



(12) 发明专利申请

(10) 申请公布号 CN 104469932 A

(43) 申请公布日 2015. 03. 25

(21) 申请号 201410670795. 7

(22) 申请日 2014. 11. 21

(71) 申请人 北京拓明科技有限公司

地址 100083 北京市海淀区志新路二里庄  
35 号万和大厦 4 层

(72) 发明人 张光辉 王广善 常青

(74) 专利代理机构 北京天悦专利代理事务所  
(普通合伙) 11311

代理人 田明 张海秀

(51) Int. Cl.

H04W 64/00(2009. 01)

H04W 24/10(2009. 01)

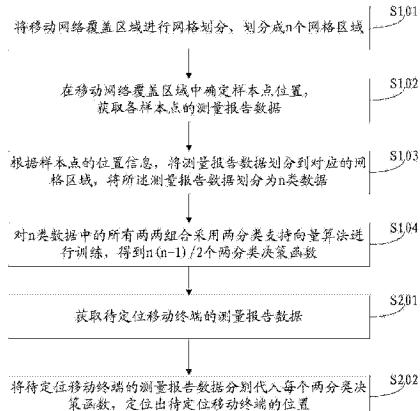
权利要求书2页 说明书7页 附图3页

(54) 发明名称

一种基于支持向量机的位置指纹定位方法

(57) 摘要

本发明公开了一种基于支持向量机的位置指纹定位方法，属于移动通信邻域的无线业务支撑领域。所述方法主要包括离线训练和在线定位两个阶段，离线训练阶段通过设置样本点，对样本点采集路测数据，通过对路测数据采用基于支持向量机算法进行训练得到定位模型。在线定位阶段获取待定位移动终端的数据后，通过定位模型实现快速定位。采用本发明所述的方法，能够有效降低在线定位阶段的定位复杂度，减小定位时延。



1. 一种基于支持向量机的位置指纹定位方法,包括以下步骤:

(1) 离线训练阶段;所述离线训练阶段包括:

(1-1) 将移动网络覆盖区域进行网格划分,划分成 n 个网格区域,每个网格区域对应相应的位置编号;其中,n ≥ 2;

(1-2) 在移动网络覆盖区域中确定样本点位置,获取各样本点的测量报告数据;所述测量报告数据中包括样本点接收到的来自各扇区的接收电平;

(1-3) 根据样本点的位置信息,将所述测量报告数据划分到对应的网格区域,将所述测量报告数据划分为 n 类数据;

(1-4) 将每个网格区域的位置编号以及与其对应的一类数据作为训练样本数据,遍历 n 类训练样本数据中所有的两两组合,并对所有两两组合的两类训练样本数据采用基于一阶多项式核函数的两分类支持向量算法进行训练,得到  $n(n-1)/2$  个两分类决策函数,将得到的  $n(n-1)/2$  个两分类决策函数存储,作为定位模型;

(2) 在线定位阶段;所述在线定位阶段包括:

获取待定位移动终端的测量报告数据;

遍历所述  $n(n-1)/2$  个两分类决策函数,将待定位移动终端的测量报告数据分别代入每个两分类决策函数,定位出待定位移动终端所在的网格区域,将定位出的网格区域的中心经纬度确定为待定位移动终端的位置。

2. 根据权利要求 1 所述的一种基于支持向量机的位置指纹定位方法,其特征在于,步骤(1-4)中,对两类训练样本数据采用基于一阶多项式核函数的两分类支持向量算法进行训练,得到两分类决策函数并存储,包括:

对于所述两类训练样本数据,计算每类训练样本数据所对应的网格区域接收到的来自每个扇区的平均接收电平,得到两类新的训练样本数据;来自某个扇区的平均接收电平是指网格区域所包含的所有样本点接收到的来自该扇区的接收电平的平均值;

采用一阶多项式核函数将所述两类新的训练样本数据映射到高维空间,采用两分类支持向量算法对映射到高维空间的两类新的训练样本数据进行训练,得到一个两分类决策函数,两分类决策函数的表达式为:

$$\text{两分类决策函数 } f(x) = \operatorname{sgn} \left[ \sum_{i=1}^l \alpha_i y_i (\gamma x_i x + r) + b \right]$$

其中,  $\operatorname{sgn}[\cdot]$  是符号函数, l 表示新的训练样本数据中涉及到的扇区的个数,  $x_i$  表示新的训练样本数据所对应的网格区域接收到的来自第 i 个扇区的平均接收电平, 为待定位移动终端的测量报告数据中所接收到的来自各个扇区的平均接收电平,  $x_i x$  表示  $x_i$  与 x 的向量内积,  $y_i$  为新的训练样本数据中所涉及到的扇区的分类标识,  $f(x)$  为 x 的分类判断标识, 根据  $f(x)$  判断待定位移动终端的测量报告数据属于两类新的训练样本数据中的哪一类,  $\alpha_i$  和 b 为修正系数, 为经验值;

