设 $U(P)$, 其中P为S的子集的任何集, 表示P的所有元素的联合。 $U(\text{单个}(S)) = S$ 。

[0143] 如果并且仅如果P'中的每个元素x'为P中元素x的联合, 则S上的分区P'比S上的另一分区P更粗糙。在各种实施例中, S上的每个分区比单元素(S)更粗糙, 并且 $\{S\}$ 比S上的每个分区更粗糙。例如, $\{\{a, b, c\}, \{d, e, f\}, \{g\}\}$ 是比 $\{\{a, b\}, \{c\}, \{d, e\}, \{f\}, \{g\}\}$ 更粗糙的分区。

[0144] 图15描绘了在一些实施例中的示例自动分组模块214。自动分组模块214可包括数据结构模块1502、分区生成模块1504、评分函数模块(例如, Q_subset评分模块1506、Q_max评分模块1508、Q_partition评分模块1510)、分区选择模块1512以及数据控制模块1514。尽管评分函数模块论述为包括三个模块, 每个模块执行不同的评分函数, 但应认识到可存在执行任何数量的评分函数的任何数量的评分函数模块(例如, 执行单个评分函数的一个模块能够生成任何数量或类型的评分)。例如, 评分函数可生成和/或最大化任何数量的度量函数的度量值。

[0145] 在各种实施例中, 数据结构模块1502接收包括多个数据集的数据。该数据可从任何数量的数字装置接收。

[0146] 分区生成模块1504(例如“组块器”)利用由数据结构模块1502接收的多个数据集形成森林F。例如, 分区生成模块1504可使用由数据结构模块1502接收的数据生成森林F的第一分区。在一些实施例中, 第一分区可包括叶子, 其为来自该数据的所有元素的单个元素。在各种实施例中, 第一分区可包括任何数量的数据集。第一分区可包括森林的叶子、单个元素、根、多个元素的集合等。

[0147] 分区生成模块1504可使用第一分区生成森林F的第二分区。例如, 第二分区可包括第一分区中的至少两个集合的至少一个联合。随后的分区可以相似方式形成(例如, 至少部分基于包括来自先前分区的至少两个集合的至少一个联合)。

[0148] 分区生成模块1504可在将分区(或多个分区集)评分前生成整个森林F。例如, 分区生成模块1504可在任何或全部评分函数模块将森林F的分区的全部或部分评分之前生成整个森林F。

[0149] 在一些实施例中, 分区生成模块1504可在执行评分时或与分区评分(例如, 多个分区集的评分)相连来生成整个森林F。例如, 分区生成模块1504可在任何或全部评分函数模块将森林F的分区的全部或部分评分时生成整个森林F。在另一示例中, 分区生成模块1504可生成森林F的一个或多个分区, 并然后任何数量的评分函数模块可在分区生成模块1504生成森林F的一个或多个另外分区之前将已生成的分区评分。

[0150] 在各种实施例中,分区生成模块1504可通过森林F的(多个)先前生成的分区(或多个分区集)的任何数量的评分函数模块至少部分基于评分来生成森林F的分区。

[0151] 应认识到分区生成模块1504可能不生成整个森林F,而是可在森林F完成前终止生成森林F的分区。分区生成模块1504可基于森林F的任何数量的先前生成分区或与(多个)先前生成分区的全部或部分相关联的评分,来确定是否构建森林F的新分区。

[0152] 如本文所论述,分区生成模块1504可能不生成森林F的所有可能的数据集和/或所有可能的分区。

[0153] 应认识到,分区生成模块1504可利用任何数量的层次聚类技术与本文所述的技术。在一个示例中,数据和/或节点由 ϵ (epsilon) 联结(如果两个数据集或节点在彼此的距离 ϵ 之内,那么它们联结在一起)。尽管该示例标准技术具有传统限制(“固定的 ϵ ”),由此单个 ϵ 可能无法以优选方式分解空间,通过将分区的每个子集评分,我们可跨森林选择子集以识别和/或生成所选择的分区(例如,通过将多个分区的子集自动分组)。

[0154] 层次聚类技术的一个示例,有限度量空间X上的KNN是为每个点计算K个最近邻点(nearest neighbor),其中例如 $K=50$ 。分区生成模块1504可以为单元素(X) (Singletons (X))的INITIAL()开始。然后在对于 $1 \leq k \leq 50$ 在每个步骤处,如果x和y在彼此的对称的k个最近邻点中,则分区生成模块1504可将x连接到y。注意如果对于 $k < 50$,KNN(P,k)返回P,则分区生成模块1504可提高k并再次尝试,而不是断定P为稳定的。

[0155] 另一层次聚类技术实施例关于在点集S上的加权图表G(带有正权重)被定义。该层次聚类技术被预先确定的实数 Δ (delta) 参数化,其中 $1 > \Delta > 0$ 。分区生成模块1504以 $\Delta = 0$ 开始,因此INITIAL()为单元素(S)。对于每个分区P,我们对于在P中 $p \neq q$,将 $w_t(p,q)$ 定义为在图表中的节点之间的边缘权重之和除以 $|p| * |q|$,这些节点为G中的子集p的一部分和子集q的一部分。分区生成模块1504被配置为取得分区P并通过在 $w_t(a,b) \geq \Delta * \max(w_t(p,q))$ (其中最大值(max)高于分区P中子集p和q的所有对)时联结子集(a,b)的所有对(其中,a,b为分区P中的子集)来形成新的分区P'。

[0156] 存在用于层次聚类的任何数量的技术,并且它们中的任何技术可与满足对本文所论述的评分函数的示例限制的评分函数进行结合。

[0157] 自动分组模块214包括可分别利用三个评分函数的Q_Subset评分模块1506、Q_Max评分模块1508和Q_Partition评分模块1510。Q_Subset评分模块1506为一个或多个分区的子集计算Q_Subset评分。Q_Max评分模块1508基于子集的Q_Subset评分计算Q_Max评分(例如,基于Q_Subset评分为分区计算最大评分)。Q_Partition评分模块1510利用子集的至少一个Q_Subset评分为森林的两个或多个分区计算Q_Partition评分。

[0158] 在各种实施例中,Q_Subset评分模块1506计算Q_Subset评分(例如,为分区的每个子集计算一次评分)。函数Q被定义在空间S的子集上并且为将要在自动分组过程中群集在一起的属性评分。例如,在一些实施例中,Q_Subset评分为图表上的模块性评分(因此S为图表中节点)。分区选择模块1512可为带有(多个)最大模块性评分的图表S的分区检查数据结构。

[0159] 第二评分函数Q_Partition评分可以是将要定义在空间S的分区上的第一评分函数Q的延伸。如果评分函数Q被定义在S的子集上,则其可以各种方式延伸到分区函数Q_Partition。将函数Q延伸到分区的最简单方式中的一种为通过将Q_Partition(P)定义为Q

(p)的P中的p之和(例如,对于分区P, $Q_{\text{Partition}}(P) = \sum_{\{P\text{中的子集}p\}} Q(p)$)。

[0160] 在各种实施例中,Q_Partition必须具有以下属性:假设P为S的子集的任意分区,假设p属于P,并且假设q为p的分区。P(q)被定义为通过用q的元素替换P中的p获得的分区。那么在此示例中,Q_Partition必须对于如上所述的所有P、p、q具有以下属性:

[0161] (1) 如果且仅如果 $QP(q) \geq Q(\{p\})$, $QP(P(q)) \geq QP(P)$

[0162] 在一些实施例中,函数Q在此情况下不需要来自集合函数。满足属性(1)的函数Q_Partition按照定义为稳定分区函数。此类函数的种类如下描述。

[0163] 假设Q为定义在S的非空子集的集合上的任何实值函数。假设A(p,q)为定义在多对非空子集上的任何函数,使得p为q的子集。如果:

[0164] (2) 对于所有合规的p、q、r, $A(p,p) = 1$ 和 $A(p,q) * A(q,r) = A(p,r)$

[0165] 然后我们可通过以下公式将集合函数Q()延伸到所有分区P:

[0166] (3) $QP(P) = \sum_{p \in U(P)} A(p,U(P)) Q(p)$

[0167] p在P中

[0168] 注意所有实数k, $A(p,q) = (|p|/|q|)^k$ 满足该属性。此外, $k=0$ 意味着 $A(p,q) = 1$ 。

[0169] (1) 适用于在(3)中定义的Q。如果QP和QP'为稳定分区函数,那么因此对于 $x, y \geq 0$ 为 $x * QP + y * QP'$ 。我们还将S上的稳定分区函数称为F的“分区评分函数”。

[0170] 对于形式(3)的任何评分函数,单调增加的函数f可从实数到其自身被选择,并且以 $Q'() = f(Q())$ 替代Q。具体地,如果f()为“充分可逆的”(例如,A()和Q()为 ≥ 0 并且f()在非负实数上可逆)。QP(P)可由以下公式定义:

[0171] (3') $QP'(P) = f\text{-inverse}(\sum_{p \in U(P)} A(p,U(P)) f(Q(p)))$

[0172] p在P中

[0173] 由于f(QP(P))满足(1)并且f()是单调增加的,因此(3')中的QP'还满足(1)并且在S的子集上延伸Q()。具体地,如果在集合上 $A = 1$ 并且 $Q() \geq 0$,则QP(P)可被定义为在P的单个元素上的Q()的欧几里得标准,并仍得到评分函数。还可使用f()的指数函数而无需Q为非负的。

[0174] 在各种实施例中,对于在S的子集的分区上定义的函数Q利用 \leq 或 \geq ,在比较下存在极值。由于如果比较为 \leq 则Q可被 $-Q$ 替换,因此可以不失一般性地假设Q的最大值(即, \geq)是所关注的。具体地,在Q满足简单属性的情况下寻找Q在其上为最大的F-从属分区(subordinate partition)的方法在本文中公开。

