(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
2019 年 7 月 4 日 (04.07.2019)

WIPO | PCT

(10) 国际公布号
WO 2019/128355 A1

(51) 国际专利分类号:
H04L 29/06 (2006.01)

(21) 国际申请号: PCT/CN2018/108635

(22) 国际申请日: 2018 年 9 月 29 日 (29.09.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201711481337.9 2017年12月29日 (29.12.2017) CN

(71) 申请人: 北京京东尚科信息技术有限公司 (**BEIJING JINGDONG SHANGKE INFORMATION TECHNOLOGY CO., LTD.**) [CN/CN]; 中国北京市海淀区杏石口路65号西杉创意园四区11号楼东段1-4层西段1-4层, Beijing 100195 (CN)。
北京京东世纪贸易有限公司(**BEIJING JINGDONG CENTURY TRADING CO., LTD.**) [CN/CN]; 中国北京市北京经济技术开发区科创十一街18号C座2层201室, Beijing 100176 (CN)。

(72) 发明人: 肖明科(**XIAO, Mingke**); 中国北京市海淀区杏石口路65号西杉创意园四区11号楼东段1-4层西段1-4层, Beijing 100195 (CN)。

(74) 代理人: 中原信达知识产权代理有限责任公司 (**CHINA SINDA INTELLECTUAL PROPERTY LTD.**); 中国北京市西城区金融街19号富凯大厦B座11层, Beijing 100033 (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, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, 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,

(54) **Title:** METHOD AND DEVICE FOR DETERMINING ACCURATE GEOGRAPHIC LOCATION

(54) **发明名称**: 确定精确地理位置的方法和装置

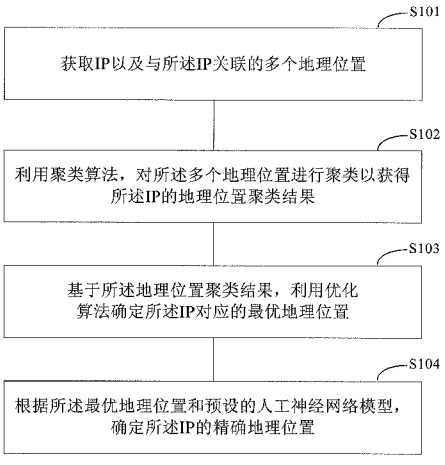

图 1

S101 Acquire an IP and a plurality of geographic locations associated with the IP
S102 Cluster the plurality of geographic locations by using a clustering algorithm, so as to obtain a geographic location clustering result of the IP
S103 Determine, according to the geographic location clustering result, an optimal geographic location corresponding to the IP by using an optimization algorithm
S104 Determine a precise geographic location of the IP according to the optimal geographic location and a preset artificial neural network model

(57) **Abstract:** The present invention relates to the technical field of the Internet, and disclosed thereby are a method and device for determining an accurate geographic location. A preferred embodiment of the method comprises: acquiring an IP and a plurality of geographic locations associated with the IP; clustering the plurality of geographic locations by using a clustering algorithm, so as to obtain a geographic location clustering result of the IP; determining, according to the geographic location clustering result, an optimal geographic location corresponding to the IP by using an optimization algorithm; and determining a precise geographic location of the IP according to the optimal geographic location and a preset artificial neural network model. According to said preferred embodiment, the precise geographical location of the IP may be determined, thereby improving positioning accuracy without needing to install a large number of monitoring points, so that cost is reduced while the positioning accuracy is increased.

(57) **摘要**: 本发明公开了一种确定精确地理位置的方法和装置，涉及互联网技术领域。该方法的一具体实施方式包括：获取IP以及与所述IP关联的多个地理位置；利用聚类算法，对所述多个地理位置进行聚类以获得所述IP的地理位置聚类结果；基于所述地理位置聚类结果，利用优化算法确定所述IP对应的最优地理位置；根据所述最优地理位置和预设的人工神经网络模型，确定所述IP的精确地理位置。该实施方式能够确定IP的精确地理位置，提高了定位精度，而且不需要铺设大量监测点，进而在提高定位精度的同时降低了成本。

WO 2019/128355 A1

[见续页]

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))。

确定精确地理位置的方法和装置

技术领域

本申请是以申请号为 201711481337.9，申请日为 2017 年 12 月 29 日的中国申请为基础，并主张其优先权，该中国申请的内容在此作为整体引入本申请中。