将所述两分类函数进行转换并存储,包括:将两两分类决策函数转换为:

$$f(x) = \operatorname{sgn} [\beta x + b']$$

其中,

$$\sum_{i=1}^l \alpha_i y_i \gamma x_i = \beta$$

$$\sum_{i=1}^l \alpha_i y_i r + b = b'$$

将转换后的两分类决策函数及每个决策函数对应的  $\beta$  和  $b'$  以文件的形式进行存储，作为定位模型。

3. 根据权利要求 2 所述的一种基于支持向量机的位置指纹定位方法，其特征在于，步骤 (2) 中，所述定位出待定位移动终端所在的网格区域，包括：

(2-1) 将待定位移动终端的测量报告数据按照其通话 ID 和时间戳分为若干段，得到每段中所包含的测量报告数据中所接收到的来自各个扇区的平均接收电平，包括：

假设属于某一通话 ID 的测量报告数据有  $m$  条，将  $m$  条测量报告数据按照时间戳由小到大排列，并标记为  $1, 2, \dots, m$ ；

设定分段时间间隔  $t_1$ ，假设  $t_1$  的时间间隔内对应着  $m_1$  条测量报告数据，则将排列好的测量报告数据按照  $m_1$  进行分割， $m$  条测量报告数据共分割为  $m/m_1$  段；

统计每段中所有测量报告数据所涉及到的所有不同扇区，计算每段中所有测量报告数据中所接收到的来自各个扇区的平均接收电平；

(2-2) 将每段中所包含的测量报告数据中所接收到的来自各个扇区的平均接收电平作为该段新的测量报告数据，遍历所述  $n(n-1)/2$  个两分类决策函数，将所述新的测量报告数据作为两分类决策函数中的  $x$ ，代入每一个两分类决策函数进行计算，根据计算出的  $f(x)$  的结果判断出新的测量报告数据在每一个两分类决策函数所对应的两类训练样本数据中的所属类，并根据判断结果确定出新的测量报告数据的最终所属类；

(2-3) 将每段的新的测量报告数据的最终所属类所对应的网格区域的中心经纬度作为该段所对应的时间间隔内待定位移动终端的定位结果。

4. 根据权利要求 3 所述的一种基于支持向量机的位置指纹定位方法，其特征在于，步骤 (2-2) 中，在计算机上开启若干个线程，将  $m/m_1$  段中的新的测量报告数据平均在若干个线程上进行计算。

## 一种基于支持向量机的位置指纹定位方法

### 技术领域

[0001] 本发明涉及移动通信邻域的无线业务支撑领域,具体涉及一种基于支持向量机的位置指纹定位方法。

### 背景技术

[0002] 随着无线通信的发展,无线定位技术被越来越多的人所关注。现有比较热门的定位技术有以下几种,到达时间 (ToA),到达时间差 (TDOA),到达角 (AOA),全球定位 (GPS),位置指纹定位 (LFP)。这几种定位方法各有其优缺点,然而,随着移动互联网的发展,3G 手机的普及和 4G 手机的推广,高精度的定位方法无疑是未来移动定位技术的关键需求。以上几种的定位方法中,指纹定位方法定位精度最高,在样本足够密集的时候定位误差可以低于 100 米,然而指纹定位方法一个缺点就是计算复杂度很高,对每一个定位点,都需要大量的样本比对和计算,运算速度较低,且对硬件环境要求较高,成为实际应用推广的瓶颈。

[0003] 目前常用的指纹定位方法主要有最近邻算法 (Nearest Neighborhood, NN),神经网络算法、支持向量机算法 (Support Vector Machine, SVM) 等,虽然算法本身各有区别,但是基本上都分为两个阶段:首先,离线训练阶段,将海量的样本数据训练成定位模型;第二,在线定位阶段,对于实时的待定位的测量报告,通过和定位模型中的海量样本数据逐一比对,得到最终的定位结果。

[0004] 现有指纹定位方法存在以下问题:

[0005] 1. 需要对所有的历史数据集进行全局搜索,计算开销大,速度慢,难以满足海量用户数据实时定位处理的需要。

[0006] 2. 支持向量机中的核函数多数使用高斯(径向基)核函数或是多阶的多项式核函数,导致分类决策函数的计算复杂度过高。

[0007] 3. 目前支持向量机方法中的分类判决采取的策略是遍历所有不同的两两分类组合,导致分类判决的计算复杂度极高。

[0008] 4. 没有运用多线程技术进行加速优化。

[0009] 总的来说,现有的基于支持向量机的定位方法算法复杂度较高,计算速度不能满足实际需要。

### 发明内容

[0010] 针对现有技术中存在的缺陷,本发明的目的在于提供一种基于支持向量机的位置指纹定位方法,实现对待定位移动终端的快速定位。

[0011] 为实现上述目的,本发明采用的技术方案为:一种基于支持向量机的位置指纹定位方法,包括以下步骤:

[0012] (1) 离线训练阶段;所述离线训练阶段包括:

[0013] (1-1) 将移动网络覆盖区域进行网格划分,划分成 n 个网格区域,每个网格区域对应相应的位置编号;其中, n ≥ 2;

