



(12) 发明专利申请

(10) 申请公布号 CN 103440421 A

(43) 申请公布日 2013. 12. 11

(21) 申请号 201310391339. 4

(22) 申请日 2013. 08. 30

(71) 申请人 上海普之康健康管理有限公司  
地址 200120 上海市浦东新区陆家嘴东路  
161 号招商局大厦 3105

(72) 发明人 张明延

(74) 专利代理机构 北京邦信阳专利商标代理有  
限公司 11012

代理人 金玺

(51) Int. Cl.  
G06F 19/00(2011. 01)

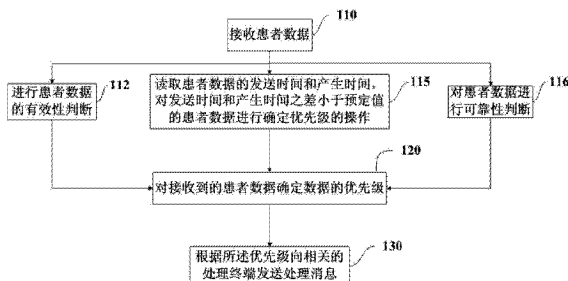
权利要求书2页 说明书6页 附图1页

(54) 发明名称

医学数据处理方法和系统

(57) 摘要

本发明提供了一种医学数据处理方法,所述方法包括:接收患者数据;将接收到的患者数据确定优先级;根据所述优先级处理终端发送相关的医学处理消息,以合理匹配到对应服务流程、优先级确定模块、处理消息发送模块。采用本发明的医学数据处理方法和处理系统后,在相同诊疗时间及病患数量情况及相同复杂程度下,由于优先级的设定,在不同级别患者中的医疗资源分配将更合理化,也使得错诊率、误诊率大大降低。基于该智能平台的拓展性设计,该平台可以整合不同的物联网设备用于不同的专科疾病的管理。



1. 一种医学数据处理方法,其特征在于,所述方法包括:
  - a. 接收患者数据;
  - b. 将接收到的患者数据确定优先级;
  - c. 根据所述优先级向特定处理终端发送相关的处理消息,以匹配相应的服务流程。
2. 根据权利要求1所述的方法,其特征在于,所述患者数据包括以下信息:疾病种类、检测参数、诊断时间、就诊地点、诊断医生、病历类型、数据输入方式、确诊时间、数据产生时间、数据发送时间、确诊机构。
3. 根据权利要求2所述的方法,其特征在于,在步骤a与步骤b之间还包括步骤a11:读取患者数据中的数据产生时间和数据发送时间,当数据产生时间和数据发送时间之差小于预定值时,则该患者数据进入步骤b进行处理。
4. 根据权利要求1或3所述的方法,其特征在于,在步骤a与步骤b之间还包括步骤a12:对患者数据进行有效性分析,当数据为有效数据时,则该患者数据进入步骤b进行处理。
5. 根据权利要求1所述的方法,其特征在于,在步骤a与步骤b之间还包括步骤a13:对患者数据进行可靠性判断,可靠性高的患者数据进入步骤b进行处理。
6. 根据权利要求4所述的方法,其特征在于,在步骤a与步骤b之间还包括步骤a13:对患者数据进行可靠性判断,可靠性高的患者数据进入步骤b进行处理。
7. 根据权利要求1所述的方法,其特征在于,步骤c为:根据所述优先级,通过短信、电子邮件、电话中任何一种方式或多种方式的组合向特定处理终端发送相关的处理消息。
8. 根据权利要求1所述的方法,其特征在于,所述服务流程包括以下任意一种:短信用药提醒、生活习性短信提醒、一对多的现场教育、初级医生药物调整、初级医生与专科会诊、专科远程面诊、专科医生面诊。
9. 根据权利要求2所述的方法,其特征在于,步骤b包括:将接收到的检测参数加入到多元线性回归模型,所述多元线性回归模型为:
$$Y_i = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_q X_q + \epsilon$$
,其中, $X_1 \dots X_q$ 是多个检测参数, $\beta_1, \beta_q$ 是每个检测参数的权重, $Y_i$ 用于确定相应的服务流程。
10. 根据权利要求9所述的方法,其特征在于,选定多个检测参数后,先将各项检验参数 $X_i$ 与 $Y_i$ 的 $R^2$ 都计算一遍,再将与 $Y_i$ 的相关系数大于预定值的参数引入多元线性回归模型。
11. 一种医学数据处理系统,其特征在于,所述系统包括:患者数据接收模块、优先级确定模块、处理消息发送模块,其中,
  - 患者数据接收模块用于接收患者数据;
  - 优先级确定模块用于确定接收到的患者数据的优先级;
  - 处理消息发送模块用于根据所述优先级向特定处理终端发送相关的处理消息,以匹配相应的服务流程。
12. 根据权利要求11所述的系统,其特征在于,所述系统包括有效性判断模块,所述有效性判断模块用于进行患者数据的有效性判断,以对有效的患者数据进行优先级判断。
13. 根据权利要求11所述的系统,其特征在于,所述系统包括数据时间模块,所述数据时间模块用于读取患者数据的发送时间和产生时间,并且根据所述发送时间和产生时间判