本发明涉及互联网技术领域，尤其涉及一种确定精确地理位置的方法和装置。

背景技术

IP 定位技术，简而言之，是通过设备的 IP 地址来确定其地理位置的技术。IP 定位具有极其广泛的应用，主要包括定向广告、社交网络、网络安全、性能优化等。在移动互联网的大背景下，手机等包含 GPS 信息模块的终端设备，通过数据上报能够很容易获取用户的街道级别的地理位置。但是如果是台式电脑、笔记本等不含 GPS 硬件设备的终端，就无法通过 GPS 等技术获取用户的地理位置，这时候就需要使用高精度 IP 定位技术。而传统的 IP 定位只能定位到市级，区级数据的准确性也值得商榷。

传统的 IP 定位算法，根据时延与地理距离之间的线性关系来估测位置，并通过拓扑结构减小误差。

具体的，是基于 BGP（Border Gateway Protocol 边界网关协议）/ASN（Autonomous System Number 自治系统号）数据分析而得来的，同时在全球自建网络监测点，根据待定位 IP 与监测点之间的网络返回延时值划分网络拓扑结构，从而进一步确认待定位 IP 的地理位置，此种方式定位较为可信，但是精度仍然不高（区级精度）。

在实现本发明过程中，发明人发现现有技术中至少存在如下问题：

此技术需要铺设足够多的监测点，用于确认 IP 的物理地址，成本较高，而且需要较为复杂的步骤，并且由于是通过网络链路延时来反推地理位置，此种方式定位虽然较为可信，但是精度仍然不高。

发明内容

有鉴于此，本发明实施例提供一种确定精确地理位置的方法和装置，提高了定位精度，而且本发明不需要铺设大量监测点，进而在提高定位精度的同时降低了成本。

为实现上述目的，根据本发明实施例的一个方面，提供了一种确定精确地理位置的方法，包括：获取 IP 以及与所述 IP 关联的多个地理位置；利用聚类算法，对所述多个地理位置进行聚类以获得所述 IP 的地理位置聚类结果；基于所述地理位置聚类结果，利用优化算法确定所述 IP 对应的最优地理位置；根据所述最优地理位置和预设的人工神经网络模型，确定所述 IP 的精确地理位置。

可选地，所述聚类算法为 k-means 算法，所述优化算法为加权最小二乘法。

可选地，利用聚类算法，对所述多个地理位置进行聚类以获得所述 IP 的地理位置聚类结果的步骤包括：从所述 IP 关联的多个地理位置中选取两个地理位置作为第一初始质心和第二初始质心；计算所述多个地理位置中每个地理位置与所述第一初始质心之间的第一球面距离以及与第二初始质心之间的第二球面距离；根据所述第一球面距离和所述第二球面距离，对所述 IP 关联的多个地理位置进行聚类以获得高密度簇，以所述高密度簇作为所述 IP 的地理位置聚类结果。

可选地，根据如下公式（1）计算每个地理位置与所述第一初始质

心之间的第一球面距离以及与第二初始质心之间的第二球面距离：

$$S = R \bullet ar\cos(\cos\beta 1 \bullet \cos\beta 2 \bullet \cos(\alpha 1 - \alpha 2) + \sin\beta 1 \bullet \sin\beta 2) \quad (1)$$

其中，R 表示地球长轴半径，S 表示地理位置 A 与地理位置 B 之间的球面距离，$\beta 1$为地理位置 A 的纬度，$\alpha 1$为地理位置 A 的经度，$\beta 2$为地理位置 B 的纬度，$\alpha 2$为地理位置 B 的经度。

可选地，根据所述地理位置聚类结果，利用优化算法确定每个 IP 对应的最优地理位置包括：对于高密度簇中的每个地理位置，根据每个地理位置与高密度簇质心的球面距离，确定所述每个地理位置的权重；根据所述权重，利用加权最小二乘法确定每个 IP 对应的最优地理位置。

可选地，根据下式（2）确定所述每个地理位置的权重：

$$\lambda_i = \frac{\frac{1}{d_i}}{\sum_{i=1}^{n}\frac{1}{d_i}} \quad (2)$$

其中，λ_i表示第 i 个地理位置的权重，d_i表示第 i 个地理位置与高密度簇质心之间的球面距离，n 为大于或等于 1 的整数；

根据下式（3）确定所述 IP 对应的最优地理位置：

$$Min = \sum_{i=1}^{n}\lambda_i \bullet ((x_i - \hat{x})^2 + (y_i - \hat{y})^2) \quad (3)$$

