(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2016 年 8 月 18 日 (18.08.2016)

WIPO | PCT

(10) 国际公布号
WO 2016/127880 A1

(51) 国际专利分类号:
G01S 5/02 (2010.01)

(21) 国际申请号: PCT/CN2016/073271

(22) 国际申请日: 2016 年 2 月 3 日 (03.02.2016)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201510075882.2 2015 年 2 月 12 日 (12.02.2015) CN

(71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 英国开曼群岛大开曼乔治城资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。

(72) 发明人; 及

(71) 申请人 (仅对美国): 吴兴昊 (WU, Xinghao) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 刘欣 (LIU, Xin) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 房宇星 (FANG, Yuxing) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 邵茂材 (SHAO, Maocai) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 张杨 (ZHANG, Yang) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 林海涛 (LIN, Haitao) [CN/CN]; 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

(74) 代理人: 北京国昊天诚知识产权代理有限公司 (CO-HORIZON INTELLECTUAL PROPERTY INC.); 中国北京市朝阳区小关北里甲 2 号渔阳置业大厦 B 座 605, Beijing 100029 (CN)。

(81) 指定国 (除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR,

[见续页]

(54) Title: METHOD AND DEVICE FOR DETERMINING QUALITY OF OFFLINE POSITIONING DATA

(54) 发明名称 : 一种确定离线定位数据质量的方法和装置

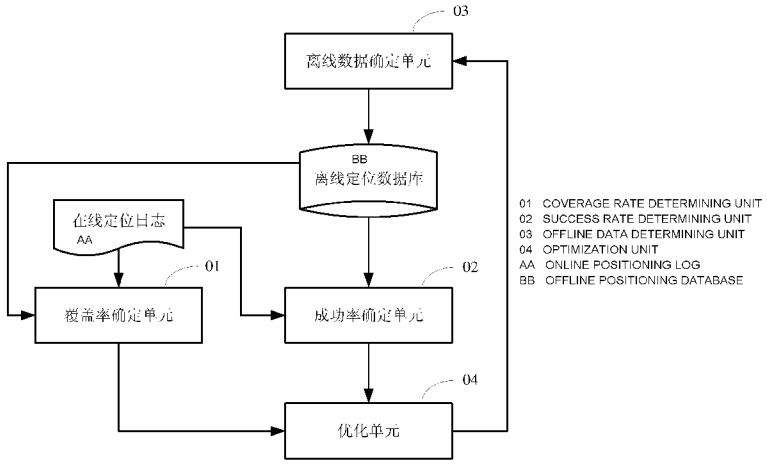


图 1

(57) Abstract: The present invention provides a method and device for determining the quality of offline positioning data. The method comprises: determining the coverage rate and/or success rate of offline positioning data according to an online positioning log on the side of a server and the offline positioning data, wherein the determining the coverage rate comprises: computing the total number X of records in the online positioning log within a set time period, computing the number Y of records, existing in an offline positioning database, of LAC identification information in the online positioning log within the set time period, and determining the coverage rate of the offline positioning data according to the ratio of Y to X; and the determining the success rate comprises: computing the total number X of records in the online positioning log within a set time period, computing the number Z of positioning results in the online positioning log within the set time period that are consistent with the offline positioning data, and determining the success rate of the offline positioning data according to the ratio of Z to X. By means of the present invention, the quality of offline positioning is measured according to the coverage rate and/or success rate of offline positioning data, so as to give guidance for optimization of the offline positioning data.

(57) 摘要:

[见续页]

WO 2016/127880 A1

CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) **指定国** (除另有指明，要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

本发明提供了一种确定离线定位数据质量的方法和装置，其中方法包括：利用服务器端的在线定位日志和离线定位数据，确定离线定位数据的覆盖率和/或成功率；其中覆盖率的确定包括：统计在线定位日志中设定时间段内的总记录数 X，以及统计在线定位日志中设定时间段内的 LAC 标识信息在离线定位数据库中存在的记录数目 Y，利用 Y 与 X 的比值确定离线定位数据的覆盖率；成功率的确定包括：统计在线定位日志中设定时间段内的总记录数 X，以及统计设定时间段内在线定位日志中的定位结果与离线定位数据一致的数量 Z，利用 Z 与 X 的比值确定离线定位数据的成功率。本发明通过离线定位数据的覆盖率和/或成功率衡量离线定位的质量，从而对离线定位数据的优化起指导作用。

一种确定离线定位数据质量的方法和装置

【技术领域】

本发明涉及计算机应用技术领域，特别涉及一种确定离线定位数据质量的方法和装置。

【背景技术】

随着移动设备的不断普及和发展，用户越来越多的使用移动设备进行通讯、信息获取、娱乐甚至工作。在用户享受移动设备的服务的过程中，往往会涉及到对移动设备定位的需求。目前对移动设备进行定位的常用方式主要是 GPS（Global Positioning System，全球定位系统）定位和基站定位。其中 GPS 定位需要移动终端具备 GPS 模块，且移动设备位置受到遮挡时常常无法定位。基站定位是移动设备将基站 ID（标识）通过网络请求发送给服务端，服务端根据数据库中基站的经纬度信息确定移动设备的位置，将移动设备的位置通过网络下发给移动设备。基站定位要求移动设备每次定位必须访问网络，属于在线定位（也称为网络定位）方式，会消耗网络流量，且没有网络情况下则无法实现定位。