断是否进行优先级的判断。

14. 根据权利要求 11 所述的系统,其特征在於,所述系统包括可靠性判断模块,用于对患者数据进行可靠性判断,以对可靠性高的患者数据进行优先级判断。

## 医学数据处理方法和系统

### 技术领域

[0001] 本发明涉及医学数据处理技术,尤其涉及一种医学数据处理方法和系统。

### 背景技术

[0002] 慢性病患者会在长期的疾病治疗过程中需要多次使用专科的医疗资源,以便实现调整用药、指标监控等基础的疾病护理 / 监控,而且导致了专科医生接诊量的剧增,以及专科医生资源的浪费。

[0003] 同时,由于医生对于患者的处理是按照时间顺序进行的,但患者病情直接轻重程度不同,所以很难对紧急患者的医疗需求做出及时并且合理的响应,从而会耽误治疗。另外,专科医生每天的工作时间有限,巨大的接诊量使得他们不得不将缩短花在每个病人身上的时间,因此会导致紧急患者诊治的不准确、误诊等。

[0004] 可以看出,传统的医学数据处理系统的弊端在于:医疗资源没有得到良好的利用,医患匹配不均。

### 发明内容

[0005] 针对现有技术的缺陷,本发明的目的是提供一种能够实现医疗资源良好匹配的医学数据处理方法和系统。

[0006] 本发明提供了一种医学数据处理方法,其特征在于,所述方法包括:a. 接收患者数据;b. 将接收到的患者数据确定优先级;c. 根据所述优先级动态匹配合理的服务流程;d. 向特定处理终端发送相关的处理消息。优选地,医疗服务包括以下任何一种:短信用药提醒、生活习性短信提醒、一对多的现场教育、初级医生药物调整、初级医生与专科会诊、专科远程面诊、专科医生面诊。

[0007] 步骤 b 包括:将接收到的检测参数加入到多元线性回归模型,所述多元线性回归模型为:

[0008]  $Y_i = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_q X_q + \epsilon$ , 其中,  $X_1 \dots X_q$  是多个检测参数,  $\beta_1$ 、 $\beta_q$  是每个检测参数的权重,  $Y_i$  用于确定相应的服务流程。

[0009] 优选地,选定多个检测参数后,先将各项检验参数  $X_i$  与  $Y_i$  的  $R^2$  都计算一遍,再将与  $Y_i$  的相关系数大于预定值的参数引入多元线性回归模型。

[0010] 优选地,所述患者数据包括以下信息:疾病种类、检测参数、诊断时间、就诊地点、诊断医生、病历类型、数据输入方式、确诊时间、数据产生时间、数据发送时间、确诊机构。

[0011] 更优选地,在步骤 a 与步骤 b 之间还包括步骤 a11:读取患者数据中的数据产生时间和数据发送时间,当数据产生时间和数据发送时间之差小于预定值时,则该患者数据进入步骤 b 进行处理。