其中，(x_i, y_i)表示第 i 个地理位置，$(\hat{x}, \hat{y})$表示最优地理位置。

可选地，根据所述最优地理位置和预设的人工神经网络模型，确定所述 IP 的精确地理位置包括：将所述最优地理位置输入所述预设的人工神经网络模型，获取输出结果；若所述输出结果为预设的目标结果，则所述最优地理位置为所述 IP 的精确地理位置。

可选地，所述预设的人工神经网络模型的输入层具有 3 个神经元

节点，隐含层具有 5 个神经元节点，输出层具有 1 个神经元节点。

为实现上述目的，根据本发明实施例的一个方面，提供了一种确定精确地理位置的装置，包括：获取模块，用于获取 IP 以及与所述 IP 关联的多个地理位置；聚类模块，用于利用聚类算法，对所述多个地理位置进行聚类以获得所述 IP 的地理位置聚类结果；最优地理位置确定模块，用于基于所述地理位置聚类结果，利用优化算法确定所述 IP 对应的最优地理位置；精确地理位置确定模块，用于根据所述最优地理位置和预设的人工神经网络模型，确定所述 IP 的精确地理位置。

可选地，所述聚类算法为 k-means 算法，所述优化算法为加权最小二乘法。

可选地，所述聚类模块还用于：从所述 IP 关联的多个地理位置中选取两个地理位置作为第一初始质心和第二初始质心；计算所述多个地理位置中每个地理位置与所述第一初始质心之间的第一球面距离以及与第二初始质心之间的第二球面距离；根据所述第一球面距离和所述第二球面距离，对所述 IP 关联的多个地理位置进行聚类以获得高密度簇，以所述高密度簇作为所述 IP 的地理位置聚类结果。

可选地，所述聚类模块根据如下公式（1）计算每个地理位置与所述第一初始质心之间的第一球面距离以及与第二初始质心之间的第二球面距离：

$$S = R \bullet ar\cos(\cos\beta 1 \bullet \cos\beta 2 \bullet \cos(\alpha 1 - \alpha 2) + \sin\beta 1 \bullet \sin\beta 2) \quad (1)$$

其中，R 表示地球长轴半径，S 表示地理位置 A 与地理位置 B 之间的球面距离，$\beta 1$为地理位置 A 的纬度，$\alpha 1$为地理位置 A 的经度，$\beta 2$为地理位置 B 的纬度，$\alpha 2$为地理位置 B 的经度。

可选地，所述最优地理位置确定模块还用于：对于高密度簇中的每个地理位置，根据每个地理位置与高密度簇质心的球面距离，确定

所述每个地理位置的权重；根据所述权重，利用加权最小二乘法确定每个 IP 对应的最优地理位置。

可选地，根据下式（2）确定所述每个地理位置的权重：

$$\lambda_i = \frac{\frac{1}{d_i}}{\sum_{i=1}^{n}\frac{1}{d_i}} \qquad (2)$$

其中，λ_i表示第 i 个地理位置的权重，d_i表示第 i 个地理位置与高密度簇质心之间的球面距离，n 为大于或等于 1 的整数；

根据下式（3）确定所述 IP 对应的最优地理位置：

$$\mathrm{Min} = \sum_{i=1}^{n} \lambda_i \bullet ((x_i - \hat{x})^2 + (y_i - \hat{y})^2) \qquad (3)$$

其中，（x_i, y_i）表示第 i 个地理位置，（$\hat{x}$，$\hat{y}$）表示最优地理位置。

可选地，所述精确地理位置确定模块还用于：将所述最优地理位置输入所述预设的人工神经网络模型，获取输出结果；若所述输出结果为预设的目标结果，则所述最优地理位置为所述 IP 的精确地理位置。

可选地，所述预设的人工神经网络模型的输入层具有 3 个神经元节点，隐含层具有 5 个神经元节点，输出层具有 1 个神经元节点。

为实现上述目的，根据本发明实施例的一个方面，提供了一种电子设备，包括：一个或多个处理器；存储装置，用于存储一个或多个程序，当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现本发明实施例所述的确定精确地理位置的方法。