[0014] (1-2) 在移动网络覆盖区域中确定样本点位置, 获取各样本点的测量报告数据; 所述测量报告数据中包括样本点接收到的来自各扇区的接收电平 RSS;

[0015] (1-3) 根据样本点的位置信息, 将所述测量报告数据划分到对应的网格区域, 将所述测量报告数据划分为 n 类数据;

[0016] (1-4) 将每个网格区域的位置编号以及与其对应的一类数据作为训练样本数据, 遍历 n 类训练样本数据中所有的两两组合, 并对所有两两组合的两类训练样本数据采用基于一阶多项式核函数的两分类支持向量算法进行训练, 得到  $n(n-1)/2$  个两分类决策函数, 将得到的  $n(n-1)/2$  个两分类决策函数存储, 作为定位模型;

[0017] (2) 在线定位阶段; 所述在线定位阶段包括:

[0018] 获取待定位移动终端的测量报告数据;

[0019] 遍历所述  $n(n-1)/2$  个两分类决策函数, 将待定位移动终端的测量报告数据分别代入每个两分类决策函数, 定位出待定位移动终端所在的网格区域, 将定位出的网格区域的中心经纬度确定为待定位移动终端的位置。

[0020] 可选的, 如上所述的一种基于支持向量机的位置指纹定位方法, 步骤 (1-4) 中, 对两类训练样本数据采用基于一阶多项式核函数的两分类支持向量算法进行训练, 得到两分类决策函数并存储, 包括:

[0021] 对于所述两类训练样本数据, 计算每类训练样本数据所对应的网格区域接收到的来自每个扇区的平均接收电平, 得到两类新的训练样本数据; 来自某个扇区的平均接收电平是指网格区域所包含的所有样本点接收到的来自该扇区的接收电平的平均值;

[0022] 采用一阶多项式核函数将所述两类新的训练样本数据映射到高维空间, 采用两分类支持向量算法对映射到高维空间的两类新的训练样本数据进行训练, 得到一个两分类决策函数, 两分类决策函数的表达式为:

$$[0023] \text{两分类决策函数 } f(x) = \operatorname{sgn} \left[ \sum_{i=1}^l \alpha_i y_i (\gamma x_i x + r) + b \right]$$

[0024] 其中,  $\operatorname{sgn}[\cdot]$  是符号函数,  $l$  表示新的训练样本数据中涉及到的扇区的个数,  $x_i$  表示新的训练样本数据所对应的网格区域接收到的来自第  $i$  个扇区的平均接收电平,  $x$  为待定位移动终端的测量报告数据中所接收到的来自各个扇区的平均接收电平,  $x_i x$  表示  $x_i$  与  $x$  的向量内积,  $y_i$  为新的训练样本数据中所涉及到的扇区的分类标识,  $f(x)$  为  $x$  的分类判断标识, 根据  $f(x)$  判断待定位移动终端的测量报告数据属于两类新的训练样本数据中的哪一类,  $\alpha_i$  和  $b$  为修正系数, 为经验值;

[0025] 将所述两分类函数进行转换并存储, 包括: 将两两分类决策函数转换为:

$$[0026] f(x) = \operatorname{sgn} [\beta x + b']$$

[0027] 其中,

$$[0028] \sum_{i=1}^l \alpha_i y_i \gamma x_i = \beta$$

$$[0029] \quad \sum_{i=1}^l \alpha_i y_i r + b = b'$$

[0030] 将转换后的两分类决策函数及每个决策函数对应的  $\beta$  和  $b'$  以文件的形式进行存储,作为定位模型。

[0031] 可选的,如上所述的一种基于支持向量机的位置指纹定位方法,步骤(2)中,所述定位出待定位移动终端所在的网格区域,包括:

[0032] (2-1) 将待定位移动终端的测量报告数据按照其通话 ID 和时间戳分为若干段,得到每段中所包含的测量报告数据中所接收到的来自各个扇区的平均接收电平,包括:

[0033] 假设属于某一通话 ID 的测量报告数据有  $m$  条,将  $m$  条测量报告数据按照时间戳由小到大排列,并标记为  $1, 2, \dots, m$ ;

[0034] 设定分段时间间隔  $t_1$ ,假设  $t_1$  的时间间隔内对应着  $m_1$  条测量报告数据,则将排列好的测量报告数据按照  $m_1$  进行分割,  $m$  条测量报告数据共分割为  $m/m_1$  段;

[0035] 统计每段中所有测量报告数据所涉及到的所有不同扇区,计算每段中所有测量报告数据中所接收到的来自各个扇区的平均接收电平;

[0036] (2-2) 将每段中所包含的测量报告数据中所接收到的来自各个扇区的平均接收电平作为该段新的测量报告数据,遍历所述  $n(n-1)/2$  个两分类决策函数,将所述新的测量报告数据作为两分类决策函数中的  $x$ ,代入每一个两分类决策函数,根据计算出的  $f(x)$  的结果判断出新的测量报告数据在每一个两分类决策函数所对应的两类训练样本数据中的所属类,并根据判断结果确定出新的测量报告数据的最终所属类;