[0175] 假定F上的评分函数Q_Partition,如果p为叶子,则我们可将评分函数Q_max定义为Q(p),并且如果p不是叶子,则可将评分函数Q_max定义为 $\max(Q(p), Q_{\text{max}}(C(p)))$ 。该定义和针对Q_Partition的需求(1)的一个结论为子集p的最大分区(即, $Q_{\text{max}}(V)$ 为最大的p的分区V)为p或C(p)的每个元素的最大分区的联合(连结可通过取得子集p而不是子元素(children))。

[0176] 在各种实施例中,自动分组方法对S使用层次聚类过程以计算F(即,构造森林F),并且如果Q_Partition为对F的根R的评分函数,则我们可寻找从属于F的S的Q_Max最大分区。如我们上面所述,这里直觉为将评分函数Q()与层次聚类联结提供了用于在分区中选择“Q-最大分区”的原理方法。

[0177] 分区生成模块1504以原空间S开始并形成上述的森林F。具体地,在一些实施例中,生成模块1504取得分区P并返回比P更粗糙的新分区P'。注意 $(\{S\}) = \{S\}$ 。使得生成模块1504 $(P) = P$ 的任何分区P被称为组块器-末端,并且重复的应用必须最终达到组块器-末端分区。序列Singletons(S) (单元素(S))、组块器(单元素(S))、组块器(组块器(单元素(S)))等在有限数量的步骤中终止,并且这些分区的联合形成原子森林F,该原子森林F的根为C-末端分区R中的元素,该元素为F的根。

[0178] 利用评分函数并生成分区的一个示例过程在以下伪码中如下:

[0179]

$P = \text{INITIAL}(S)$ //一些初始分区- 通常单元素(), 但其可为任何物

$F = \text{Tree}(P)$ //每个子集, 记住连接, 并且具有最大隙来保持其具有

//最大评分的节点集的分区

for (S 中的 x) $\{\{x\}.max = \{x\}\}$

开始

$P' = \text{clumper}(P)$

如果 $P == P'$

则

退出

否则

UPDATE_Qmax(P',P)

结束

UPDATE_Qmax(P',P)

for (p in P') {

 如果(!(p in P)) {

[0180] Subset pSubset = AddSubset(p,F);

 如果(Q(p) >= QP(C(p)))

 pSubset.最大分区 = p

 pSubset.Qmax = Q(p)

 否则

 pSubset.Qmax = QP(C(p))

 pSubset.最大分区 = MAX_UNION(C(p))

 }

}

MAX_UNION({Ni})

 返回 Ni.max 的联合

[0181] 当该过程终止时，F的根R的元素可包含它们的最大分区，其联合为S的F中的最优分区。

[0182] 分区选择模块1512寻找从属于森林F的将至少一个评分函数最大化的分区。例如，分区选择模块1512可选择从属于森林F的将评分函数QP最大化的分区。

[0183] 在各种实施例中，分区的每个子集(如本文所论述)可与其自己的评分相关联。例如，分区的每个子集可与不同的Q_Max评分相关联。分区选择模块1512可使用Q_Max评分从森林F的任何数量的不同分区选择独特元素的子集以生成并选择分区。

[0184] 例如，参见图19d，分区选择模块1512可基于评分函数从一个分区选择子集{A,B,C}，并从另一分区选择子集{D,E}、{F}和{G}。所选择的子集可然后形成(例如，生成)新选择的分区P1(例如，包括子集{A,B,C}、{D,E}、{F}和{G}的分区)。在该示例中，分区选择模块1512可利用Q_Max评分从第一分区选择子集{A,B,C}。在进一步的示例中，包括A、B或C中的任何元素的所有分区的每个子集可与单独的Q_Max评分相关联。包括A、B或C中的任何元素的所有集合的最大Q_Max评分为子集{A,B,C}。结果，在该示例中，分区选择模块1512选择该子集{A,B,C}。

[0185] 相似地，包括D、E、F或G中的任何元素的所有分区的每个子集可与单独的Q_Max评

分相关联。包括D、E、F或G中的任何元素的所有集合的最大Q_Max评分为子集{D,E}、{F}和{G} (即,与子集{D,E,F,G}、{D,E,F}和{G}相关联的Q_Max评分在与子集{D,E}、{F}和{G}的Q_Max评分比较时不是最大值)。结果,在该示例中分区选择模块1512选择子集{D,E}、{F}和{G}。

[0186] 本文中提到的评分函数的一个示例包括用于节点集S上加权图表的模块性评分。在一些实施例中,在子集内边缘的图表比例的子集(e和a)的模块性评分,其为与子集边界交叉的边缘的比例。最终评分可为 $e-a^2$ 。在各种实施例中,分区选择模块1512通过最大化该评分来选择和/或生成分区。模块性评分因子QP可为对该分区内的子集上的模块性评分之和。

[0187] 评分函数的另一示例为具有相关联的集合S的熵的变型:即,函数 $cls:S \rightarrow \{1,2,\dots,k\}$ (即,你具有集合并且每个元素具有一些有限标签)。对于S的子集,我们定义如果 $|s| \neq 0$,则 $p_i(s) = |\{s \text{ 中的 } x: cls(x) = i\}| / |s|$ 。那么 $Q(s) = \sum_{\text{classes } i} (p_i(s) * \log(p_i(s)))$ 。熵评分因子Q到分区评分因子的延伸QP由延伸属性(3)给出,其中 $A(p,q) = |p| / |q|$ 。即,对于分区P, $QP(P) = \sum_{\text{P中的 } p} (Q(p) * |p| / |U(P)|)$ 。通常人们要求最小化熵,并且通过最大化评分函数,这里的子集评分因子为传统熵评分的负值。

[0188] 数据控制模块1514被配置为从分区选择模块1512提供所选择的和/或所生成的分区。在各种实施例中,数据控制模块1514生成指示来自分区选择模块1512的所选择的分区和/或所生成的分区的报告。该报告可包括例如数据集、分区、子集、元素、数据集标识符、分区标识符、子集标识符、元素标识符等。在一些实施例中,报告可包括带有所选择节点的指示的图表(例如,参见图14),该所选择节点的(多个)成员包括来自分区选择模块1512的所选择的分区和/或所生成的分区的数据。

[0189] 图16是在一些实施例中的用于自动分组的示例流程图。在该示例中,自动分组模块214接收集合 $S = \{A,B,C,D,E,F,G\}$ 并执行自动分组以基于S识别森林的所选择分区。描述图16中步骤中的至少一些步骤的非限制性示例将使用在图17中描绘的图表来描述。在该示例中Q_Partition的实施例简单地为在每个子集上的Q_Subset评分的分区P的子集之和。例如,如果 $P = \{\{A,B,C\}, \{D\}, \{E,F\}, \{G\}\}$,那么 $Q_Partition(P) = Q_Subset(\{A,B,C\}) + Q_Subset(\{D\}) + Q_Subset(\{E,F\}) + Q_Subset(\{G\})$ 。

[0190] 在步骤1602中,数据结构模块1502接收集合S并且分区生成模块1504生成初始分区,其为集合 $S = \{A,B,C,D,E,F,G\}$ 的单元元素,即, $P_0 = \{\{A\}, \{B\}, \{C\}, \{D\}, \{E\}, \{F\}, \{G\}\}$ 。这在图17中被示为所描绘森林的底部行{1702}。

[0191] 在步骤1604中,Q_subset评分模块1506对分区P_0的每个子集计算Q_Subset评分。在该示例中,Q_subset评分模块1506利用值0.5将每个单元元素子集评分。该评分在图17中针对分区1702的每个子集被示为 $Q_Sub = 0.5$ 。

[0192] 在步骤1606中,Q_partition评分模块1510从构造的森林中的子集a的子元素计算P_0的每个子集a的最大分区。由于P_0中的子集a在森林中没有子元素,因此子集a的子元素的最大分区为其自身。即,对于P_0中的每个子集a,最大分区子元素(a) = a。

[0193] 在该示例中,Q_partition评分模块1510计算作为其自身的每个子集的最大分区。这在图17中针对分区1702的每个子集被示为对于子集{A}, $MaxP = \{A\}$,对于子集{C}, $MaxP = \{C\}$,对于子集{D}, $MaxP = \{D\}$,对于子集{E}, $MaxP = \{E\}$,对于子集{F}, $MaxP = \{F\}$,以及

对于子集 $\{G\}$, $MaxP = \{G\}$ 。

[0194] 在步骤1608中, Q_{max} 评分模块对 P_0 的每个子集计算 Q_{max} 。回调 (recall) 其是由于对于 P_0 中的每个子集 a , P_0 的子集没有任何子元素,

[0195] $Q_{Max}(a) = \max(Q_{Subset}(a), Q_{Partition}(\text{最大分区子元素}(a)))$

[0196] $= \max(Q_{Subset}(a), Q_{Partition}(a))$

[0197] $= \max(Q_{Subset}(a), Q_{Subset}(a)) = Q_{Subset}(a)$

[0198] $= 0.5$

[0199] 在该示例中, Q_{max} 评分模块1508以值0.5将每个子集进行评分。该 Q_{max} 评分在图17中针对分区1702的每个子集被示为 $Q_{Sub} = 0.5$ 。

[0200] 在步骤1610中, 我们可选地记录 P_0 中的每个子集 a 的最大分区为生成该子集的 Q_{Max} 的子集 a 的分区。因此我们在该初始分区中记录: 最大分区 $(a) = a$ 。

[0201] 在步骤1612中, 数据结构模块1502计算下个分区 P_1 (在图17中标记为1704的行)。即, 在该示例中, 数据结构模块1502将子集 $\{A\}$ 和 $\{B\}$ 群集成子集 $\{A, B\}$, 并将子集 $\{D\}$ 和 $\{E\}$ 群集成子集 $\{D, E\}$ 。数据结构模块1502在分区 P_1 中保留来自分区 P_0 的子集 $\{C\}$ 、 $\{F\}$ 和 $\{G\}$ 。