考虑到 GPS 定位和基站定位的上述限制，提出了一种离线定位的设想，即在对定位精度要求不高的情况下，移动设备能够利用本地存储的离线定位数据进行定位。离线定位数据是离线定位的前提，其质量直接影响离线定位的质量，现有技术中并没有确定离线定位数据质量的方法，也就无法衡量离线定位的质量以及无法对离线定位数据的优化起指导作用。

【发明内容】

有鉴于此，本发明提供了一种确定离线定位数据质量的方法和装置，以便于能够衡量离线定位的质量，从而能够对离线定位数据的优化起指导作用。

具体技术方案如下：

本发明提供了一种确定离线定位数据质量的方法，该方法包括:

利用服务器端的在线定位日志和离线定位数据,确定离线定位数据的覆盖率和/或成功率;

其中所述确定离线定位数据的覆盖率包括:统计所述在线定位日志中设定时间段内的总记录数 X，以及统计在线定位日志中设定时间段内的 LAC 标识信息在离线定位数据库中存在的记录数目 Y，利用 Y 与 X 的比值确定离线定位数据的覆盖率;

所述确定离线定位数据的成功率包括:统计所述在线定位日志中设定时间段内的总记录数 X，以及统计设定时间段内在线定位日志中的定位结果与离线定位数据一致的数量 Z，利用 Z 与 X 的比值确定离线定位数据的成功率。

根据本发明一优选实施方式，所述 LAC 数据包括: LAC 标识信息及其对应的位置信息。

根据本发明一优选实施方式,所述 LAC 数据是根据 LAC 对基站数据进行合并后得到的，所述基站数据包括基站标识、基站所属 LAC 标识以及基站的位置信息。

根据本发明一优选实施方式，所述根据 LAC 对基站数据进行合并包括:

将属于同一 LAC 的基站数据基于位置信息进行聚类，得到至少一个类;

记录 LAC 的标识信息及其对应类的中心点位置信息，以及轮廓信息或半径信息。

根据本发明一优选实施方式,所述统计设定时间段内在线定位日志中的定位结果与离线定位数据一致的数量 Z 包括:

比较在线定位日志中某基站 ID 对应的定位结果与该基站 ID 所属 LAC 标识信息在离线定位数据中对应的位置信息是否一致，如果是，则统计该定位结果;

针对设定时间段内在线定位日志中的所有定位结果均执行所述比较的操作，最终统计的定位结果数量为所述 Z。

根据本发明一优选实施方式，在所述统计定位结果时，进一步记录离线定位

数据中与所述定位结果一致的位置信息的定位精度信息;

该方法进一步包括：从统计的定位结果中确定预设精度信息的定位结果数目，利用该数目与 X 的比值确定所述预设精度对应的离线定位数据的成功率。

根据本发明一优选实施方式，利用确定的覆盖率和/成功率，优化离线定位数据的确定方法。

根据本发明一优选实施方式，所述利用确定的覆盖率和/成功率，优化离线定位数据的确定方法包括:

针对确定离线定位数据时采用的不同聚类方法分别计算得到的离线定位数据的覆盖率和/或正确率，选择覆盖率和/或正确率满足预设覆盖率要求和/或正确率要求的聚类方式；或者，

依据采用某聚类方法得到的离线定位数据的覆盖率和/或成功率，调整所述某聚类方法中的聚类参数值，直至得到的离线定位数据的覆盖率和/或成功率满足预设覆盖率要求和/或成功率要求。

根据本发明一优选实施方式，该方法进一步包括:

根据对离线定位数据的数据量要求，调整所述覆盖率要求和/或成功率要求，从而选择出得到的离线定位数据的数据量能够满足所述数据量要求的聚类方式和/或聚类参数。

本发明还提供了一种确定离线定位数据质量的装置，用于利用服务器端的在线定位日志和离线定位数据，确定离线定位数据的覆盖率和/或成功率；该装置包括：覆盖率确定单元和/或成功率确定单元;

所述覆盖率确定单元，用于统计所述在线定位日志中设定时间段内的总记录数 X，以及统计在线定位日志中设定时间段内的 LAC 标识信息在离线定位数据库中存在的记录数目 Y，利用 Y 与 X 的比值确定离线定位数据的覆盖率;

所述成功率确定单元，用于统计所述在线定位日志中设定时间段内的总记录数 X，以及统计设定时间段内在线定位日志中的定位结果与离线定位数据一致的数量 Z，利用 Z 与 X 的比值确定离线定位数据的成功率。

根据本发明一优选实施方式，所述 LAC 数据包括：LAC 标识信息及其对应的位置信息。

根据本发明一优选实施方式，该装置还包括：离线数据确定单元，用于根据 LAC 对基站数据进行合并后得到所述 LAC 数据，所述基站数据包括基站标识、基站所属 LAC 标识以及基站的位置信息。

根据本发明一优选实施方式，所述离线数据确定单元，具体用于将属于同一 LAC 的基站数据基于位置信息进行聚类，得到至少一个类；记录 LAC 的标识信息及其对应类的中心点位置信息，以及轮廓信息或半径信息。