[0037] (2-3) 将每段的新的测量报告数据的最终所属类所对应的网格区域的中心经纬度作为该段所对应的时间间隔内待定位移动终端的定位结果。

[0038] 可选的,如上所述的一种基于支持向量机的位置指纹定位方法,步骤(2-2)中,在计算机上开启若干个线程,将  $m/m_1$  段中的新的测量报告数据平均在若干个线程上进行计算。

[0039] 本发明的有益效果在于:本发明所述的方法,采用基于一阶多项式核函数的支持向量机算法计算两分类判决函数,并通过两分类判决函数的转换存储,能够有效降低在线定位阶段的分类判决的复杂度。此外,通过对测量报告数据的分段合并处理的方式,并以合并报告的定位结果作为所在分段中所有测量报告的定位结果,大大减小了定位时延。

## 附图说明

[0040] 图 1 为本发明具体实施方式中一种基于支持向量机的位置指纹定位方法的总体流程图;

[0041] 图 2 为本发明具体实施方式中一种基于支持向量机的位置指纹定位方法的具体流程图;

[0042] 图 3 为本发明具体实施方式中离线训练阶段的示意图;

[0043] 图 4 为本发明具体实施方式中在线定位阶段的示意图。

## 具体实施方式

[0044] 下面结合说明书附图与具体实施方式对本发明做进一步的详细说明。

[0045] 图 1 和图 2 分别示出了本发明具体实施方式中一种基于支持向量机的位置指纹定位方法的总体流程图和具体流程图,由图中可以看出,该方法主要包括离线训练和在线定位两个阶段。每个阶段的具体流程如下:

[0046] 第一阶段:离线训练阶段

[0047] 该阶段的目的主要是采集路测数据,并对采集到的路测数据进行训练,得到定位模型。该阶段主要包括以下四个步骤,如图 2 所示和图 3 所示。

[0048] 步骤 S101:将移动网络覆盖区域进行网格划分,划分成 n 个网格区域,每个网格区域对应相应的位置编号;其中,  $n \geq 2$ ;

[0049] 本实施例中,可以按照一定的边长对移动网络覆盖区域进行网格划分,并对划分后的网格区域进行位置编号,可以是标注为区域 1, 2, 3, ……, n。

[0050] 步骤 S102:在移动网络覆盖区域中确定样本点位置,获取各样本点的测量报告数据;

[0051] 在划分后的所有网格区域中确定样本点位置,并采集样本点的路测信息。本实施例中,可以用带有 GPS 设备的移动终端在网格区域中进行路测,获取到待有位置信息的测量报告数据。所述测量报告数据中包括但不限于样本点接收到的来自各扇区的接收电平 RSS、样本点的位置信息(经纬度)、所在的服务小区、邻小区等。

[0052] 其中,样本点的选择可以根据需要进行确定,选择的样本点越多,得到的定位模型越精确,当然定位模型的计算复杂度也越高。

[0053] 步骤 S103:根据样本点的位置信息,将所述测量报告数据划分到对应的网格区域,将所述测量报告数据划分为 n 类数据;

[0054] 根据各样本点的位置信息以及步骤 S101 所划分的网格区域,把步骤 S102 中所获取的测量报告数据分别划分到对应的网格区域中,由于网格区域共划分成了 n 个,对应的所述测量报告数据也划分成了 n 类。

[0055] 步骤 S104:对 n 类数据中的所有两两组合采用两分类支持向量算法进行训练,得到  $n(n-1)/2$  个两分类决策函数。

[0056] 将每个网格区域的位置编号即位置信息以及与其对应的一类数据作为训练样本数据,得到 n 类训练样本数据,遍历 n 类训练样本数据中所有的两两组合,并对所有两两组合的两类训练样本数据采用基于一阶多项式核函数的两分类支持向量算法进行训练,得到  $n(n-1)/2$  个两分类决策函数,将得到的  $n(n-1)/2$  个两分类决策函数存储,作为定位模型。

[0057] 本实施例中,对两类训练样本数据采用基于一阶多项式核函数的两分类支持向量算法进行训练,得到两分类决策函数并存储,具体包括以下步骤:

[0058] (1) 对于两类训练样本数据,计算每类训练样本数据所对应的网格区域接收到的来自每个扇区的平均接收电平,得到两类新的训练样本数据;来自某个扇区的平均接收电平是指网格区域所包含的所有样本点接收到的来自该扇区的接收电平的平均值。

[0059] 假设一类训练样本数据中所包含的测量报告的个数为 10 个,每个测量报告接收到的来自 A 扇区的接收电平 RSS 分别为  $A_1, A_2, \dots, A_{10}$ , 则该类训练样本数据所对应的网格区域接收到的来自 A 扇区的平均接收电平  $\overline{Rxlev_A}$  为:

[0060]  $\overline{Rxlev_A} = \sum_{i=1}^{10} A_i / 10$  。

[0061] 需要说明是，在实际应用中，每类训练样本数据所对应的扇区的个数应该是相同的，如果一类训练样本数据中的某个测量报告在某一扇区上没有接收电平信息，则可以定义该测量报告接收到的来自该扇区的接收电平为一个固定的低电平，低电平的具体取值为经验值。