[0202] 在各种实施例中, 数据结构模块1502可确定系统是否结束和/或是否将计算新分区。应认识到数据结构模块1502可基于任何数量的方式来进行确定。在一些实施例中, 数据结构模块1502确定下个生成的分区是否等于先前分区。如果两个分区相等 (例如, 具有相同子集), 则方法可终止, 否则方法可继续到步骤1614。

[0203] 在一些实施例中, 如果发现预定数量的根, 等等, 则数据结构模块1502在生成预定数量的分区后终止方法。在各种实施例中, 如果预定数量的子集在计算的分区中存在, 则数据结构模块1502可终止方法。在另一示例中, 数据结构模块1502可在预定时段、预定存储器使用时段后或基于任何阈值 (例如, 基于接收的数据量计算的阈值) 终止方法。

[0204] 在步骤1614中, Q_{subset} 评分模块1506对分区 P_1 的每个子集计算 Q_{Subset} 评分。在该示例中, Q_{subset} 评分模块1506计算 $Q_{Subset}(\{A, B\}) = 0.5$ 和 $Q_{Subset}(\{D, E\}) = 2$ 。如在上面描述1604的段落中论述, 每个单元元素子集的 Q_{Subset} 为0.5 (例如, 在1704中的单元元素子集的先前 Q_{Subset} 评分保持不从1702改变)。这些评分与每个子集相关联并且在图17中被可视化为1704中的 Q_{Sub} 。

[0205] 在步骤1616中, $Q_{partition}$ 评分模块1510然后计算在 P_1 的每个子集的子元素处的最大分区。子集 $\{C\}$ 、 $\{F\}$ 和 $\{G\}$ 的子元素的最大分区再次为原始的单元元素子集。子元素 $\{A, B\}$ 的最大分区为包括 $\{A, B\}$ 的子元素的最大分区的集合, 即如图17中的分区1704中所描绘的 $\{\{A\}, \{B\}\}$ 。相似地, $\{D, E\}$ 的子元素的最大分区为集 $\{\{D\}, \{E\}\}$, 同样如图17中的分区1704中所描绘。

[0206] 在步骤1618中, Q_{max} 评分模块1508对 P_1 的每个子集计算 Q_{max} 。回调 $Q_{Max}(a) = \max(Q_{Subset}(a), Q_{Partition}(\text{最大分区子元素}(a)))$ 。对于子集 $\{A, B\}$:

[0207] $Q_{Max}(\{A, B\}) = \max(Q_{Subset}(\{A, B\}), Q_{Partition}(\{\{A\}, \{B\}\}))$

[0208] $= \max(.5, Q_{Subset}(\{A\}) + Q_{Subset}(\{B\}))$

[0209] $= \max(0.5, 1)$

[0210] $= 1$

[0211] 对于子集 $\{D, E\}$:

[0212] $Q_Max(\{D,E\}) = \max(Q_Subset(\{D,E\}), Q_Partition(\{\{D\}, \{E\}\}))$

[0213] $= \max(2, Q_Subset(\{D\}) + Q_Subset(\{E\}))$

[0214] $= \max(2, 1)$

[0215] $= 2。$

[0216] 如在图17的分区1704中显示, {A,B}的 Q_Max 为1和{D,E}的 Q_Max 为2。分区1704中单元素{C}、{F}和{G}的 Q_Max 保持与分区1702中的相应子集一致。即, {C}、{F}和{G}中的每个的 Q_Max 为0.5。

[0217] 在步骤1620中,我们任选地记录产生 Q_Max 评分的 P_1 中每个子集a的最大分区。如上文和图17中所见,最大分区({A,B}) = {{A}, {B}}和最大分区({D,E}) = {D,E}。

[0218] 现在重复步骤1612。数据结构模块1502计算在图17中描绘为行(分区)1706的下一个分区 P_2 。在各种实施例中,数据结构模块1502可确定系统是否结束和/或是否将计算新分区。应认识到数据结构模块1502可基于任何数量的方式执行确定。

[0219] 在步骤1614中, Q_subset 评分模块1506对分区 P_2 的每个子集计算 Q_Subset 评分。在该示例中, Q_subset 评分模块1506计算 $Q_Subset(\{A,B,C\}) = 2$ 和 $Q_Subset(\{D,E,F\}) = 1.5$ 。再次, $Q_Subset(\{G\}) = 0.5$ 。这些评分与每个子集一起记录并且在图17中在分区1706中可视化。

[0220] 在步骤1616中, $Q_partition$ 评分模块1510然后计算在 P_2 的每个子集的子元素处的最大分区。子元素{G}的最大分区为子集{G}。子元素{A,B,C}的最大分区为由{A,B,C}的子元素的最大分区构成的集,即{最大分区({A,B}), 最大分区({C})} = {{A}, {B}, {C}}。相似地, {D,E,F}的子元素的最大分区为集合{最大分区({D,E}), 最大分区({F})} = {{D,E}, {F}}。

[0221] 这在图17中为分区1706的每个子集示为对于子集{A,B,C}, $MaxP = \{A,B,C\}$, 对于子集{D,E,F}, $MaxP = \{D,E,F\}$, 并且对于子集{G}, $MaxP \{G\}$ 。

[0222] 在步骤1618中, Q_max 评分模块1508对 P_2 的每个子集计算 Q_max 。回调 $Q_Max(a) = \max(Q_Subset(a), Q_Partition(\text{最大分区子元素}(a)))$ 。对于子集{A,B,C}:

[0223]

$$Q_Max(\{A,B,C\}) = \max(Q_Subset(\{A,B,C\}), Q_Partition(\{\{A\}, \{B\}, \{C\}\}))$$

[0224]

$$\begin{aligned} &= \max(2, Q_Subset(\{A\}) + Q_Subset(\{B\}) + \\ &Q_Subset(\{C\})) \\ &= \max(2, 1.5) \\ &= 2 \end{aligned}$$

[0225] 对于子集{D,E,F}:

[0226] $Q_Max(\{D,E,F\}) = \max(Q_Subset(\{D,E,F\}), Q_Partition(\{\{D,E\}, \{F\}\}))$

[0227] $= \max(1.5, Q_Subset(\{D,E\}) + Q_Subset(\{F\}))$

[0228] $= \max(1.5, 2.5)$

[0229] $= 2.5$

[0230] 如在图17的分区1706中显示, {A,B,C}的 Q_Max 为2和{D,E,F}的 Q_Max 为2.5。分区1706中单元素的 Q_Max 保持与分区1704中的相应子集一致。即, $Q_Max \{G\}$ 为0.5。

[0231] 在步骤1620中,我们任选地记录产生Q_Max评分的P_2中每个子集a的最大分区。如上所见,最大分区($\{A,B,C\}$)= $\{\{A,B,C\}\}$ 和最大分区($\{D,E,F\}$)= $\{\{D,E\},\{F\}\}$ 。

[0232] 现在重复步骤1612。数据结构模块1502计算在图17中描绘为行(分区)1708的下一个分区P_3。数据结构模块1502可确定系统是否结束和/或是否将计算新分区。

[0233] 在步骤1614中,Q_subset评分模块1506对分区P_3的每个子集计算Q_Subset评分。在该示例中,Q_subset评分模块1506计算 $Q_Subset(\{A,B,C\})=2$ 和 $Q_Subset(\{D,E,F,G\})=1$ 。这些评分与每个子集一起记录并且在图17中在分区1708中可视化。

[0234] 在步骤1616中,Q_partition评分模块1510计算在P_3的每个子集的子元素处的最大分区。子元素 $\{A,B,C\}$ 的最大分区为由 $\{A,B,C\}$ 的子元素的最大分区构成的集,即 $\{\text{最大分区}(\{A,B,C\})\}=\{A,B,C\}$ 。相似地, $\{D,E,F,G\}$ 的子元素的最大分区为集 $\{\text{最大分区}(\{D,E,F\}),\text{最大分区}(\{G\})\}=\{\{D,E\},\{F\},\{G\}\}$ 。

[0235] 这在图17中为分区1708的每个子集示为对于子集 $\{A,B,C\}$, $\text{MaxP}=\{A,B,C\}$,并且对于子集 $\{D,E,F,G\}$, $\text{MaxP}=\{\{D,E\},\{F\},\{G\}\}$ 。

[0236] 在步骤1618中,Q_max评分模块1508对P_3的每个子集计算Q_max。回调 $Q_Max(a)=\max(Q_Subset(a),Q_Partition(\text{最大分区子元素}(a)))$ 。对于子集 $\{A,B,C\}$:

[0237] $Q_Max(\{A,B,C\})=\max(Q_Subset(\{A,B,C\}),Q_Partition(\{A,B,C\}))$

[0238] $=\max(2,Q_Subset(\{A,B,C\}))$

[0239] $=2$

[0240] 对于子集 $\{D,E,F,G\}$:

[0241]

$$\begin{aligned} Q_Max(\{D,E,F,G\}) &= \max(Q_Subset(\{D,E,F,G\}),Q_Partition(\{\{D,E\},\{F\},\{G\}\})) \\ &= \max(1,Q_Subset(\{D,E\}) + Q_Subset(\{F\}) + Q_Subset(\{G\})) \\ &= \max(1.5,3) \\ &= 3 \end{aligned}$$

[0242] 如在图17的分区1708中显示, $\{A,B,C\}$ 的Q_Max为2和 $\{D,E,F,G\}$ 的Q_Max为2.5。

[0243] 在步骤1620中,我们任选地记录产生Q_Max评分的P_3中每个子集a的最大分区。如上所见,最大分区($\{A,B,C\}$)= $\{\{A,B,C\}\}$ 和最大分区($\{D,E,F,G\}$)= $\{\{D,E\},\{F\},\{G\}\}$ 。