根据本发明一优选实施方式，所述成功率确定单元在统计所述 Z 时，具体执行：

比较在线定位日志中某基站 ID 对应的定位结果与该基站 ID 所属 LAC 标识信息在离线定位数据中对应的位置信息是否一致，如果是，则统计该定位结果；

针对设定时间段内在线定位日志中的所有定位结果均执行所述比较的操作，最终统计的定位结果数量为所述 Z。

根据本发明一优选实施方式，所述成功率确定单元在所述统计定位结果时，进一步记录离线定位数据中与所述定位结果一致的位置信息的定位精度信息；从统计的定位结果中确定预设精度信息的定位结果数目，利用该数目与 X 的比值确定所述预设精度对应的离线定位数据的成功率。

根据本发明一优选实施方式，该装置还包括：优化单元，用于利用所述覆盖率确定单元确定的覆盖率和/或成功率确定单元确定的成功率，优化离线定位数据的确定方法。

根据本发明一优选实施方式，所述优化单元，具体用于针对确定离线定位数据时采用的不同聚类方法分别计算得到的离线定位数据的覆盖率和/或正确率，选择覆盖率和/或正确率满足预设覆盖率要求和/或正确率要求的聚类方式；或者，

依据采用某聚类方法得到的离线定位数据的覆盖率和/或成功率，调整所述

某聚类方法中的聚类参数值，直至得到的离线定位数据的覆盖率和/或成功率满足预设覆盖率要求和/或成功率要求。

根据本发明一优选实施方式，所述优化单元，还用于根据对离线定位数据的数据量要求，调整所述覆盖率要求和/或成功率要求，从而选择出得到的离线定位数据的数据量能够满足所述数据量要求的聚类方式和/或聚类参数。

由以上技术方案可以看出，本发明利用服务器端的在线定位日志和离线定位数据，确定离线定位数据的覆盖率和/或成功率，利用覆盖率和/或成功率反映离线定位数据的质量，从而衡量离线定位的质量并对离线定位数据的优化起指导作用。

【附图说明】

图 1 为本发明实施例提供的装置结构图。

【具体实施方式】

为了使本发明的目的、技术方案和优点更加清楚，下面结合附图和具体实施例对本发明进行详细描述。

本发明的核心思想在于，利用服务器端的在线定位日志和离线定位数据库，确定离线定位数据的覆盖率和/或成功率，以此衡量离线定位数据的质量。并能够进一步依据离线定位数据的质量自适应地调节离线定位数据的确定方法。

本发明实施例所涉及的在线定位日志指的是对采用在线定位方式进行定位的行为记录，在线定位方式主要包括基站定位和 wifi 定位，因此，在线定位日志中可以包括：基站 ID 对应的定位结果、wifi 标识对应的定位结果等等，还可以包括定位时间，基站 ID 所属的 LAC（Location Area Code，位置区码）标识信息等。LAC 是唯一识别我国数字公共陆地移动网络中每个位置区的，是一个 2 字节 16 进制的 BCD 码。LAC 在每个小区广播信道上的系统消息中发送，移动设备在进行在线定位时获取基站 ID 的同时也会获取到该

基站所属的 LAC 标识信息。服务器端在进行在线定位日志的记录时，也会记录各基站 ID 所属的 LAC 标识信息。

离线定位数据是可以用于移动设备进行离线定位所使用的数据，在本发明实施例中离线定位数据可以包括 LAC 数据，该 LAC 数据可以包括 LAC 标识信息及其对应的位置信息。该离线定位数据可以是对基站数据进行合并后得到的。下面对离线定位数据的产生过程进行介绍。

每一个基站数据都包含基站 ID、所属 LAC、基站的位置信息（例如经纬度）等属性信息，其中基站的位置信息可以是从运营商处获取的基站位置信息，也可以是经过对基站定位结果（可以从在线定位日志中获取）中同一基站对应的定位结果进行聚类后得到的中心点位置信息。

对数据库中的基站数据基于 LAC 进行合并，得到属于同一 LAC 的所有基站数据。在进行合并时，可以采用聚类的方式。即将属于同一 LAC 的基站数据基于位置信息进行聚类，大多数下，属于同一 LAC 的基站数据会被聚类成同一类，记录该 LAC 对应类的中心点的位置信息，该位置信息可以是经纬度，也可以是所在的行政区划信息，例如在 xx 省 xx 市 xx 区，更进一步地，还可以记录 LAC 对应类的轮廓信息或覆盖区域的半径。

经过观察发现一些相同运营商、不同 MNC（移动网络号码，Mobile Network Code）、相同 LAC 对应的位置信息通常会相互重叠，若重叠范围较大，则可以考虑合并，从而减少位置信息的数量。基于此，可以将同一个运营商的不同 MNC、相同 LAC 的位置信息尽量合并，如果不能合并，则全部保留。例如，若同一 LAC 的两条位置信息分属于同一运营商的不同 MNC，则当两个位置信息的中心偏移较小（例如小于指定距离范围）时，可以将这两条位置信息合并为一条；当两条位置信息的中心偏移较大（大于指定距离范围）时，认为两条位置信息不能被合并，需要分别作为一条位置信息。通过合并有利于减少数据量，节约存储资源。

在有些情况下，例如 LAC 覆盖的区域过大，则可能会产生同一 LAC 的

基站数据会被聚类成多个类（本发明实施例中所述“多个”指代多于一个），此时可以对得到的类分别判断是否符合预设的属性标准，如果不符合，则将该类过滤掉。记录该 LAC 对应的符合属性标准的类的中心点的位置信息，还可以进一步记录符合属性标准的类的轮廓信息。