[0062] (2) 采用一阶多项式核函数将所述两类新的训练样本数据映射到高维空间；

[0063] 本实施例中，所述的一阶多项式核函数可以是  $k(x_i, x_j) = \gamma x_i x_j + r$ ，其中， $x_i$  和  $x_j$  为带有位置信息（网格区域位置编号）的电平向量， $x_i$ 、 $x_j$  分别表示一类新的训练样本数据所对应的网格区域接收到的来自第  $i$  个和第  $j$  个扇区的平均接收电平， $x_i x_j$  表示  $x_i$  与  $x_j$  的向量内积， $\gamma$  与  $r$  为核函数参数，为经验值；

[0064] (3) 采用两分类支持向量算法，对映射到高维空间的两类新的训练样本数据进行训练（构造二次优化不等式并进行求解等过程），得到一个两分类决策函数，两分类决策函数的表达式为：

[0065] 两分类决策函数  $f(x) = \text{sgn} \left[ \sum_{i=1}^l \alpha_i y_i (\gamma x_i x + r) + b \right]$

[0066] 其中， $\text{sgn}[\cdot]$  是符号函数， $l$  表示新的训练样本数据中涉及到的扇区的个数， $x_i$  表示新的训练样本数据所对应的网格区域接收到的来自第  $i$  个扇区的平均接收电平， $x$  为待定位移动终端的测量报告数据中所接收到的来自各个扇区的平均接收电平， $x_i x$  表示  $x_i$  与  $x$  的向量内积， $y_i$  为新的训练样本数据中所涉及到的扇区的分类标识， $f(x)$  为  $x$  的分类判断标识，根据  $f(x)$  判断待定位移动终端的测量报告数据属于两类新的训练样本数据中的哪一类， $\alpha_i$  和  $b$  为修正系数，为经验值；

[0067] 本实施例中， $y_i$  在 SVM 两分类决策函数中是样本数据的分类标识，本实施例中，为新的训练样本数据中所涉及到的扇区的分类标识， $y_i$  一般取值为“+1”或者“-1”，一个扇区所对应的  $y_i$  的取值是在对映射到高维空间的两类新的训练样本数据进行训练的过程确定的，具体的确认方式为现有技术。

[0068]  $\text{sgn}[\cdot]$  是符号函数， $\sum_{i=1}^l \alpha_i y_i (\gamma x_i x + r) + b$  的计算结果决定了符号函数的结果，本实施例中，由于  $f(x)$  为两分类决策函数，因此  $f(x)$  的值作为分类判断标识只可能是两种结果，即符号函数的结果只能是两种， $\sum_{i=1}^l \alpha_i y_i (\gamma x_i x + r) + b$  结果大于或等于零时， $f(x)$  返回“+1”， $\sum_{i=1}^l \alpha_i y_i (\gamma x_i x + r) + b$  结果小于零时， $f(x)$  返回“-1”，根据  $f(x)$  的结果是正还是负来判断出待定位移动终端的测量报告数据属于两类新的训练样本数据中的哪一类。

[0069] (4) 将所述两分类函数进行转换并存储，包括：将两两分类决策函数转换为：

[0070]  $f(x) = \text{sgn} [\beta x + b']$

[0071] 其中，

$$[0072] \sum_{i=1}^l \alpha_i y_i \gamma x_i = \beta$$

$$[0073] \sum_{i=1}^l \alpha_i y_i r + b = b'$$

[0074] 将转换后的两分类决策函数及每个决策函数对应的  $\beta$  和  $b'$  以文件的形式进行存储，作为定位模型。

[0075] 通过本发明实施例中所提供的上述离线训练阶段，使的两分类决策（或判断函数）的计算复杂度由现有技术中的  $O(1)$  降为  $O(1)$ ，最终，离线训练阶段可得到  $n(n-1)/2$  个决策函数以及对应的  $\beta$  和  $b'$ ，将这些数据以文件的形式保存在存储设备中得到可用于定位的模型。

[0076] 第二阶段，在线定位阶段。

[0077] 在线定位阶段的目的是采用第一阶段得到的定位模型进行待定位移动终端的快速定位。该阶段的具体流程如图 2 和图 4 所示，主要包括以下两个步骤：

[0078] 步骤 S201：获取待定位移动终端的测量报告数据；

[0079] 本实施例中，获取到的待定位移动终端的测量报告数据中包括但不限于待定位移动终端接收到的来自各扇区的接收电平 RSS 等。

[0080] 步骤 S202：遍历所述  $n(n-1)/2$  个两分类决策函数，将待定位移动终端的测量报告数据分别代入每个两分类决策函数，定位出待定位移动终端所在的网格区域，将定位出的网格区域的中心经纬度确定为待定位移动终端的位置。该步骤包括以下几个步骤：

[0081] (1) 将待定位移动终端的测量报告数据按照其通话 ID 和时间戳分为若干段，得到每段中所包含的测量报告数据中所接收到的来自各个扇区的平均接收电平，包括：