[0244] 尽管未在方法1600中示出,但方法可以继续。例如,分区选择模块1512可识别和/或生成优选分区,从该优选分区最大化一个或多个评分函数。在该示例中,优选分区为最大分区。如上面刚才论述,P_3中每个子集的最大分区为如上所述的最大分区($\{A,B,C\}$)= $\{\{A,B,C\}\}$ 和最大分区($\{D,E,F,G\}$)= $\{\{D,E\},\{F\},\{G\}\}$ 。分区选择模块1512可识别和/或生成自动分组的分区 $\{\{A,B,C\},\{\{D,E\},\{F\},\{G\}\}\}$ 。

[0245] 数据控制模块1514可在报告中提供已识别和/或已生成的自动分组分区,和/或识别数据或图表中的自动分组分区。

[0246] 图18为示例性数字装置1800的框图。数字装置1800包括通信耦接到总线1814的数据处理装置(例如处理器)1802、存储器系统1804、存储系统1806、通信网络接口1808、I/O接

口1810,以及显示接口1812。处理器1802被配置为执行可执行指令(例如程序)。在一些实施例中,处理器1802包括能够处理可执行指令的电路系统或任何处理器。

[0247] 存储器系统1804是被配置为存储数据的任何存储器。存储器系统1804的一些示例为存储装置,诸如RAM或ROM。存储器系统1804可包括高速缓存存储器。在各种实施例中,数据存储于存储器系统1804内。在存储器系统1804内的数据可被清除或最终传递到存储系统1806。

[0248] 存储系统1806是被配置为检索和存储数据的任何存储设备。存储系统1806的一些示例为闪存驱动器、硬盘驱动器、光盘驱动器和/或磁带。存储系统1806可包括非暂时性媒体。在一些实施例中,数字装置1806包括RAM形式的存储器系统1804和闪存数据形式的存储系统1806。存储器系统1804和存储系统1806两者均包括计算机可读媒体,其可存储可由包括处理器1802的计算机处理器执行的指令或程序。

[0249] 通信网络接口(通信的网络接口)1808可经由链路1816耦接到网络(例如通信网络110)。例如,通信网络接口1808可支持经由以太网连接、串口连接、并口连接或ATA连接的通信。通信网络接口1808也可支持无线通信(例如,802.17a/b/g/n、WiMax)。显然通信网络接口1808可支持许多有线和无线标准。

[0250] 任选的输入/输出(I/O)接口1810为从用户接收输入和输出数据的任何装置。任选的显示接口1812为被配置为向显示器输出图形和数据的任何装置。在一个示例中,显示接口1812为图形适配器。应认识到并非所有的数字装置1800包括I/O接口1810或显示接口1812。

[0251] 数字装置1800的硬件元件不限于图18中描述的那些。数字装置1800可包括比所描绘的那些更多或更少的硬件元件。进一步地,硬件元件可共享功能并仍在本文所述的各种实施例之内。在一个示例中,编码和/或解码可由处理器1802和/或由协处理器,诸如位于图形处理单元(GPU)上的处理器执行。

[0252] 上述功能和部件可由存储在存储媒体诸如计算机可读媒体(例如,非暂时性计算机可读媒体)上的指令构成。指令可由处理器检索和执行。指令的一些示例为软件、程序代码和固件。存储媒体的一些示例为存储器装置、磁带、磁盘、集成电路和服务。当指令由处理器执行以引导处理器根据本发明的实施例操作时,指令为可操作的。

[0253] 本发明参考示例性实施例在上文中描述。可在不背离本发明的更广泛保护范围的情况下做出各种修改并且可使用其他实施例。因此,对示例性实施例的这些和其他变化旨在被本发明涵盖。