为实现上述目的，根据本发明实施例的一个方面，提供了一种计算机可读介质，其上存储有计算机程序，所述程序被处理器执行时实

现本发明实施例所述的确定精确地理位置的方法。

上述发明中的一个实施例具有如下优点或有益效果：因为采用聚类算法，对所述多个地理位置进行聚类以获得每个 IP 的地理位置聚类结果；基于所述地理位置聚类结果，利用优化算法确定所述 IP 对应的最优地理位置；根据所述最优地理位置和预设的人工神经网络模型，确定所述 IP 的精确地理位置的技术手段，所以提高了定位精度，而且不需要铺设大量监测点，降低了成本。具体的，通过 k-means 聚类，减少冗余数据，减少由于天气、信号、周边环境等因素造成的 GPS 定位误差；然后对不同用户（MAC）但是同一 IP 的地理位置，利用加权最小二乘法获取最优地理位置，；随着数据的积累，建立 ANN 神经网络训练模型，对于同一 IP 计算得到的最优解进行训练，排除一段时间内由于某些因素造成的脏数据（一些移动端装置可以模拟 GPS 数据，导致 GPS 数据无效），从而提高定位的准确性。

上述的非惯用的可选方式所具有的进一步效果将在下文中结合具体实施方式加以说明。

附图说明

附图用于更好地理解本发明，不构成对本发明的不当限定。其中：

图 1 是根据本发明一实施例的确定精确地理位置的方法的主要流程的示意图；

图 2 是根据本发明另一实施例的确定精确地理位置的方法的主要流程的示意图；

图 3 是根据本发明实施例的确定精确地理位置的装置的主要模块的示意图；

图 4 是本发明实施例可以应用于其中的示例性系统架构图；

图 5 是适于用来实现本发明实施例的终端设备或服务器的计算机系统的结构示意图。

具体实施方式

以下结合附图对本发明的示范性实施例做出说明，其中包括本发明实施例的各种细节以助于理解，应当将它们认为仅仅是示范性的。因此，本领域普通技术人员应当认识到，可以对这里描述的实施例做出各种改变和修改，而不会背离本发明的范围和精神。同样，为了清楚和简明，以下的描述中省略了对公知功能和结构的描述。

图 1 是根据本发明实施例的 IP-地理位置数据集确定精确地理位置的方法的主要流程图的示意图。如图 1 所示，该方法包括：

步骤 S101：获取 IP 以及与所述 IP 关联的多个地理位置；

步骤 S102：利用聚类算法，对所述多个地理位置进行聚类以获得所述 IP 的地理位置聚类结果；

步骤 S103：基于所述地理位置聚类结果，利用优化算法确定所述 IP 对应的最优地理位置；

步骤 S104：根据所述最优地理位置和预设的人工神经网络模型，确定所述 IP 的精确地理位置。

本实施例中的 IP 以及与 IP 关联的多个地理位置可以通过公开的地理信息数据库获取。也可以通过接收数据采集源上报的 IP 以及与 IP 关联的多个地理位置来获取，例如接收具有 GPS 信息模块的上报装置（例如智能手机、平板电脑等终端设备）上报的 IP 地址以及与该 IP 地址关联的地理位置。

随着移动互联网科技的发展，任何一部手机或者平板电脑等终端设备可以成本本实施例中的上报装置，都可以作为数据采集源，因此，本发明实施例不需要铺设大量的监测点，降低了成本。

在可选的实施例中，在接收上报装置上报的 IP 以及与该 IP 关联的地理位置时，还可以获取该上报装置的设备标识（例如 MAC 地址）和上报数据时的时间戳，从而将该设备标识、时间戳、IP 以及该 IP 的地

理位置组成一条有效数据，例如 IP-MAC-GPS-TIMESTAMP，其中，GPS 为上报的经纬度信息，TIMESTAMP 为上报数据时的时间戳。

上述地理位置可以表现为经纬度信息、海拔信息等卫星定位信息，也可以表现为城市、街道、商户、写字楼等位置信息。在本实施例中，该地理位置优选为经纬度信息。

上述 IP 本质上，是一个 32 位的无符号整型（unsigned int）数据，范围从 0~2^{32}，为了使用方便，一般使用字符串形式的 IP 地址，也就是平时使用的 192.168.0.1 这种形式，实际上，就是把每 8 个二进制位转换成对应的十进制整数，简称数值型 IP。比如，192.168.0.1 和 3232235521 是等价的。192.168.0.1 意指 $1*256^0 + 0*256^1 + 168*256^2 + 192*256^3 = 3232235521$。在本发明实施例中，为使用简便，所述的 IP 为数值型 IP。