[0082] 假设属于某一通话 ID 的测量报告数据有  $m$  条，将  $m$  条测量报告数据按照时间戳由小到大排列，并标记为  $1, 2, \dots, m$ ；

[0083] 设定分段时间间隔  $t_1$ ，假设  $t_1$  的时间间隔内对应着  $m_1$  条测量报告数据，则将排列好的测量报告数据按照  $m_1$  进行分割， $m$  条测量报告数据共分割为  $m/m_1$  段；

[0084] 统计每段中所有测量报告数据所涉及到的所有不同扇区，计算每段中所有测量报告数据中所接收到的来自各个扇区的平均接收电平；

[0085] (2) 将每段中所包含的测量报告数据中所接收到的来自各个扇区的平均接收电平作为该段新的测量报告数据，遍历所述  $n(n-1)/2$  个两分类决策函数，将所述新的测量报告数据作为两分类决策函数中的  $x$ ，代入每一个两分类决策函数进行计算，根据计算出的  $f(x)$  的结果判断出新的测量报告数据在每一个两分类决策函数所对应的两类训练样本数据中的所属类，并根据判断结果确定出新的测量报告数据的最终所属类；

[0086] 由于在离线训练阶段得到的  $n(n-1)/2$  个两分类决策函数中包括了根据所有两两组合的训练数据得到的两分类函数，因此，该步骤中，在遍历  $n(n-1)/2$  个两分类决策函数后，会判断出新的测量报告数据所对应的  $n(n-1)/4$  个所属类，最后根据将这  $n(n-1)/4$  个所属类便可以确定出最终所属类。

[0087] 例如，假设  $n = 3$ ，共三类训练数据，记为 A、B、C，则两分类决策函数的个数为 3 个，

分别对应的两两组合为  $\{A, B\}$ ,  $\{A, C\}$ ,  $\{B, C\}$ , 遍历三个决策函数进行初步判断, 如果三个决策函数初步判断所对应的所属类分别为 A、C、C, 那么由于在  $\{A, C\}$  对应的决策函数判断时, 所属类为 C, 因此, 最终所属类为 C。需要说明的是, 该例子只是为了更简单化的说明判决的过程, 并不对本发明实施例起任何限定作用。

[0088] (3) 将每段的新的测量报告数据的最终所属类所对应的网格区域的中心经纬度作为该段所对应的时间间隔内待定位移动终端的定位结果。

[0089] 在实际的在线定位阶段, 将  $m/m_1$  段中的新的测量报告数据作为两分类决策函数中的 x, 代入每一个两分类决策函数进行计算时, 可以根据计算机上的硬件资源, 在计算机上开启若干个线程, 如开启 k 个线程, 将  $m/m_1$  段中的新的测量报告数据进行任务均分, 每个线程进行  $\frac{m}{km_1}$  条测量报告的定位操作。由于多线程可执行并行计算, 计算性能可成倍提升。

[0090] 采用本发明实施例所提供的方法, 对杭州某地区进行了 6 个月的数据采集, 使用本发明的方法对其中 5 个月数据进行处理即离线训练, 生成定位模型, 并采用如下配置的计算机对另一个月份数据进行定位。

[0091] 机器配置: 服务器, windows 操作系统, 12 核, 2.4GHz 主频 CPU

[0092] 定位性能如下表所示:

[0093]

开辟线程数	1	4	8	16	32
定位性能	2.38 万 /s	7.48 万 /s	10.67 万 /s	19.14 万 /s	29.85 万 /s

[0094] 由上表可以看到, 本发明提供的方法, 定位性能与计算机器开辟线程数成正比。在开辟线程数为 32 时, 可达到近 30 万条 MR 每秒的处理性能, 可充分满足实际应用系统的需求。

[0095] 显然, 本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的精神和范围。这样, 倘若本发明的这些修改和变型属于本发明权利要求及其同等技术的范围之内, 则本发明也意图包含这些改动和变型在内。