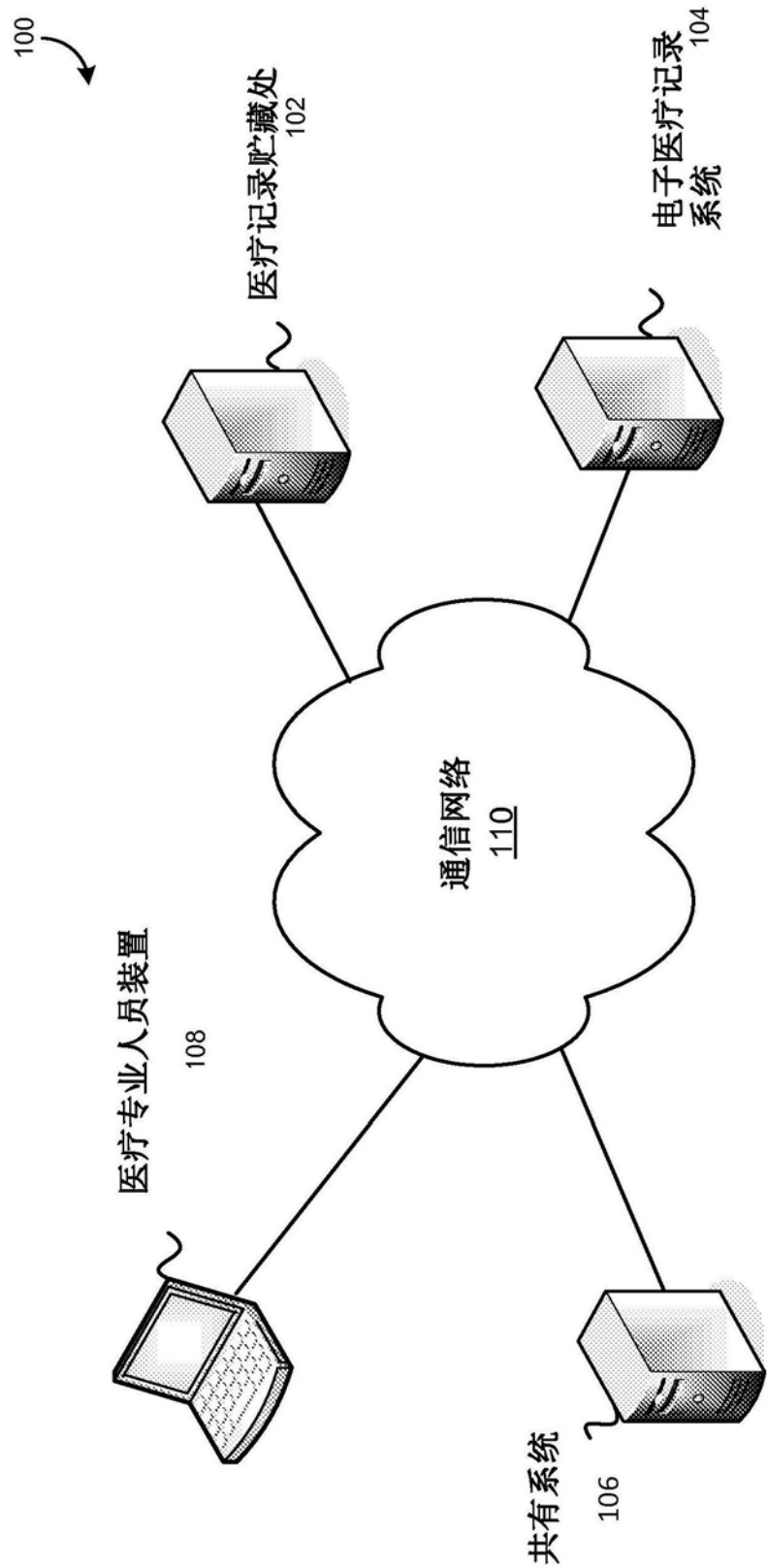


图1

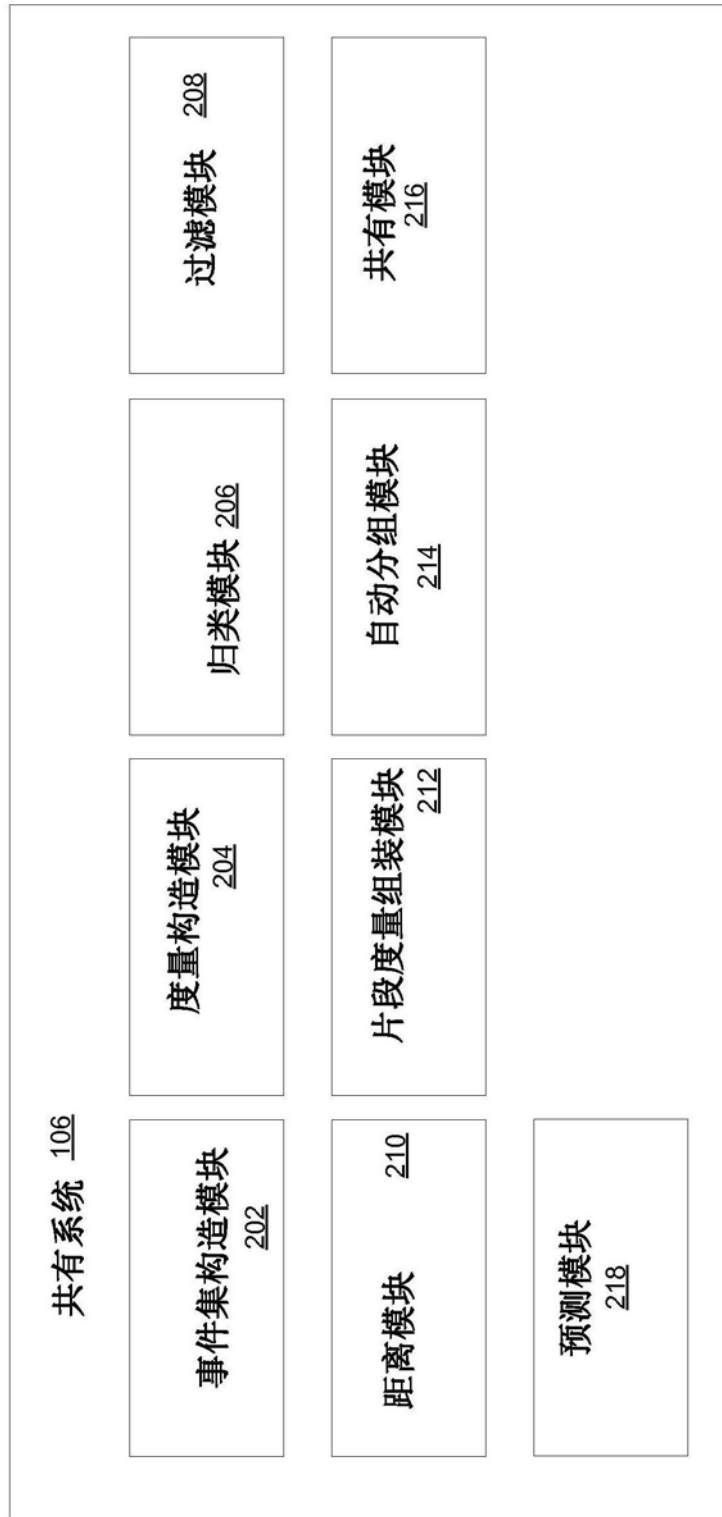


图2

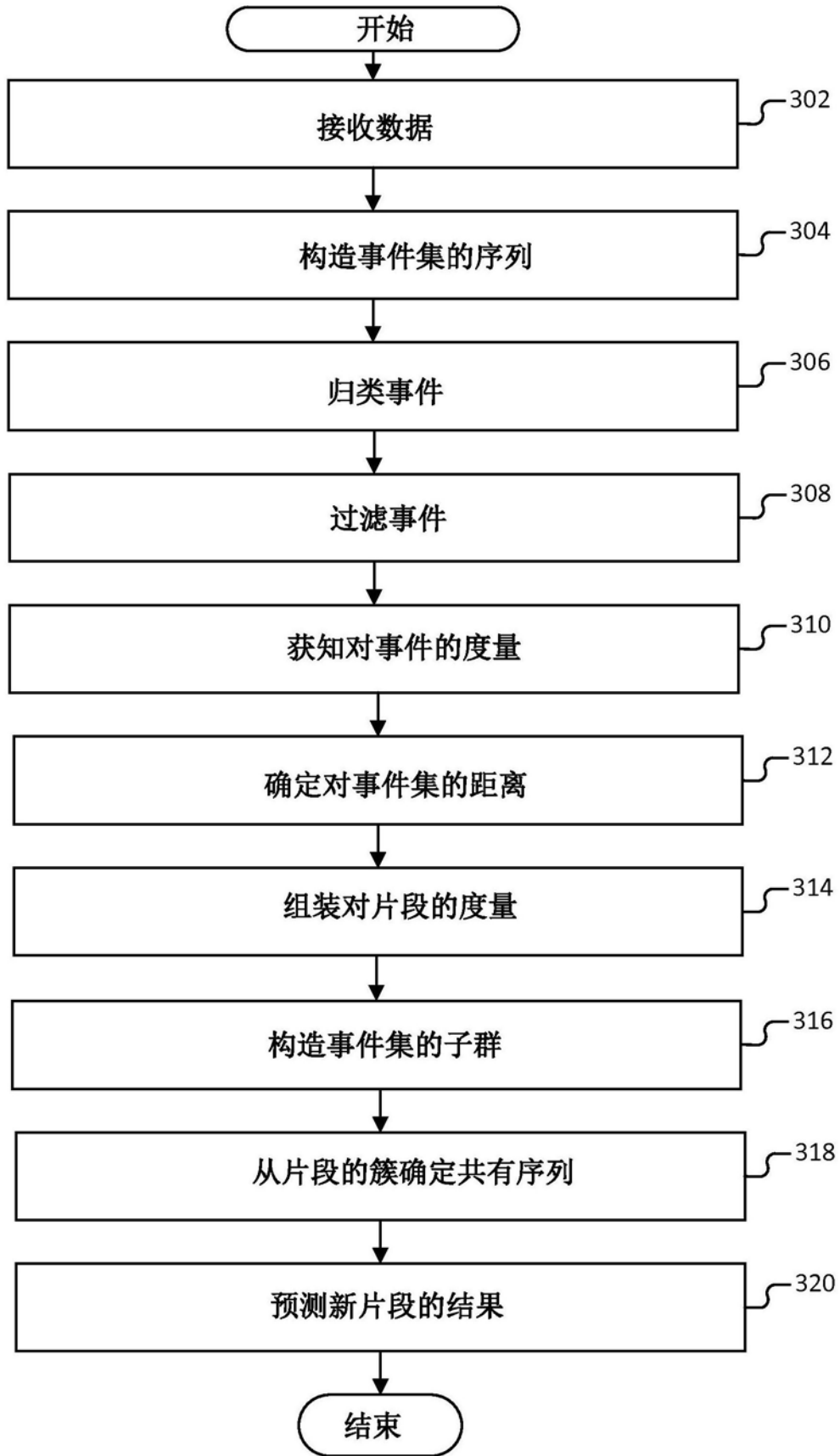


图3

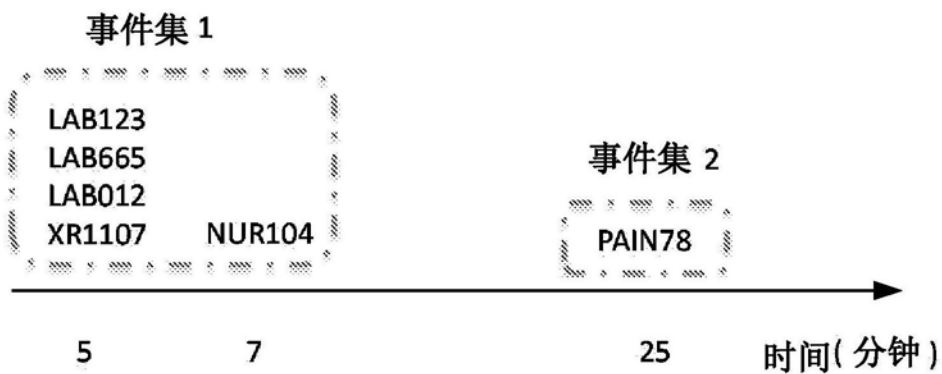


图4

	LAB	XR	...	痛 (PAIN)
LAB123	0.213	0.00012	...	0.5
...
属类 LAB	0.647	0.12	...	0.01
XR1107	0.476	0.066	...	0.0003
...

图5

G1: LAB123, LAB665, LAB012, 属类XR

G2: LAB123, LAB665, LAB013, XR1107

图6

a) ABC和abc的DTW匹配为
 $A^*BC \leftrightarrow ab^*c$:

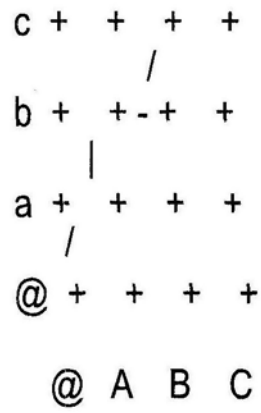


图7a

b) ABC和abc的DTW匹配为
 $*AB^*C^* \leftrightarrow a^*b^*c$:

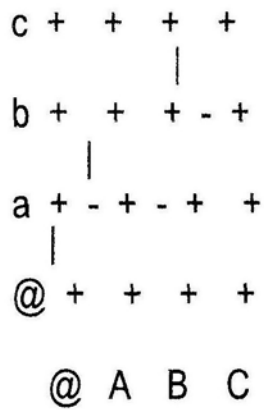
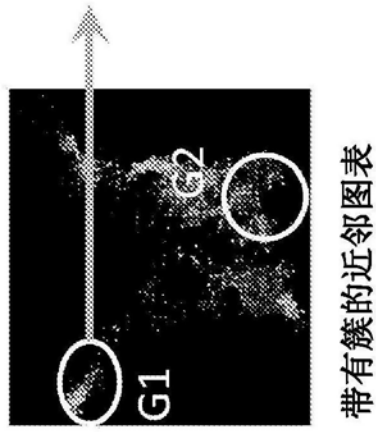
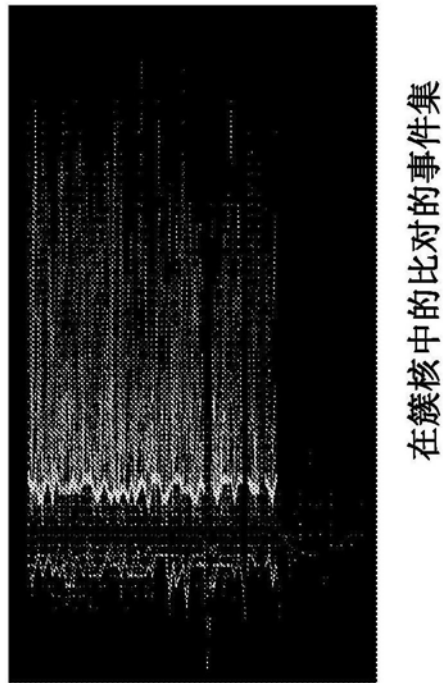


图7b



带有簇的近邻图表

图8a



在簇核中的比对的事件集

图8b

中值时间 行为	
5 分钟	LAB123
5 分钟	LAB665
5 分钟	XR1107
7 分钟	NUR104
...	...