其中，上述属性标准可以采用但不限于以下所列中之一或任意组合:

1）类内基站的创建时间在最近的预设时间范围内。这种情况是为了防止有些基站创建时间很久，有可能实际的基站已经不再使用，但数据库中的基站数据未及时更新，对于这种情况产生的类可以将其过滤掉，不进行记录。

2）类的覆盖面积大于或等于预设的面积标准。有些基站离散于其他大多数基站而分布成一小类，用户在该类所覆盖区域的概率相比较其他类通常要低得多，往往是一些在距离市区很远的某些地方，基于某种特殊的需求而创建的基站，因此对于这种情况产生的类可以将其过滤掉。

3）类内基站的数量大于或等于预设的数量标准。通常对于用户较密集、需求量比较高的地方，通常会设置较多数量的基站。与上述第 2）种情况类似的，有些基站离离散于其他大多数基站而分布成一小类，用户在该类所覆盖区域的概率相比较其他类通常要低得多，因此对于这种情况产生的类可以将其过滤掉。

在对基站数据基于位置信息进行聚类时，可以采用诸如 K-means、K-medoids、DBSCAN（Density-Based Spatial Clustering of Applications with Noise，,具有噪声的基于密度的聚类方法）等聚类方法，本发明对聚类方式不加以限制。

最终得到的 LAC 数据可以包括: LAC 标识信息及其对应的位置信息，其中 LAC 对应的位置信息可以包括所覆盖区域的中心点位置信息,还可以包括所覆盖区域的轮廓信息或半径信息。其中中心点位置信息可以采用经纬度描述，也可以采用行政区划信息描述。

确定 LAC 数据所基于的基站数据可以是从运营商处获取，也可以通过

收集各种具有定位功能的应用的定位日志，从定位日志中获取。具有定位功能的应用可以是但不限于：各种地图。

本发明实施例中，离线定位数据的质量可以通过覆盖率和成功率中的至少一个来反映。

在确定离线定位数据的覆盖率时，首先统计在线定位日志中设定时间段内的总记录数 X，以及统计在线定位日志中设定时间段内的 LAC 标识信息在离线定位数据库中存在的记录数目 Y；然后利用 Y 与 X 的比值确定离线定位数据的覆盖率。例如，统计在线日志中一天内的总记录数，有 X 条定位记录，通常一天内在线定位日志的总记录数为亿次级别。这 X 条定位记录中的 LAC 标识信息在离线定位数据库中存在的记录数目为 Y 条，例如这 X 条定位记录中存在 LAC1、LAC2、LAC3、LAC4（可能每个 LAC 都存在于多条定位记录中），其中 LAC1、LAC2 和 LAC3 在离线定位数据库中存在记录，那么 Y 为 3。那么就可以将 $\frac{Y}{X}$ 的值作为离线定位数据的覆盖率。

在确定离线定位数据的成功率时，可以统计设定时间段内在线定位结果与离线定位数据一致的数量 Z，利用 Z 与 X 的比值确定离线定位数据的成功率。其中离线定位结果就是离线定位数据库中的位置信息。

具体地，可以将一天内的在线定位日志中的一条定位数据与离线定位数据进行比较，即首先确定该条定位数据中基站 ID 所属的 LAC 标识信息，再确定该 LAC 标识信息在离线定位数据中对应的位置信息，判断该条定位数据中的基站定位数据与 LAC 标识信息在离线定位数据中对应的位置信息一致，例如是否落入 LAC 标识信息对应的区域。举个例子，假设在线定位日志中的一条定位数据为：[基站 a，LAC1，（39.98，116.30）]，其中基站 a 为基站 ID，LAC1 为基站 a 所属的 LAC 标识，（39.98，116.30）为经纬度。利用该 LAC1 查询离线定位数据，假设得到的对应位置信息为：北京市海淀区，判断（39.98，116.30）是否落入北京市海淀区的范围，判断结果为是，说明该条定位数据与离线定位数据一致。对一天内的在线定位日志中的所有定位数

据都执行类似比较，统计得出一致结果的数目，该数目就是 Z。这里“一天”是预设时间段的一个例子。

另外，由于各 LAC 标识信息在离线定位数据库中对应的位置信息是存在不同精度的，即离线定位精度，该精度可以由区域半径或者行政区划信息（例如城市级别、区县级别等）来反映。因此可以分别针对不同精度统计出成功率。例如，在比较出在线定位日志中基站定位数据与 LAC 标识信息在离线定位数据中对应的位置信息一致时，除了统计得出一致结果的数目之外，进一步记录得出一致结果对应的离线定位精度。统计预设精度对应的一致结果数目，将该数目与 X 的比值确定该预设精度对应的离线定位数据的成功率。

在确定出离线定位数据的覆盖率和成功率后，可以自适应地优化离线定位数据的确定方法。其中对离线定位数据的确定方法进行的优化主要包括聚类方式的优化和聚类参数的优化。

1）优化离线定位数据所采用的聚类方式。

由之前描述的确定离线定位数据的过程中可以得知，离线定位数据是利用基站数据得到的，具体地，是基于位置信息对属于同一 LAC 的基站数据进行聚类后得到的。在聚类时，可以采用多种聚类方式，诸如 K-means、K-medois、DBSCAN 等聚类方法。