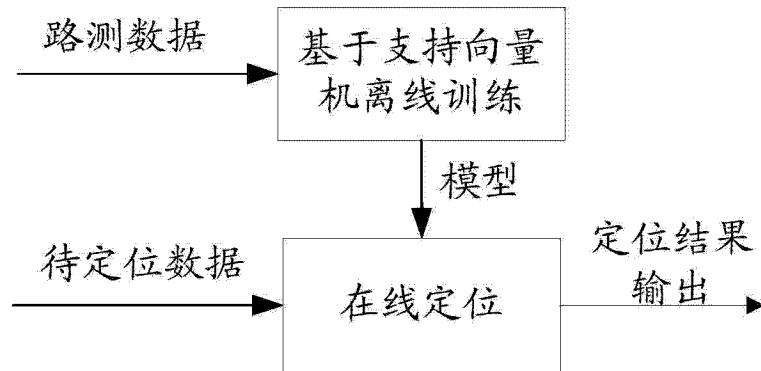


图 1

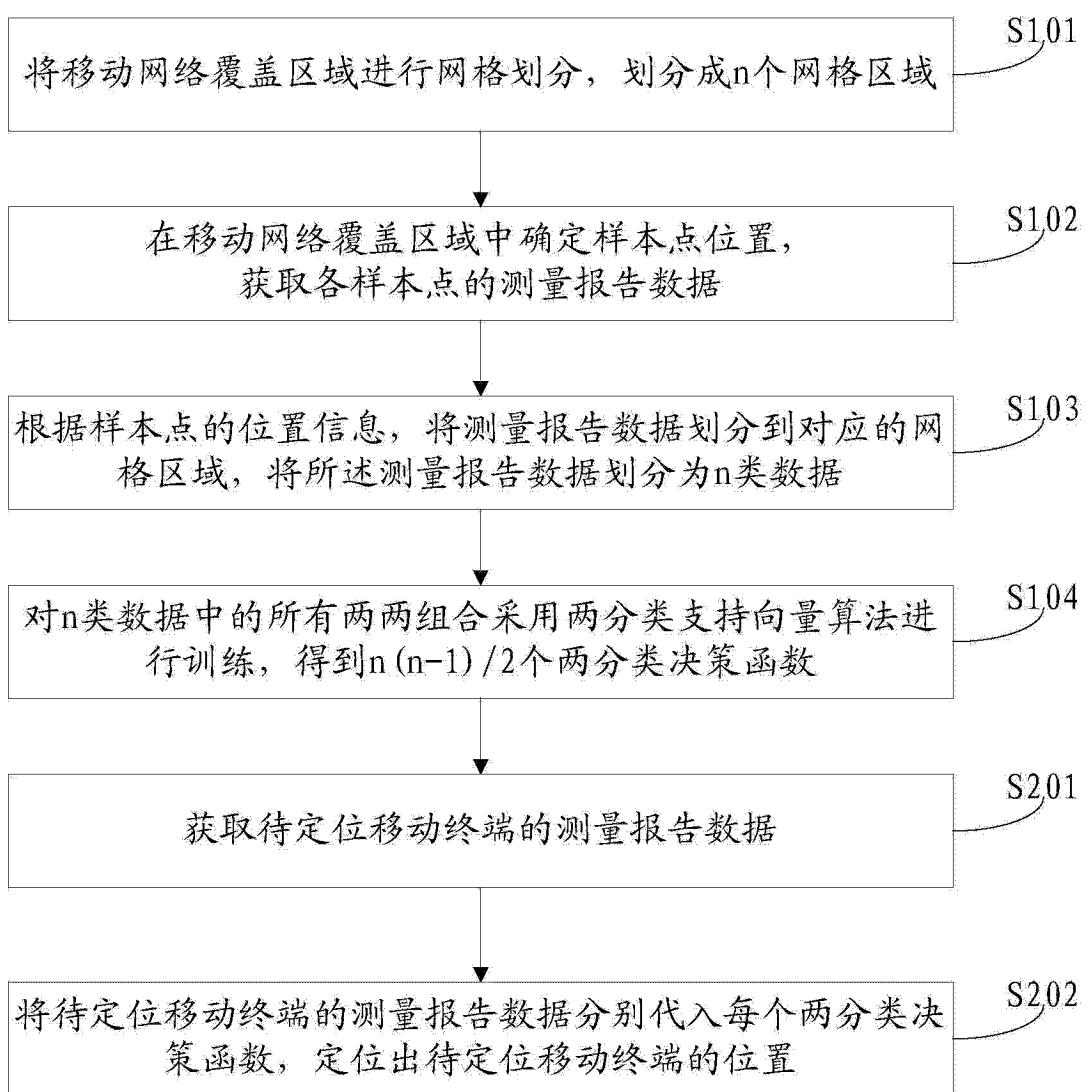


图 2