[0012] 优选地,在步骤 a 与步骤 b 之间还包括步骤 a12:对患者数据进行有效性分析,当数据为有效数据时,则该患者数据进入步骤 b 进行处理。

[0013] 优选地,在步骤 a 与步骤 b 之间还包括步骤 a13:对患者数据进行可靠性判断,可

靠性高的患者数据进入步骤 b 进行处理。

[0014] 优选地,步骤 c 为:根据所述优先级匹配合理的服务流程;d. 通过短信、电子邮件、电话中任何一种方式向特定处理终端发送相关的处理消息。

[0015] 相应地,本发明还提供了一种医学数据处理系统,所述系统包括:患者数据接收模块、优先级确定模块、处理消息发送模块。其中,患者数据接收模块用于接收患者数据;优先级确定模块用于确定接收到的患者数据的优先级;处理消息发送模块用于根据所述优先级动态匹配合理的服务流程;并且向特定处理终端发送相关的处理消息。

[0016] 优选地,所述系统包括有效性判断模块,所述有效性判断模块用于进行患者数据的有效性判断,以对有效的患者数据进行优先级判断。

[0017] 优选地,所述系统包括数据时间模块,所述数据时间模块用于读取患者数据的发送时间和产生时间,并且根据所述发送时间和产生时间判断是否进行优先级的判断。

[0018] 优选地,所述系统包括可靠性判断模块,用于对患者数据进行可靠性判断,以对可靠性高的患者数据进行优先级判断。

[0019] 相对于现有技术,本发明的医学数据处理方法和系统能够实现医疗资源的良好匹配,提高患者的就诊效率。这种技术平台可以通过连接不同的设备以及数据参数,沿用到不同的医疗数据管理中。采用本发明的医学数据处理方法和处理系统后,在相同诊疗时间及病患数量情况及相同复杂程度下,由于优先级的设定,在不同级别患者中的医疗资源动态分配将更合理化,也使得错诊率、误诊率大大降低。基于该智能平台的拓展性设计,该平台可以整合不同的物联网设备用于不同的专科疾病,例如糖尿病等疾病患者的数据管理。

## 附图说明

[0020] 图 1 是本发明一种优选实施方式的医学数据处理方法的流程图;

[0021] 图 2 是本发明一种优选实施方式的医学数据处理系统的结构框图。

## 具体实施方式

[0022] 图 1 是本发明一种优选实施方式的医学数据处理方法的流程图。如图所示,所述方法包括如下步骤:

[0023] 在步骤 110,接收患者数据。由于是针对患者群体,因此可以有多患者多类型的数据来源。所有患者信息可以收集到相应计算机系统上的患者数据库,数据的来源可以是多种途径,例如:通过病历自动导入或进行手工整理获得的患者数据,也可以是通过电子诊疗系统采集的患者数据。接收患者数据的途径可以是手工录入、无线通信传输、有线通信传输等方式采集。患者数据的格式可以采用各种计算机可以识别的格式,以数据库能够处理的数据类型存储在数据库中。

[0024] 优选地,所述患者数据包括时间标记,因此,多条患者信息之间可以通过时间排序,依次进行处理。所述时间标记包括该条患者数据的产生时间、发送时间等。

[0025] 优选地,所述患者数据可以包括疾病种类、检测参数(例如,血液检测参数、血压、心率等),诊断时间、就诊地点、诊断医生、病历类型、数据输入方式、确诊时间、确诊机构等信息。

[0026] 优选地,包括步骤 112,在该步骤,进行患者数据的有效性判断,对有效的患者数据

进行后续处理。可以有多种手段进行有效性判断,在一个实施例中,获取患者数据中的确诊机构,如果确诊机构为空,即未填写确诊机构,则该患者数据无效,不对该患者数据进行后续处理。