簇核的共有片段

图8c

时间（相对于
外科手术的分
钟）

902	904	906	908	910	912
父类别	类别	代码	间隔	描述	描述
-17114 命令 (Orders)	EKG	EKG1	未知	未知	EKG 12-引导
-17114 实验室测试 (Labs)	LAB	LAB1725	未知	未知	具有差异的CBC
-17114 实验室测试	LAB	LAB276	未知	未知	血型与抗体筛检 (TYPE AND SCREEN)
-17114 实验室测试	LAB	LAB2038	未知	未知	具有反射培养的尿检
-17114 实验室测试	LAB	LAB15	未知	未知	基础代谢检查
-17114 实验室测试	LAB	LAB347	未知	未知	尿检
-17114 实验室测试	LAB	LAB2338	未知	未知	MRSA 主动监测
-17114 实验室测试	LAB	LAB320	未知	未知	PROTIME-INR
-17114 实验室测试	LAB	LAB325	未知	未知	PTT
-143 命令	DIET	DIET40	未知	未知	饮食 NPO
-143 命令	IVT	IVT3	未知	未知	插入外周静脉
-143 命令	NUR	NUR542	未知	未知	检查知情同意书
-143 命令	COD	COD2	未知	未知	全码 (FULL CODE)
-143 命令	NUR	NUR2007	未知	未知	尿检
-143 命令	NUR	NUR185	未知	未知	看护交流
-143 命令	NUR	NUR1016	未知	未知	准备手术部位
-143 命令	ADT	ADT1	未知	未知	住院
-143 实验室测试	LAB	LAB2776	未知	未知	仅抽血
-111 药物 (Meds)	ELECT/CALORIC/H2O	LACTATED RINGERS IV	未知	未知	LACTATED RINGERS IV
-111 药物	ANTIARTHRITICS	CELECOXIB 200 MG CAPS	未知	未知	CELECOXIB 200 MG CAPSULE
-111 药物	ANALGESICS	OXYCODONE ER 10 MG T	未知	未知	oxyCODONE CR (OXYCONTIN) tablet 20 mg
-111 药物	CNS DRUGS	PREGABALIN 75 MG CAP	未知	未知	PREGABALIN 75 MG CAPSULE
-82 药物	ELECT/CALORIC/H2O	LACTATED RINGERS IV	未知	未知	LACTATED RINGERS IV
-82 药物	ANESTHETICS	MIDAZOLAM 1 MG/ML	未知	未知	MIDAZOLAM 1 MG/ML INJECTION
-41 命令	ADT	ADT7	未知	未知	转院
-34 实验室测试	LAB	LAB2038	未知	未知	具有反射培养的尿检
-27 药物	ANTIBIOTICS	CEFAZOLIN 2000 MG INF	未知	未知	CEFAZOLIN 2000 MG INFUSION SO
-14 实验室测试	LAB	LAB239	未知	未知	尿培养
-14 实验室测试	LAB	LAB347	未知	未知	尿检
32 其他	NUR	NUR2007	未知	未知	方案分析生命体征
32 其他	NUR	NUR502	未知	未知	教育、暂时停烟和SECO