在本发明实施例中，可以分别计算不同聚类方式得到的离线定位数据的覆盖率和/正确率，选择覆盖率和/或正确率满足预设覆盖率要求和/或正确率要求的聚类方式。然后可以将利用选择的聚类方式得到的离线定位数据设置于移动设备，以便进行离线定位时使用。

2）优化确定离线定位数据所使用的聚类方法中的聚类参数。

当确定离线定位数据时，在基于位置信息对属于同一 LAC 的基站数据进行聚类时，执行聚类算法过程中会涉及到一些聚类参数，例如 k-means 算法中会涉及到每轮迭代得到的新的聚类中心与上一轮迭代得到的聚类中心之间的距离，当距离小于或等于预设距离阈值时，迭代结束。那么其中的距离

阈值就是一个聚类参数。DBScan 聚类方法涉及到的参数主要是扫描半径 eps 和最小包含点数 minPts，大概的聚类过程为：任选一个未被访问的点开始，找出与其距离在 eps 之内的所有附近点，如果附近点的数量大于或等于 minPts，则当前点与其附近点形成一个类，并且触发点被标记为已访问，然后递归，以相同的方法处理该类内所有未被标记为已访问的点，从而对类进行扩展。如果附近点的数量小于 minPts，则该触发点被暂时标记为噪声点。

当确认使用一种聚类方法来对属于同一 LAC 的基站数据进行聚类从而得到离线定位数据时，可以对得到的离线定位数据确定覆盖率和/或成功率，依据得到的覆盖率和/或成功率调整聚类方法中的聚类参数值，直至得到的离线定位数据的覆盖率和/或成功率满足预设覆盖率要求和/或成功率要求。然后采用调整后的聚类参数值进行聚类后得到离线定位数据，使用该离线定位数据配置于移动设备，以便进行离线定位时使用。

更进一步地，根据应用需要的定位精度不同，可以分别针对不同定位精度来分别优化聚类方法和聚类参数，从而得到针对不同定位精度的离线定位数据。

例如，可以针对城市级别的精度分别计算离线定位数据的成功率和覆盖率，采用上述第 1）或第 2）所述的方式选择聚类方法和调整聚类参数，并利用选择的聚类方法和调整后的聚类参数得到城市级别的离线定位数据。再针对区县级别的精度分别计算离线定位数据的成功率和覆盖率，采用上述第 1）或第 2）所述的方式选择聚类方法和调整聚类参数，并利用选择的聚类方法和调整后的聚类参数得到城市区县级别的离线定位数据。

然后可以将城市级别的离线定位数据和区县级别的离线定位数据配置于移动设备，以便移动设备中的应用需要城市级别的定位精度时，采用城市级别的离线定位数据进行离线定位，当移动设备中的应用需要区县级别的定位精度时，采用区县级别的离线定位数据进行离线定位。

更进一步地，还可以根据对离线定位数据的数据量的要求，调整对覆盖

率和/或成功率的预设要求，从而选择出得到的离线定位数据的数量能够满足要求的聚类方式和聚类参数。由于移动设备的存储量是限制的，那么对离线定位数据的数据量也会有一定要求，如果数据量太大移动设备可能无法存储或对性能有很大影响。通常覆盖率越高离线定位数据的数据量越大，成功率越高离线定位数据的数据量越大。因此，在选择聚类方式和调整聚类参数时，可以调整对离线定位数据库的覆盖率和/或成功率的要求，从而选择聚类方式和调整聚类参数使得在满足离线定位数据库的覆盖率和/或成功率的要求同时，满足离线定位数据的数据量要求。

图 1 为本发明实施例提供的装置结构图，该装置用于利用服务器端的在线定位日志和离线定位数据，确定离线定位数据的覆盖率和/或成功率，以此衡量离线定位数据的质量。如图 1 所示，该装置具体包括：覆盖率确定单元 01 和/或成功率确定单元 02，图 1 中以同时包含这两个单元为例，还可以包括离线数据确定单元 03 和优化单元 04。

覆盖率确定单元 01 负责统计在线定位日志中设定时间段内的总记录数 X，以及统计在线定位日志中设定时间段内的 LAC 标识信息在离线定位数据库中存在的记录数目 Y，利用 Y 与 X 的比值确定离线定位数据的覆盖率。

成功率确定单元 02 负责统计在线定位日志中设定时间段内的总记录数 X，以及统计设定时间段内在线定位日志中的定位结果与离线定位数据一致的数量 Z，利用 Z 与 X 的比值确定离线定位数据的成功率。

具体地，可以将一天内的在线定位日志中的一条定位数据与离线定位数据进行比较，即首先确定该条定位数据中基站 ID 所属的 LAC 标识信息，再确定该 LAC 标识信息在离线定位数据中对应的位置信息，判断该条定位数据中的基站定位数据与 LAC 标识信息在离线定位数据中对应的位置信息一致，例如是否落入 LAC 标识信息对应的区域。对一天内的在线定位日志中的所有定位数据都执行类似比较，统计得出一致结果的数目，该数目就是 Z。

另外，成功率确定单元 02 在统计定位结果时，可以进一步记录离线定位数

据中与定位结果一致的位置信息的定位精度信息；从统计的定位结果中确定预设精度信息的定位结果数目，利用该数目与 X 的比值就能够确定预设精度对应的离线定位数据的成功率。