[0027] 优选地,还包括:读取患者数据的发送时间和产生时间,所述发送时间是指患者数据发送给客户的时间,产生时间是指所述用户数据的生成时间。发送时间和产生时间可以用于参与可靠性鉴定的工作。具体地,如果发送时间和产生时间之差大于或等于预定值,则说明该数据已经较旧,进而判定该患者数据不再进入下一步的处理步骤,即不再进行步骤 120 的确定优先级步骤,该患者数据可以重新生成后再进入后续程序。在步骤 115:读取患者数据的发送时间和产生时间,如果发送时间和产生时间之差小于预定值,则所述患者数据进行步骤 120 的确定优先级操作。

[0028] 更优选地,还可以包括步骤 116:对患者数据进行可靠性判断,对可靠性高的患者数据进行后续处理。所述可靠性判断可以通过多种手段实现,在一个实施例中,读取患者数据中的数据输入方式,如果数据输入方式为患者自己手动输入,则判定该患者数据的可靠性低;在另一个实施例中,读取患者数据中的病历类型,如果病历类型为电子病历,则判定该患者数据的可靠性高。当患者数据的可靠性高时,该患者数据进入后续步骤,例如步骤 120 中进行处理。

[0029] 本领域技术人员可以理解,可以在上述步骤 112、115、116 中的一个或多个步骤的处理后再进入步骤 120。

[0030] 在步骤 120,对接收到的患者数据确定数据的优先级。优先级的确定可以根据患者数据的属性确定。所述属性是指患者数据所包含的一个或多项信息,例如:以检测参数种的血压作为一项属性,当某个患者数据中的收缩压超过了 180 毫米汞柱或舒张压超过了 110 毫米汞柱时,则该患者数据的优先级为最高;又例如:以疾病类型作为一项属性,当疾病类型属于特定类型(例如糖尿病或其它慢性病症)时,优先级设置为最低。

[0031] 优先级的高级根据需要处理的紧急程度确定,例如当检测参数反映出患者需要进行紧急治疗时,就需要将优先级设定为高级。优先级可以划分为不同的数量,例如高、中、低三级,或 1、2、3、4……等从高到低(或从低到高)多级,具体情况可以根据实际应用来确定。

[0032] 在实际的医学数据处理问题中,一个判断结果  $Y_i$  往往受到多个生理指标的影响。例如,在糖尿病数据处理中,除了受其血糖含量变化的影响外,还受诸如饮食、运动习惯,遗传,其他疾病等多种因素的影响,因此表现在线性回归模型中的解释变量有多个。这样的模型被称为多元线性回归模型。

[0033] 在本发明中,优选地,根据多种检验参数,例如生理指标、历史用药情况、饮食运动等数据,即下文中的  $X_1, X_2, X_3, \dots, X_q$  的大数据库积累,通过建立计算结果( $Y_i$ )和检测参数的关联,即多元线性回归模型:

$$[0034] \quad Y_i = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_q X_q + \varepsilon$$

[0035] 其中,  $P$  值确保在小于 0.05(即生物学上认可的统计学意义),随着数据库数据量的不断积累和生理指标的不断加入,模型会不断优化筛选出具有统计学意义的  $X$  以及  $X$  之间的关联效应。其中, $P$  值定义为,原假设  $H_0$  在本次试验中出现的概率。 $p$  值的统计学意义是结果真实程度(能够代表总体)的一种估计方法,为结果可信程度的一个递减指标,是将观察结果认为有效即具有总体代表性的犯错概率。如  $p=0.05$  提示样本中变量关联有

5%的可能是由于偶然性造成的。在许多研究领域,0.05的p值通常被认为是可接受错误的边界水平。 $\beta_1$ 、 $\beta_q$ 为每个生理指标 $X_i$  ( $i=1, 2, \dots, q$ )的权重,所述权重也会根据数据的不断增加而优化调整。

[0036] P值的计算方法如下:

[0037] 一般地,用X表示检验的统计量,当 $H_0$ 为真时,可由样本数据计算出该统计量的值C,根据检验统计量X的具体分布,可求出P值。具体地说:

[0038] 左侧检验的P值为检验统计量X小于样本统计值C的概率,即: $P=P\{X<C\}$

[0039] 右侧检验的P值为检验统计量X大于样本统计值C的概率: $P=P\{X>C\}$

[0040] 双侧检验的P值为检验统计量X落在样本统计值C为端点的尾部区域内的概率的2倍: $P=2P\{X>C\}$ (当C位于分布曲线的右端时)或 $P=2P\{X<C\}$ (当C位于分布曲线的左端时)。若X服从正态分布和t分布,其分布曲线是关于纵轴对称的,故其P值可表示为 $P=P\{|X|>C\}$ 。

[0041] 计算出P值后,将给定的显著性水平 $\alpha$ 与P值比较,就可作出检验的结论:

[0042] 如果 $\alpha > P$ 值,则在显著性水平 $\alpha$ 下拒绝原假设。

[0043] 如果 $\alpha \leq P$ 值,则在显著性水平 $\alpha$ 下接受原假设。

[0044] 在实践中,当 $\alpha = P$ 值时,也即统计量的值C刚好等于临界值,优选地,可增加样本容量,重新进行抽样检验。

[0045] 在一个具体实施方式中:引入相关的多项检验参数 $X_1, \dots$ ,在每步操作中进行统计学评估直到所有达到某些特定的统计指标标准( $p < 0.05$ )的参数全部被引入。即:先将各项检验参数 $X_i$ 与 $Y_i$ 的 $R^2$ 都计算一遍,再将与 $Y_i$ 的相关系数大于预定值的参数引入。

[0046]  $R^2$ 为相关系数,是在线性条件下说明两个变量之间相关关系密切程度的统计分析指标,其值分布在0-1之间,越高越表明相关性越高。

[0047] 其中,本发明中 $R^2$ 的计算公式为:

$$[0048] \quad r = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{\sqrt{\Sigma(x - \bar{x})^2} \cdot \sqrt{\Sigma(y - \bar{y})^2}} = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{\sqrt{\Sigma(x - \bar{x})^2} \cdot \sqrt{\Sigma(y - \bar{y})^2}}$$

[0049] 例如,首先引入与判断 $Y_i$ 相关性最大的参数 $X_1$ ,然后记录观察到的 $Y_i$ 最大变化比例。之后引入参数 $X_2$ ,需满足 $X_2$ 对于 $Y_i$ 相关系数 $R^2$ 提高大于预定值(例如 $R^2$ 需大于0.653),同时,保证 $X_1$ 仍保留在模型中,且引入 $X_2$ 后在统计上是可以接受的( $P < 0.05$ )。重复进行这一过程,从而不断增加新的参数,直到剩余参数对 $Y_i$ 的评估没有明显的贡献,例如,确定的限度即为( $p < 0.05$ ,  $R^2$ 增加不明显)。

[0050] 采用本发明的上述处理方法,可以克服现有处理方法的缺陷,例如,在现有处理方法中,当血压超过某个指标,则判断该患者为高血压患者。采用这种方法,由于是固定的阈值,如果阈值设置太严,则会造成大量的假阳性(即误判为疾病的情况);如果标准太宽松,则会造成假阴性(即有疾病的被判为正常的情况)。而本发明所述的动态智能数据处理方法,可以实现动态临床路径分配,打破了传统“一刀切”的固定标准模式,使得结果更加准确。

[0051] 在步骤130,根据所述优先级向相关的处理终端发送处理消息。所述消息可以通过多种形式实现,例如、短信、电子邮件、电话等方式。当优先级高时,相应的消息为应急处理消息,例如,在特定的终端设备(如预警大屏幕,PC,平板电脑,手机)上提示医生做出相应的

紧急处理的临床路径；当优先级较低时，相应的消息为日常处理临床路径所需的消息，例如向全科医生诊室终端发送提示信息。发出的处理消息表示所需的相应处理，处理消息可以是文字、图案、音频、视频等多种形式。