由于 GPS 容易受到上报装置所处环境、天气以及上报装置本身的信号等因素的影响，所以样本集中某些地理位置的误差可能较大，并不能完全采信。

因此，对于步骤 S102，需要利用聚类算法对所述 IP 的多个地理位置进行聚类来排除误差较大的地理位置，从而获得该 IP 对应的较为准确的地理位置，进而提高定位的准确性。作为具体的示例，该聚类算法可以是 k-means 聚类算法，进一步的，可以将设备标识作为维度，以时间戳进行聚类，即对同一上报装置在某一段时间内上报的数据进行聚类。

上述 k-means 算法是很典型的基于距离的聚类算法，采用距离作为相似性的评价指标，即认为两个对象的距离越近，其相似度就越大。该算法的核心在于通过解算数据点到质心的某种距离作为优化目标的函数，利用函数取极值不断迭代，因此把得到紧凑且独立的簇作为最

终目标。

进一步的，如图 2 所示，利用 k-means 聚类算法，对所述 IP 关联的多个地理位置进行聚类以获得所述 IP 的地理位置聚类结果的步骤包括如下步骤：

步骤 S201：从所述 IP 关联的多个地理位置中选取两个地理位置作为第一初始质心和第二初始质心；

步骤 S202：计算所述多个地理位置中每个地理位置与所述第一初始质心之间的第一球面距离以及与第二初始质心之间的第二球面距离；

步骤 S203：根据所述第一球面距离和所述第二球面距离，对所述 IP 关联的地理位置进行聚类以获得高密度簇和低密度簇，以所述高密度簇作为所述 IP 的地理位置聚类结果。

对于步骤 S201，对于同一上报装置（即同一 IP）在一段时间内采集的经纬度数据在分布上都是散列在该 IP 真实的地理位置附近的，此类点密度较大，但是由于受到外界因素影响，少数点与真实位置偏差较大，密度稀疏。因此，本发明实施例定义簇为被低密度区域分开的高密度区域，在选取初始化质心时，在基于密度的簇上选用两类。

具体的，可以随机选取 2 个经纬度作为第一初始质心和第二初始质心，也可以选取所有经纬度的平均值作为第一初始质心、与平均值偏差最大的经纬度作为第二初始质心。

对于步骤 S202，由于经纬度是椭球面的坐标，因此不能简单的使用欧式距离作为衡量簇的紧凑指标，在此本发明实施例使用球面距离作为衡量簇的紧凑指标。可以通过如下公式计算两个地理位置之间的球面距离：

$$S = R \bullet ar\cos(\cos\beta 1 \bullet \cos\beta 2 \bullet \cos(\alpha 1 - \alpha 2) + \sin\beta 1 \bullet \sin\beta 2) \quad (1)$$

其中，R 表示地球长轴半径，S 表示地理位置 A 与地理位置 B 之

间的球面距离，$\beta 1$为地理位置 A 的纬度，$\alpha 1$为地理位置 A 的经度，$\beta 2$为地理位置 B 的纬度，$\alpha 2$为地理位置 B 的经度。

对于步骤 S203，根据公式（1）计算得出第一球面距离和第二球面距离之后，距离第一初始质心近的地理位置为一簇，距离第二初始质心近的地理位置为另一簇。然后，重新计算每一簇的质心，重复迭代，直至最终质心不变或变化很小。选取高密度簇作为该 IP 的地理位置聚类结果，低密度簇作为误差簇进行摒弃，避免造成数据污染。

对于步骤 S103，经过聚类算法初步排除了误差较大的地理位置之后，为了进一步提高定位精度，需要利用优化算法确定每一个 IP 对应的最优地理位置。具体的，可以利用优化算法对同一 IP 的高密度簇求取最优解。作为具体的示例，该优化算法可以是加权最小二乘法。

上述加权最小二乘法是一种数学优化技术，它通过最小化误差的平方和寻找数据的最佳函数匹配。加权最小二乘法在工程技术领域有着广泛的应用，利用加权最小二乘法可以简便地求得未知的参数，并使得这些求得数据与实际数据之间误差的平方和最小。