离线定位数据库中的 LAC 数据包括：LAC 标识信息及其对应的位置信息。该 LAC 数据可以由离线数据确定单元 03 得到。具体地，离线数据确定单元 03 根据 LAC 对基站数据进行合并后得到 LAC 数据，基站数据包括基站标识、基站所属 LAC 标识以及基站的位置信息。

离线数据确定单元03可以将属于同一LAC的基站数据基于位置信息进行聚类，得到至少一个类；记录 LAC 的标识信息及其对应类的中心点位置信息，以及轮廓信息或半径信息。得到 LAC 数据过程的更多细节可以参见上述方法实施例中的描述。

在确定出离线定位数据的覆盖率和成功率后，可以自适应地优化离线定位数据的确定方法，该优化处理由优化单元 04 执行。具体地，该优化处理可以包括聚类方法的优化和聚类参数的优化。

对于聚类方法的优化，优化单元 04 可以针对确定离线定位数据时采用的不同聚类方法分别计算得到的离线定位数据的覆盖率和/或正确率，选择覆盖率和/或正确率满足预设覆盖率要求和/或正确率要求的聚类方式。

对于聚类参数的优化，优化单元 04 可以依据采用某聚类方法得到的离线定位数据的覆盖率和/或成功率，调整某聚类方法中的聚类参数值，直至得到的离线定位数据的覆盖率和/或成功率满足预设覆盖率要求和/或成功率要求。

更进一步地，优化单元 04 可以根据对离线定位数据的数据量要求，调整覆盖率要求和/或成功率要求，从而选择出得到的离线定位数据的数据量能够满足数据量要求的聚类方式和/或聚类参数。由于移动设备的存储量是限制的，那么对离线定位数据的数据量也会有一定要求，如果数据量太大移动设备可能无法存储或对性能有很大影响。通常覆盖率越高离线定位数据的数据量越大，成功率越高离线定位数据的数据量越大。因此，在选择聚类方式和

调整聚类参数时，可以调整对离线定位数据库的覆盖率和/或成功率的要求，从而选择聚类方式和调整聚类参数使得在满足离线定位数据库的覆盖率和/或成功率的要求同时，满足离线定位数据的数据量要求。

本发明提供的装置可以设置于服务器端，也可以设置为独立的装置。

由以上描述可以看出，本发明提供的方法和装置可以具备以下优点：

1）本发明利用服务器端的在线定位日志和离线定位数据，确定离线定位数据的覆盖率和/或成功率，利用覆盖率和/或成功率反映离线定位数据的质量，从而实现离线定位的质量衡量，并能够对离线定位数据的优化起指导作用。

2）本发明利用离线定位数据的覆盖率和/或成功率能够对确定离线定位数据的方法进行优化，即不需要依赖移动设备的反馈，在服务器端就能形成闭环反向进行数据优化，提高离线定位的质量。

在本发明所提供的几个实施例中，应该理解到，所揭露的装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本发明各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用硬件加软件功能单元的形式实现。

上述以软件功能单元的形式实现的集成的单元，可以存储在一个计算机可读取存储介质中。上述软件功能单元存储在一个存储介质中，包括若干指

令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）或处理器（processor）执行本发明各个实施例所述方法的部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（Read-Only Memory，ROM）、随机存取存储器（Random Access Memory，RAM）、磁碟或者光盘等各种可以存储程序代码的介质。

以上所述仅为本发明的较佳实施例而已，并不用以限制本发明，凡在本发明的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本发明保护的范围之内。

权 利 要 求 书

1、一种确定离线定位数据质量的方法，其特征在于，该方法包括:

利用服务器端的在线定位日志和离线定位数据,确定离线定位数据的覆盖率和/或成功率;

其中所述确定离线定位数据的覆盖率包括:统计所述在线定位日志中设定时间段内的总记录数 X，以及统计在线定位日志中设定时间段内的 LAC 标识信息在离线定位数据库中存在的记录数目 Y，利用 Y 与 X 的比值确定离线定位数据的覆盖率;

所述确定离线定位数据的成功率包括:统计所述在线定位日志中设定时间段内的总记录数 X，以及统计设定时间段内在线定位日志中的定位结果与离线定位数据一致的数量 Z，利用 Z 与 X 的比值确定离线定位数据的成功率。

2、根据权利要求 1 所述的方法，其特征在于，所述 LAC 数据包括: LAC 标识信息及其对应的位置信息。

3、根据权利要求 2 所述的方法，其特征在于，所述 LAC 数据是根据 LAC 对基站数据进行合并后得到的，所述基站数据包括基站标识、基站所属 LAC 标识以及基站的位置信息。

4、根据权利要求 3 所述的方法，其特征在于，所述根据 LAC 对基站数据进行合并包括:

将属于同一 LAC 的基站数据基于位置信息进行聚类，得到至少一个类;

记录 LAC 的标识信息及其对应类的中心点位置信息，以及轮廓信息或半径信息。

5、根据权利要求 2 所述的方法，其特征在于，所述统计设定时间段内在线定位日志中的定位结果与离线定位数据一致的数量 Z 包括:

比较在线定位日志中某基站 ID 对应的定位结果与该基站 ID 所属 LAC 标识信息在离线定位数据中对应的位置信息是否一致，如果是，则统计该定位结果;