[0052] 相应地，本发明还提供了一种医学数据处理系统 200。该系统包括患者数据接收模块 210、优先级确定模块 220、处理消息发送模块 230。其中：

[0053] 患者数据接收模块 210 用于接收患者数据。所有患者信息可以收集到相应计算机系统患者数据库，接收患者数据的途径可以是手工录入、无线通信传输、有线通信传输等方式采集。患者数据的格式可以采用各种计算机可以识别的格式，以数据库能够处理的数据类型存储在数据库中。在本发明的优选实施例中，患者数据接收模块 210 由各种适合的有线、无线通信接口及数据采集程序实现。

[0054] 优选地，本系统可以包括进行患者数据的有效性判断的有效性判断模块 212。可以有多种手段进行有效性判断，在一个实施例中，获取患者数据中的确诊机构，如果确诊机构为空（即未填写确诊机构），则该患者数据无效，不对该患者数据进行后续处理。

[0055] 优选地，本系统还包括数据时间模块 215，所述数据时间模块用于读取患者数据的发送时间和产生时间，并且根据所述发送时间和产生时间判断是否进行后续优先级的匹配。具体地，如果发送时间和产生时间之差大于或等于预定值，则说明该数据已经较旧，进而判定该患者数据不再进入下一步的处理步骤，即不再确定优先级，该患者数据可以重新生成后再进入后续程序。如果发送时间和产生时间之差小于预定值，则所述患者数据继续确定优先级操作。

[0056] 优选地，本系统还可以包括可靠性判断模块 216，用于对患者数据进行可靠性判断，对可靠性高的患者数据进行后续处理。所述可靠性判断可以通过多种手段实现，在一个实施例中，读取患者数据中的数据输入方式，如果数据输入方式为手动输入，则判定该患者数据的可靠性低；在另一个实施例中，读取患者数据中的病历类型，如果病历类型为电子病历，则判定该患者数据的可靠性高。当患者数据的可靠性高时，该患者数据进入后续步骤。

[0057] 优先级确定模块 220 用于确定接收到的患者数据的优先级。优先级的确定可以根据患者数据的属性确定。所述属性是指患者数据所包含的一个或多项信息，例如：以检测参数种的血压作为一项属性，当某个患者数据中的收缩压超过了 180 毫米汞柱或舒张压超过了 110 毫米汞柱时，则该患者数据的优先级会根据超标的程度判定为不同级别的优先级；又例如：以疾病类型作为一项属性，当疾病类型属于特定类型（例如糖尿病或其它慢性病）时，优先级设置为最低。

[0058] 在本发的优选实施例中，优先级确定模块 220 从患者数据接收模块 210 接收患者数据，通过解析患者数据中的特定信息并与设定值进行比较，并且根据数据来源做了一定的处理后，再确定该患者数据的优先级。

[0059] 处理消息发送模块 230 用于根据所述优先级发送相关的处理消息。所述处理消息可以通过多种形式实现，例如、短信、电子邮件、电话等方式。发出的处理消息表示所需的相应处理，处理消息可以是文字、图案、音频、视频等多种形式。因此，在本发明的具体实施方式中，处理消息发送模块 230 从优先级确定模块 220 接收根据患者数据得出的优先级信息，并且根据所述优先级信息，向特定的处理终端发送处理消息。处理消息发送模块 230 可以通过与消息发送设备（如电话、移动通信终端、电子邮件模块）通信连接的软件单元实现。



[0060] 上述模块均可以由软件模块以及相应硬件控制逻辑电路实现,例如中央处理器、存储器、输入输出接口电路及嵌入的程序代码等实现。

[0061] 采用本发明的医学数据处理方法和处理系统后,在相同诊疗时间及病患数量情况及相同复杂程度下,由于优先级的设定,在不同级别患者中的医疗资源分配将更合理化,也使得错诊率、误诊率大大降低。

[0062] 本发明已经结合其具体实施方式进行了详细描述,并且上述实施方式仅供说明本发明之用,而并非是对本发明的限制,在不脱离本发明公开范围的情况下,本领域技术人员还可以做出各种变化和变型,因此所有等同的技术方案,包括所有不背离本发明精神和范围的改变和修改都在本发明保护的范围之内。