图9

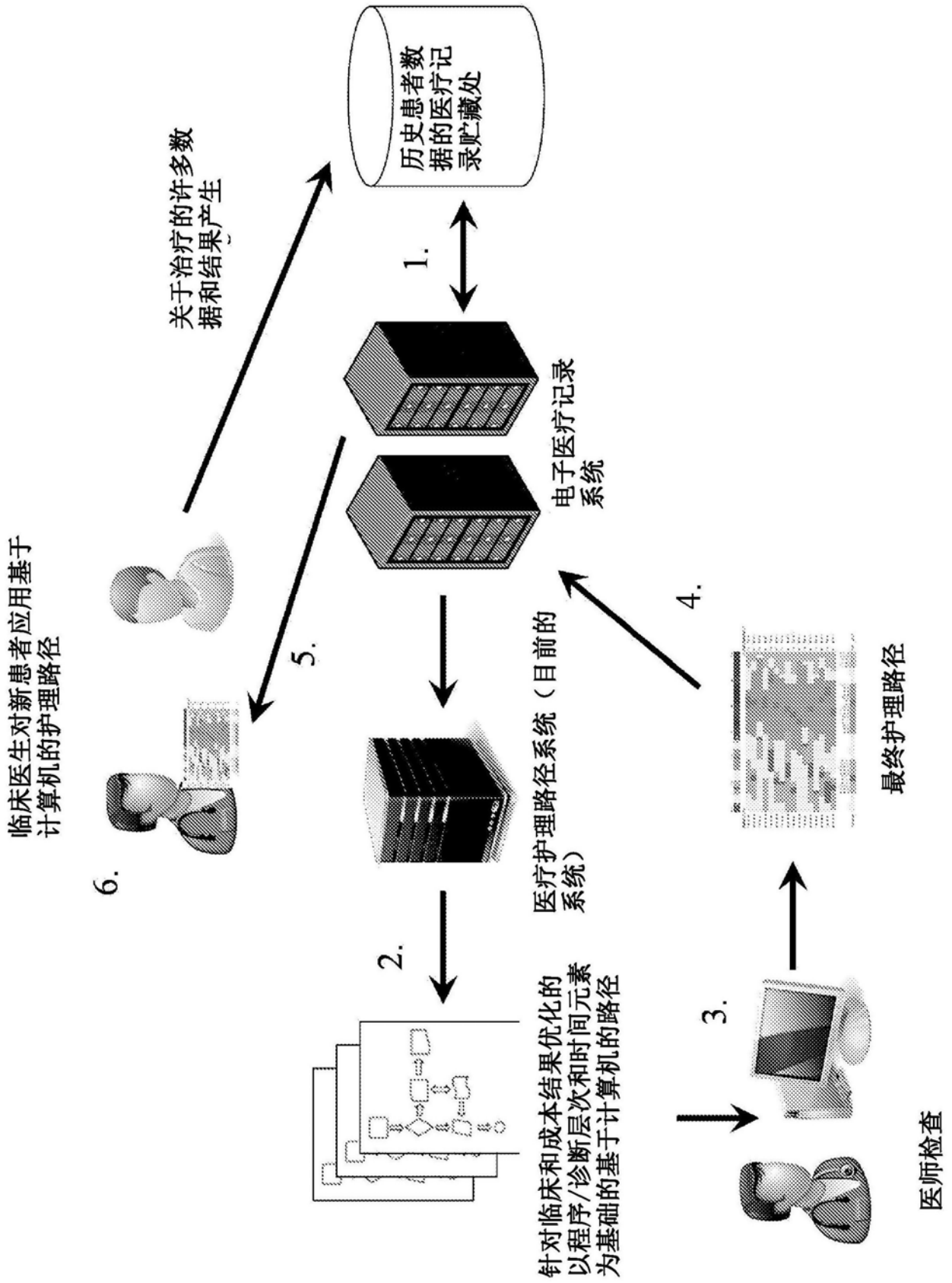


图10

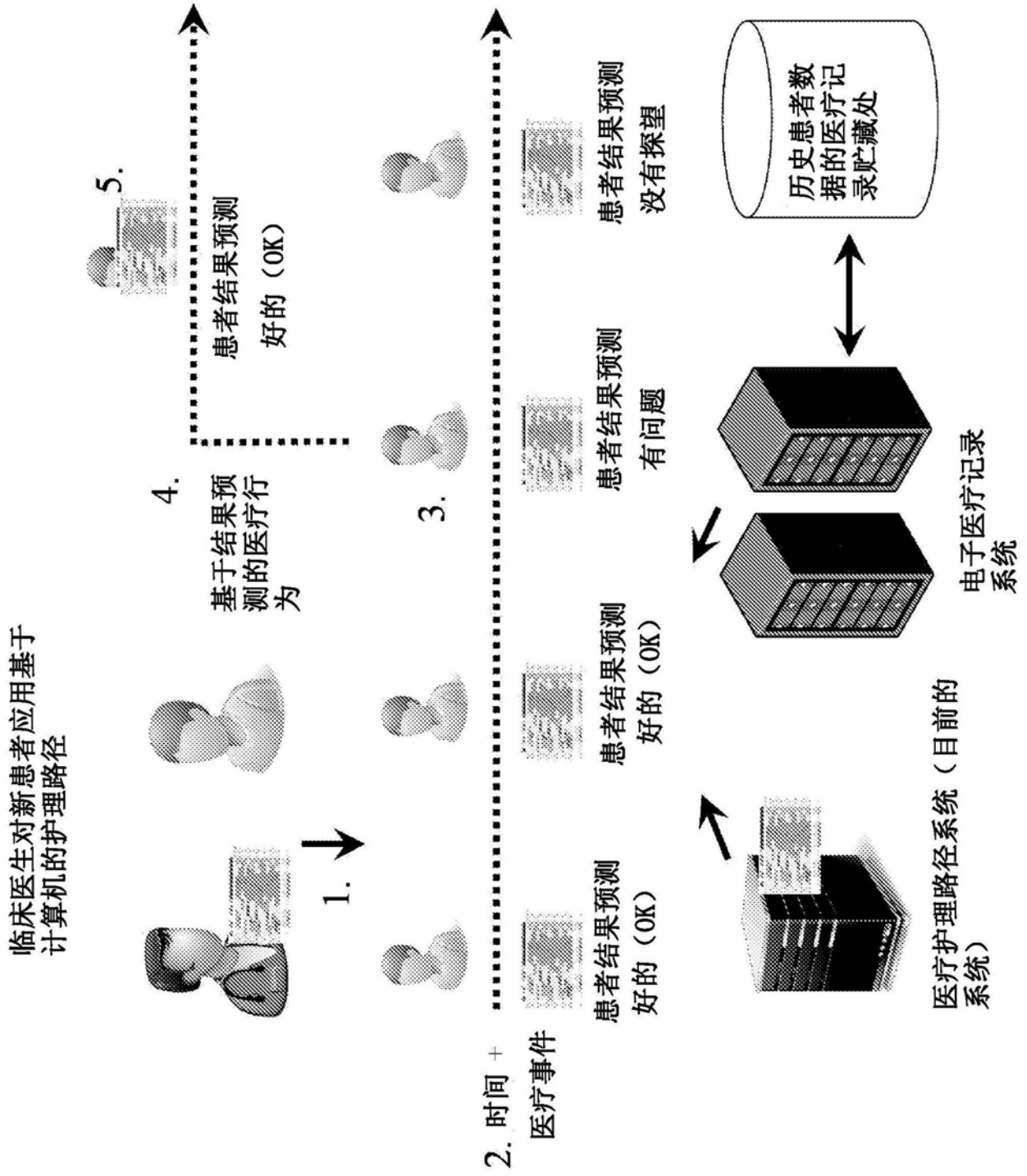


图11

1206

全膝替换		最好								
< 最好	1202	2.00	-6hr	外科手术	6hr	12hr	18hr	1d	1d 6hr	最好
> 最好	1202	2.00	停留	外科手术	6hr	12hr	18hr	1d	1d 6hr	最好
	365085_d...	2.00								未知 EKG
	236480_0...	2.00								ekg 12- 引导
	176221_0...	2.00								PRE OP 11天21小时14分
	222032_9...	2.00								未知 LAB
	230388_e...	2.00								具有差异的cbc
	217989_0...	2.00								PRE OP 11天21小时14分
	235934_9...	2.00								未知 LAB
	229147_a...	2.00								具有差异的cbc
	243632_3...	2.00								PRE OP 11天21小时14分
	177180_d...	2.00								未知 LAB
	225249_9...	2.00								做血型与抗体筛检
	246669_1...	2.00								PRE OP 11天21小时14分
	219680_d...	2.00								未知 LAB
	230508_a...	2.00								具有反射培养的尿检
	235202_7...	2.00								PRE OP 11天21小时14分
	184276_6...	2.00								未知 LAB
	243687_0...	2.00								基础代谢检查
	246515_e...	2.00								PRE OP 11天21小时14分
64 患者 (总计645)		>								未知 LAB

1204

图12

< 全膝替换:最好 最好 1306		1310					
		外科	手术	6hr	12hr		
> Best 1302	2.00					最好 成本:\$1,913,03 停留:2.00	NUR 2100 实例
365985_d...	2.00					未知 EKG ekg 12-引导 1308	看护交流 187 实例
236480_0...	2.00					PRE OP 11天21小时 14分	方案分析生命体征
176221_0...	2.00					未知 LAB 具有差异的cbc	180 实例
222032_9...	2.00					PRE OP 11天21小时14分	神经/脉管检查
230388_e...	2.00					未知 LAB 做血型与抗体筛检	119 实例
217989_0...	2.00					PRE OP 11天21小时14分	教育、戒烟和避免 二手烟
235934_9...	2.00					未知 LAB 具有反射培养的尿检	85 实例
229147_a...	2.00					PRE OP 11天21小时14分	护理刺激性肺量测定法
243632_3...	2.00					未知 LAB 基础代谢检查	85 实例
177180_d...	2.00					PRE OP 11天21小时14分	换衣 82 实例
225249_9...	2.00					未知 LAB 尿检	教育、刺激性肺量 测定法
246669_1...	2.00						
219680_d...	2.00						
230508_a...	2.00						
235202_7...	2.00						
184276_6...	2.00						
243687_0...	2.00						
246515_e...	2.00						
64 患者 (总计645)	>						