针对设定时间段内在线定位日志中的所有定位结果均执行所述比较的操作，

最终统计的定位结果数量为所述Z。

6、根据权利要求5所述的方法，其特征在于，在所述统计定位结果时，进一步记录离线定位数据中与所述定位结果一致的位置信息的定位精度信息；

该方法进一步包括：从统计的定位结果中确定预设精度信息的定位结果数目，利用该数目与X的比值确定所述预设精度对应的离线定位数据的成功率。

7、根据权利要求1至6任一权项所述的方法，其特征在于，利用确定的覆盖率和/成功率，优化离线定位数据的确定方法。

8、根据权利要求7所述的方法，其特征在于，所述利用确定的覆盖率和/成功率，优化离线定位数据的确定方法包括：

针对确定离线定位数据时采用的不同聚类方法分别计算得到的离线定位数据的覆盖率和/或正确率，选择覆盖率和/或正确率满足预设覆盖率要求和/或正确率要求的聚类方式；或者，

依据采用某聚类方法得到的离线定位数据的覆盖率和/或成功率，调整所述某聚类方法中的聚类参数值，直至得到的离线定位数据的覆盖率和/或成功率满足预设覆盖率要求和/或成功率要求。

9、根据权利要求8所述的方法，其特征在于，该方法进一步包括：

根据对离线定位数据的数据量要求，调整所述覆盖率要求和/或成功率要求，从而选择出得到的离线定位数据的数据量能够满足所述数据量要求的聚类方式和/或聚类参数。

10、一种确定离线定位数据质量的装置，其特征在于，用于利用服务器端的在线定位日志和离线定位数据，确定离线定位数据的覆盖率和/或成功率；该装置包括：覆盖率确定单元和/或成功率确定单元；

所述覆盖率确定单元，用于统计所述在线定位日志中设定时间段内的总记录数X，以及统计在线定位日志中设定时间段内的LAC标识信息在离线定位数据库中存在的记录数目Y，利用Y与X的比值确定离线定位数据的覆盖率；

所述成功率确定单元，用于统计所述在线定位日志中设定时间段内的总记

录数 X，以及统计设定时间段内在线定位日志中的定位结果与离线定位数据一致的数量 Z，利用 Z 与 X 的比值确定离线定位数据的成功率。

11、根据权利要求 10 所述的装置，其特征在于，所述 LAC 数据包括：LAC 标识信息及其对应的位置信息。

12、根据权利要求 11 所述的装置，其特征在于，该装置还包括：离线数据确定单元，用于根据 LAC 对基站数据进行合并后得到所述 LAC 数据，所述基站数据包括基站标识、基站所属 LAC 标识以及基站的位置信息。

13、根据权利要求 12 所述的装置，其特征在于，所述离线数据确定单元，具体用于将属于同一 LAC 的基站数据基于位置信息进行聚类，得到至少一个类；记录 LAC 的标识信息及其对应类的中心点位置信息，以及轮廓信息或半径信息。

14、根据权利要求 11 所述的装置，其特征在于，所述成功率确定单元在统计所述 Z 时，具体执行：

比较在线定位日志中某基站 ID 对应的定位结果与该基站 ID 所属 LAC 标识信息在离线定位数据中对应的位置信息是否一致，如果是，则统计该定位结果；

针对设定时间段内在线定位日志中的所有定位结果均执行所述比较的操作，最终统计的定位结果数量为所述 Z。

15、根据权利要求 14 所述的装置，其特征在于，所述成功率确定单元在所述统计定位结果时，进一步记录离线定位数据中与所述定位结果一致的位置信息的定位精度信息；从统计的定位结果中确定预设精度信息的定位结果数目，利用该数目与 X 的比值确定所述预设精度对应的离线定位数据的成功率。

16、根据权利要求 10 至 15 任一权项所述的装置，其特征在于，该装置还包括：优化单元，用于利用所述覆盖率确定单元确定的覆盖率和/或成功率确定单元确定的成功率，优化离线定位数据的确定方法。

17、根据权利要求 16 所述的装置，其特征在于，所述优化单元，具体用于针对确定离线定位数据时采用的不同聚类方法分别计算得到的离线定位数据的覆盖率和/或正确率，选择覆盖率和/或正确率满足预设覆盖率要求和/或正确率要

求的聚类方式；或者，

依据采用某聚类方法得到的离线定位数据的覆盖率和/或成功率，调整所述某聚类方法中的聚类参数值，直至得到的离线定位数据的覆盖率和/或成功率满足预设覆盖率要求和/或成功率要求。

18、根据权利要求 17 所述的装置，其特征在于，所述优化单元，还用于根据对离线定位数据的数据量要求，调整所述覆盖率要求和/或成功率要求，从而选择出得到的离线定位数据的数据量能够满足所述数据量要求的聚类方式和/或聚类参数。