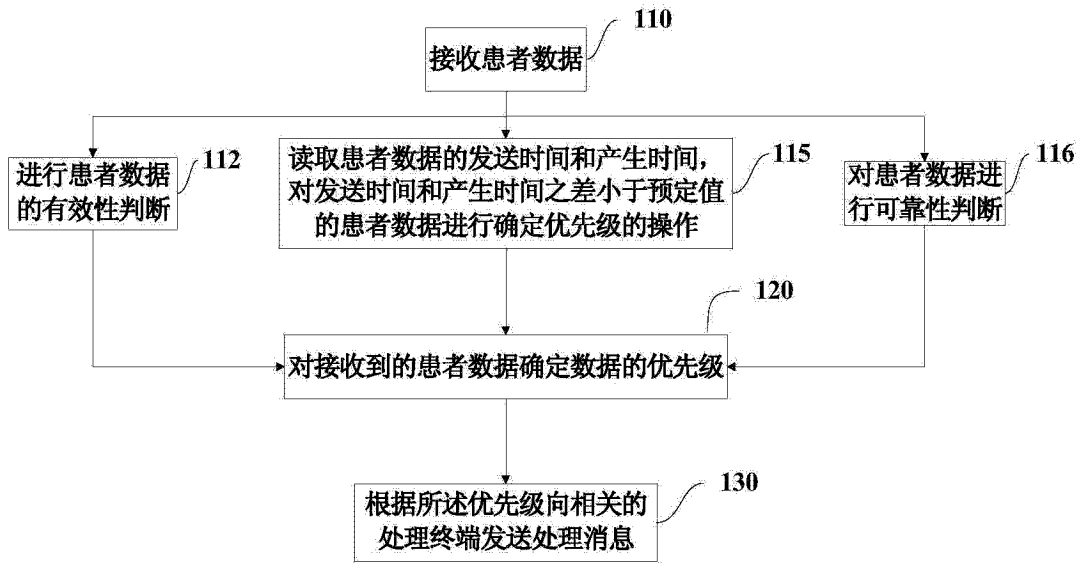


图 1

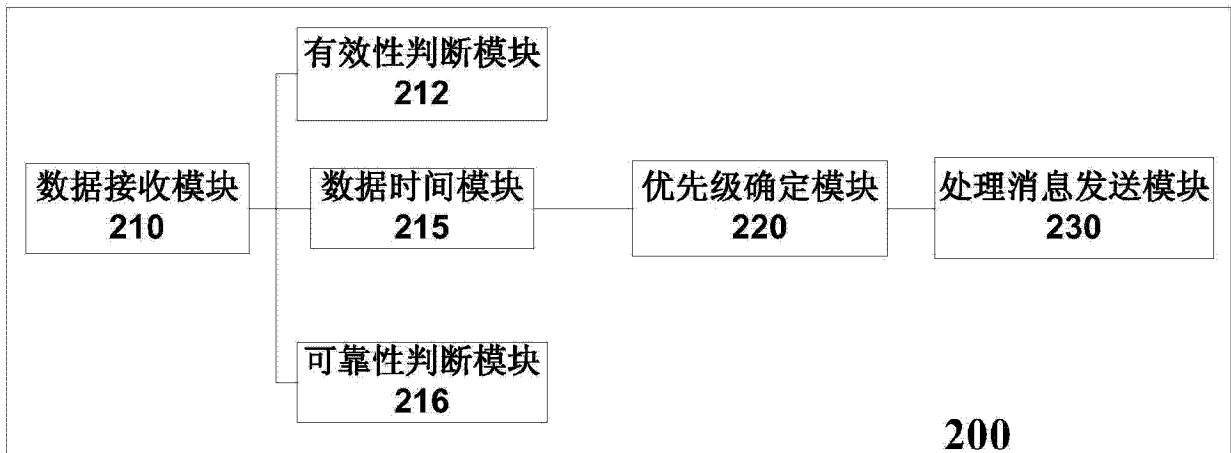


图 2