图13

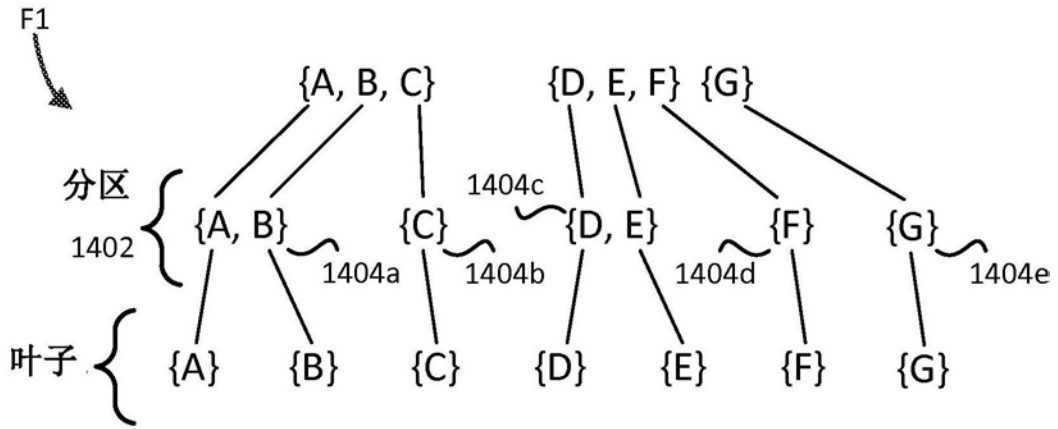


图14a

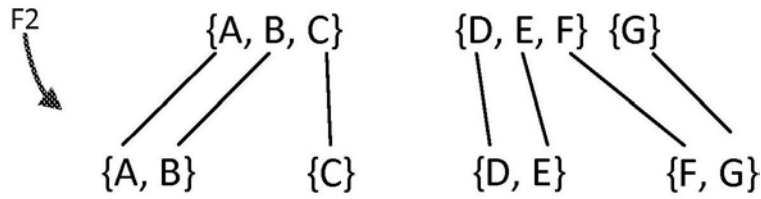


图14b

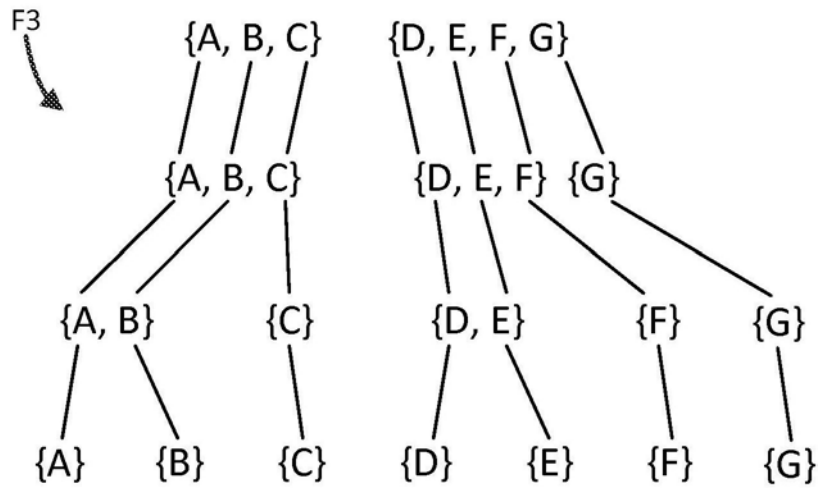


图14c

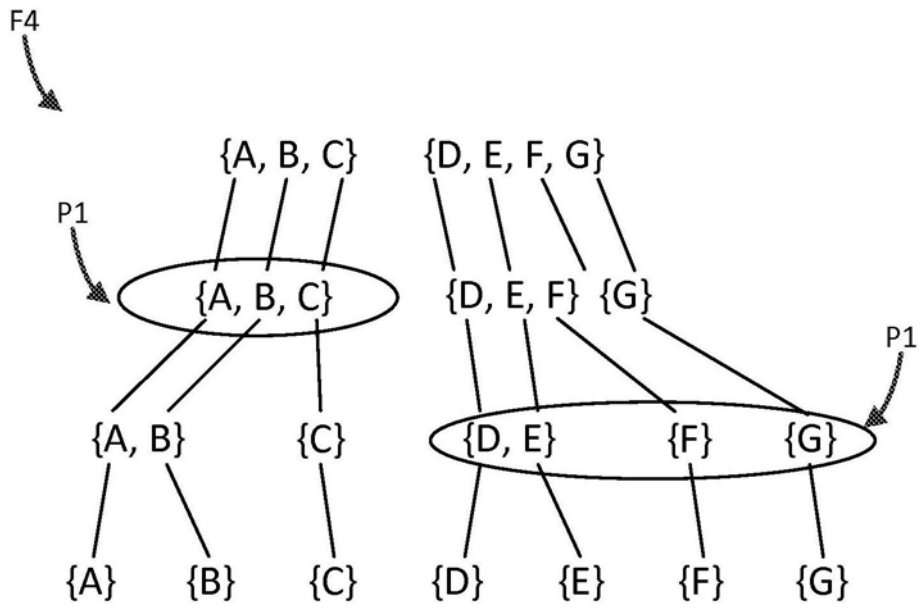


图14d



图15

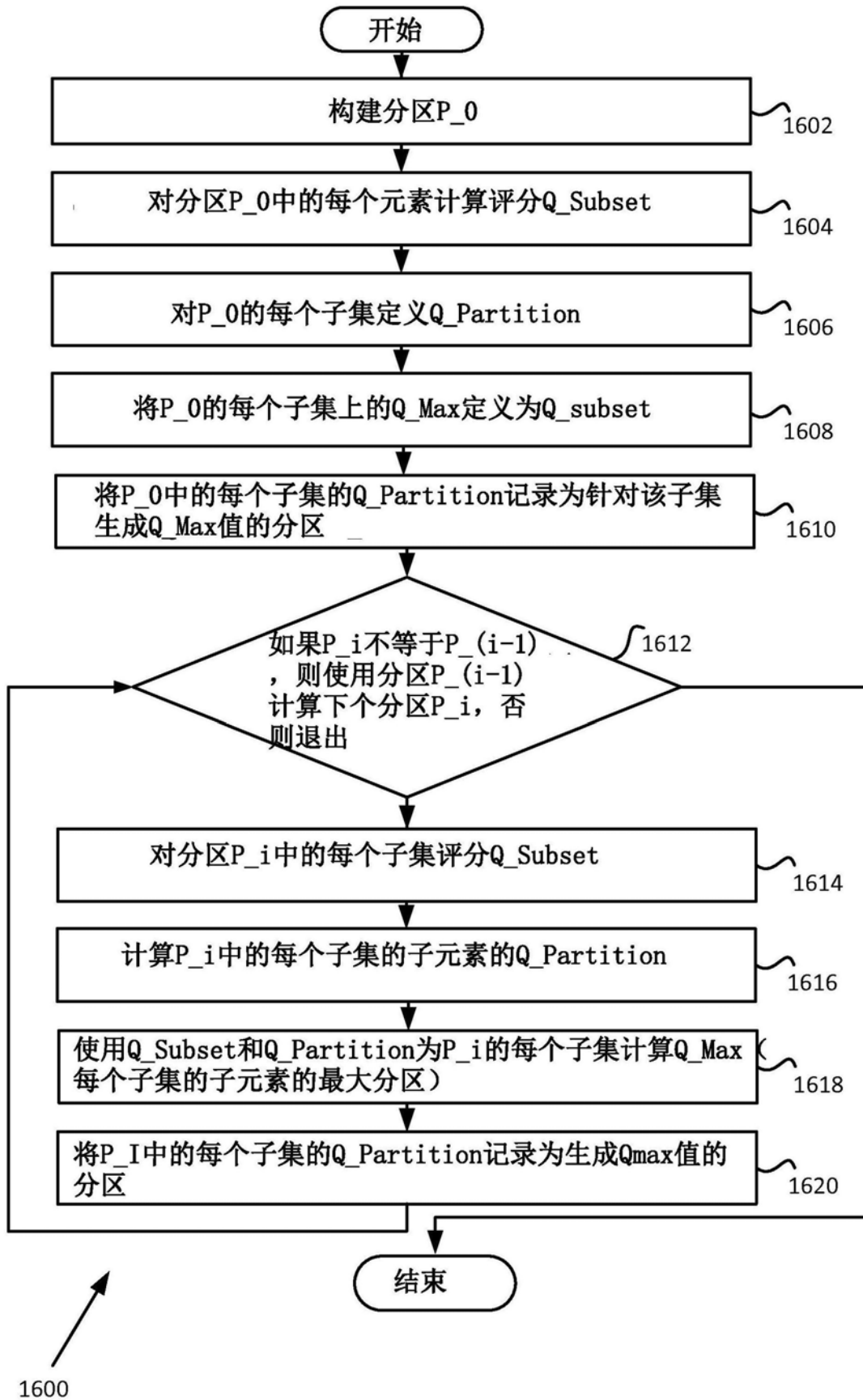


图16

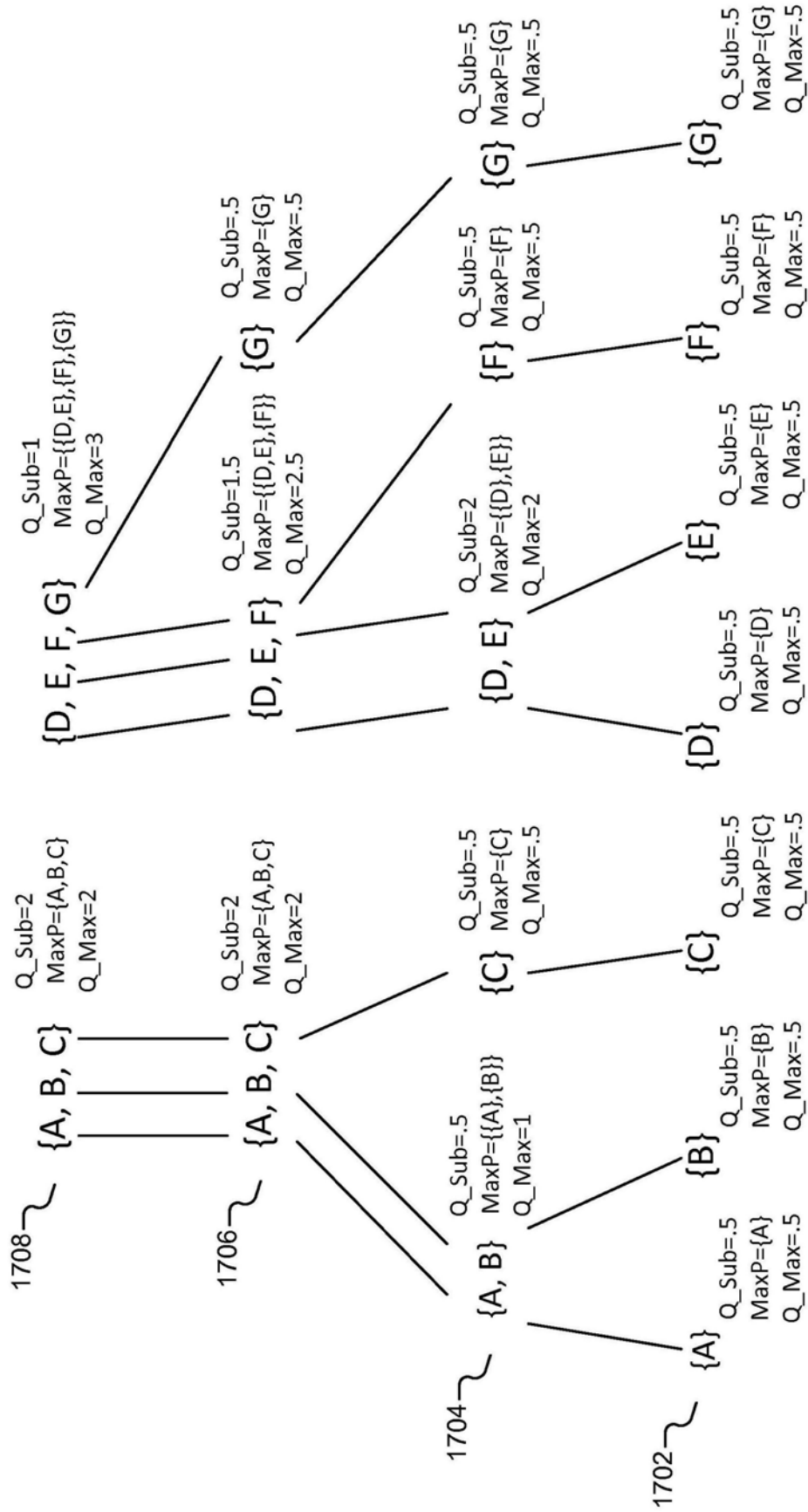


图17

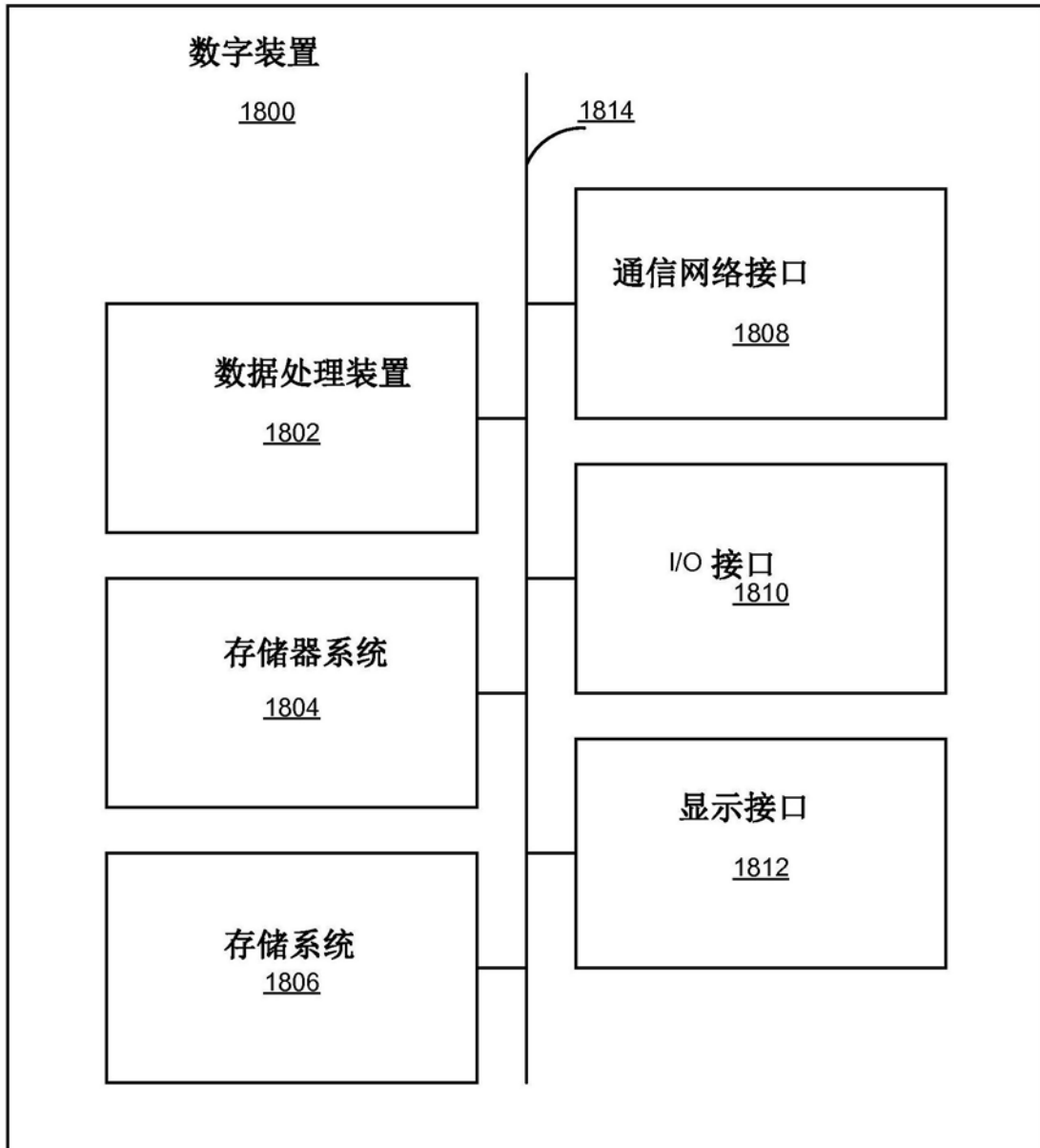


图18