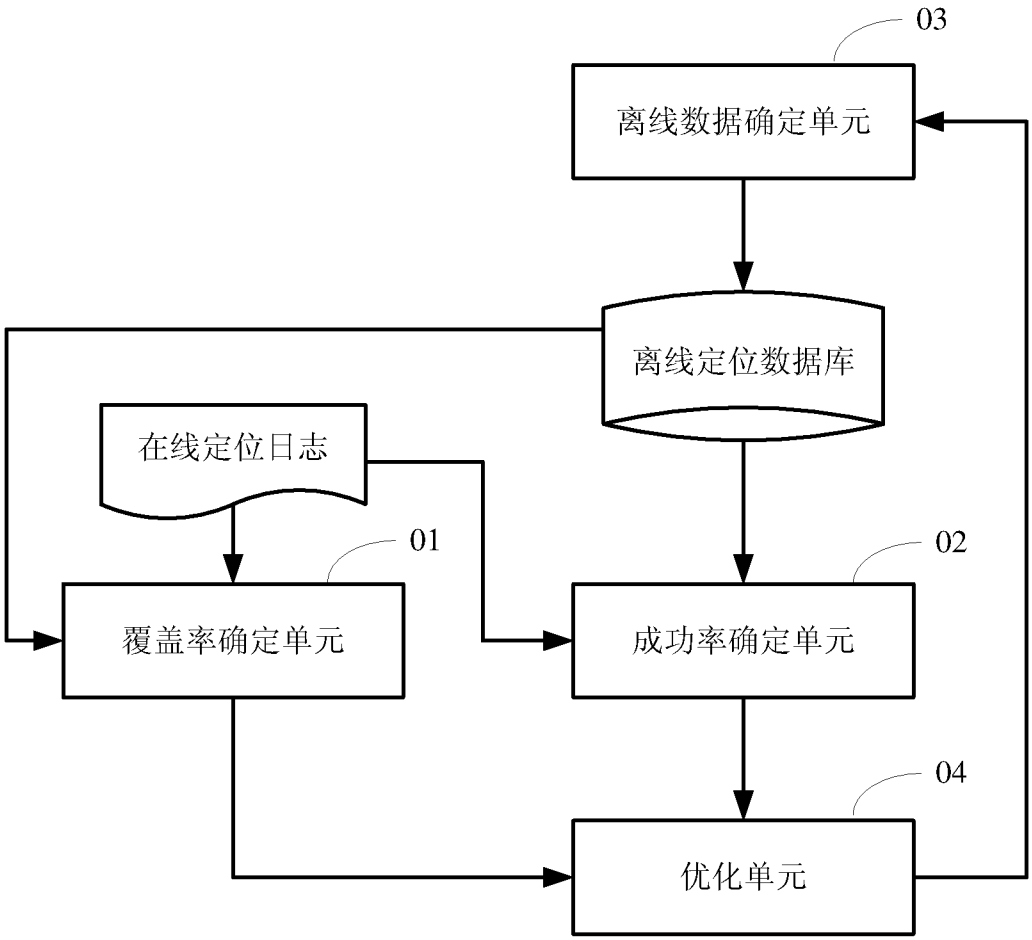


图 1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/073271

A. CLASSIFICATION OF SUBJECT MATTER

G01S 5/02 (2010.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G01S; H04W; H04Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, CNKI, WPI, EPODOC: coverage rate, success rate, area code, consistent, locat+, position+, off-line, offline, on-line, online, quality, coverage, success, ratio, LAC, same

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2015014544 A1 (HERE GLOBAL B.V.), 05 February 2015 (05.02.2015), description, paragraphs 0018-0020 and 0035-0054	1-18
A	WO 2013084032 A1 (NOKIA CORPORATION), 13 June 2013 (13.06.2013), the whole document	1-18
A	CN 103634902 A (SHANGHAI JIAO TONG UNIVERSITY), 12 March 2014 (12.03.2014), the whole document	1-18

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
07 April 2016 (07.04.2016)	**28 April 2016 (28.04.2016)**
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer **HE, Xijia** Telephone No.: (86-10) **62413281**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/073271

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
WO 2015014544 A1	05 February 2015	GB 2516832 A	11 February 2015
		CN 105431748 A	23 March 2016
WO 2013084032 A1	13 June 2013	None	
CN 103634902 A	12 March 2014	None	

国际检索报告

国际申请号
PCT/CN2016/073271

A. 主题的分类

G01S 5/02(2010.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G01S；H04W；H04Q

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用））

CNPAT, CNKI, WPI, EPODOC：定位，离线，在线，质量，覆盖率，成功率，比值，比率，区码，区域码，一致，相同，一样，locat+, position+, off-line, offline, on-line, online, quality, coverage, success, ratio, LAC, same

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	WO 2015014544 A1 (HERE GLOBAL B.V.) 2015年 2月 5日 (2015-02-05) 说明书第0018-0020，0035-0054段	1-18
A	WO 2013084032 A1 (NOKIA CORPORATION) 2013年 6月 13日 (2013-06-13) 全文	1-18
A	CN 103634902 A (上海交通大学) 2014年 3月 12日 (2014-03-12) 全文	1-18

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2016年 4月 7日	2016年 4月 28日
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	受权官员 贺希佳 电话号码 (86-10)62413281

表 PCT/ISA/210 (第2页) (2009年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2016/073271

检索报告引用的专利文件	公布日 (年/月/日)	同族专利	公布日 (年/月/日)
WO 2015014544 A1	2015年 2月 5日	GB 2516832 A	2015年 2月 11日
		CN 105431748 A	2016年 3月 23日
WO 2013084032 A1	2013年 6月 13日	无	
CN 103634902 A	2014年 3月 12日	无	