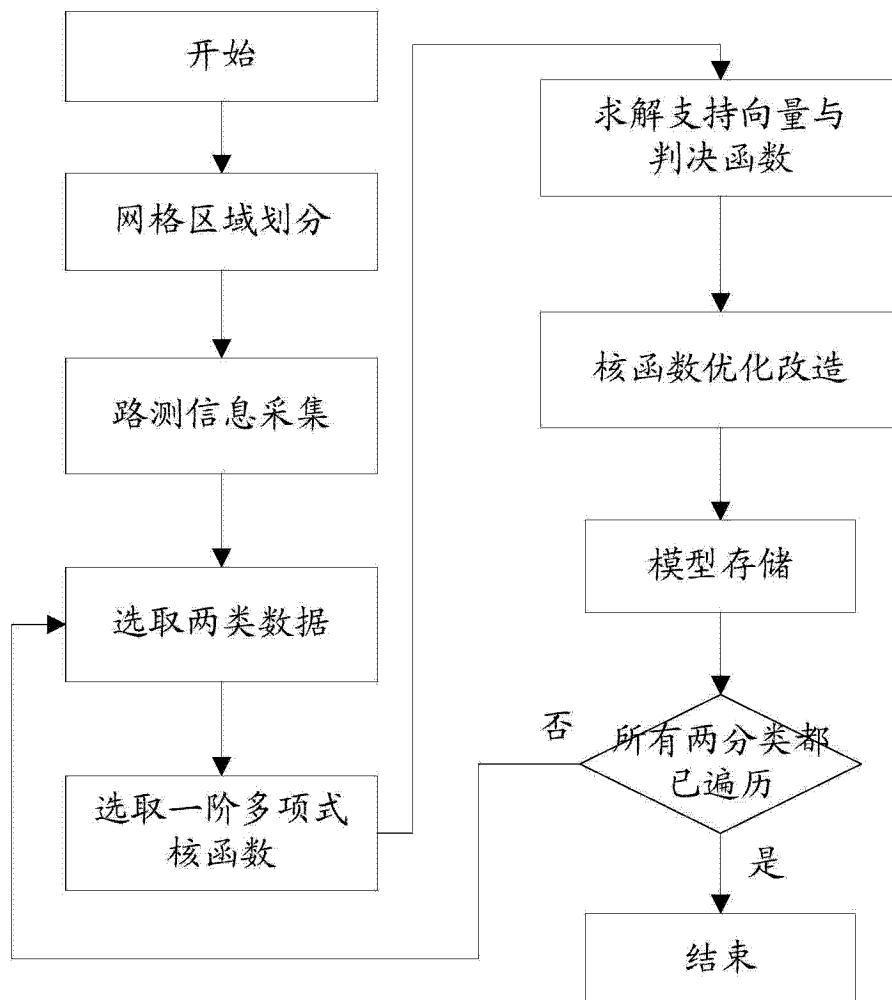


图 3

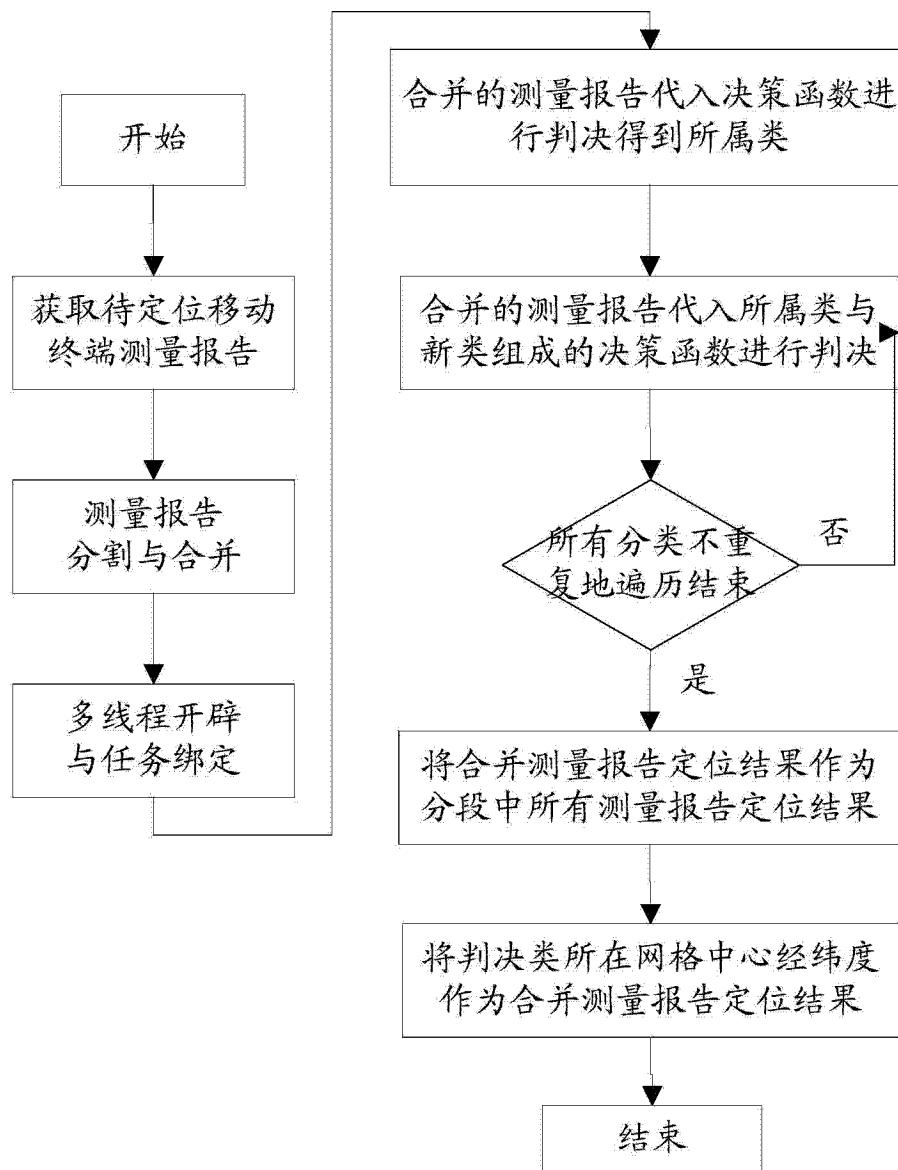


图 4