具体的，基于地理位置聚类结果，利用加权最小二乘法确定所述 IP 对应的最优地理位置的过程可以包括如下步骤：

1. 对于高密度簇中的每个地理位置，根据每个地理位置与高密度簇质心的球面距离，确定所述每个地理位置的权重；

公式如下：

$$\lambda_i = \frac{\frac{1}{d_i}}{\sum_{i=1}^{n}\frac{1}{d_i}} \quad (2)$$

λ_i表示第 i 个经纬度的权重，d_i表示第 i 个经纬度与质心之间的距离，n 为大于或等于 1 的整数。

2. 根据所述权重，利用加权最小二乘法确定每个 IP 对应的最优地理位置。在此过程中，需要对同一 IP 的经纬度建立非线性曲线拟合函数，使其方差最小，具体公式如下式（3）：

$$\text{Min} = \sum_{i=1}^{n} \lambda_i \bullet ((x_i - \hat{x})^2 + (y_i - \hat{y})^2) \quad (3)$$

其中，（x_i, y_i）表示第 i 个地理位置，（$\hat{x}$，$\hat{y}$）为该 IP 对应的最优地理位置。在实际计算时，（x_i, y_i）为第 i 个地理位置通过高斯投影将经纬度转换为大地坐标后的平面坐标。

在本发明实施例中，对同一 IP 的经纬度数据建立非线性回归模型：$(x-\hat{x})^2+(y-\hat{y})^2=r^2$，其中（$\hat{x}$，$\hat{y}$）为圆心坐标，r 为半径。求该 IP 对应的最优地理位置即求解（$\hat{x}$，$\hat{y}$）使其满足 $\sum_{i=1}^{n} \lambda_i \bullet ((x_i - \hat{x})^2 + (y_i - \hat{y})^2)$ 最小。

对于步骤 S103，经过上述的 k-means 算法以及加权最小二乘法，可以认定某一台采样装置上报数据已经得到正确处理，但是在实际过程中，由于存在模拟器等因素，上报的 IP 与经纬度数据可能存在较大偏差，这部分数据可以认为是异常数据。因此，在本实施例中可以利用人工神经网络模型，对于同一 IP 计算得到的最优地理位置进行筛选，从而排除异常数据。具体的，可以在确定该 IP 的最优地理位置之后，引入人工神经网络模型对该最优地理位置进行一个简单的‘分类’，即将所有的最优地理位置分为两类，正常和异常两类。

因此，进一步的，该方法还包括：根据所述最优地理位置和预设的人工神经网络模型，确定所述 IP 的精确地理位置。

具体的，可以包括如下步骤：

将所述最优地理位置输入所述预设的人工神经网络模型，获取输出结果；

若所述输出结果为预设的目标结果，则所述最优地理位置为所述IP的精确地理位置。

在使用预设的人工神经网络模型对最优地理位置进行筛选之前，该方法还包括：训练该人工神经网络模型，即通过训练数据调整各个神经节点的权重，使得正常最优地理位置的期望输出为1，异常最优地理位置的期望输出为0。

具体的，选取大量关联正确地理位置的IP数据作为正常数据（例如大于20000条数据），并对同一IP加入人工异常数据，利用该正常数据和人工异常数据进行人工神经网络模型隐藏层权重训练，保证最终函数收敛，此时的隐藏层权重参数作为初始化参数。

作为具体的示例，所述预设的人工神经网络模型的输入层具有3个神经元节点，分别对应IP（数值型IP）、经度和纬度；隐含层具有5个神经元节点，该节点数量由开发人员通过训练数据收敛时间以及方法确定；输出层具有1个神经元节点，通过输出结果判定该经纬度是否是异常数据，输出结果为1表示该经纬度为正常数据，输出结果为0表示该经纬度为异常数据。

因此，上述的预设的目标结果可以为1，若输出结果为1，则该最优地理位置为所述IP的精确地理位置。

在可选的实施例中，可以将获得的IP及该IP的精确地理位置保存。

在本实施例中，人工神经网络模型（Artificial Neural Network，ANN）为：从信息处理角度对人脑神经元网络进行抽象，建立某种简单模型，按不同的连接方式组成不同的网络。在工程与学术界也常直接简称为神经网络或类神经网络。神经网络是一种运算模型，由大量的节点（或称神经元）之间相互联接构成。每个节点代表一种特定的

输出函数，称为激励函数（activation function）。每两个节点间的连接都代表一个对于通过该连接信号的加权值，称之为权重，这相当于人工神经网络的记忆。网络的输出则依网络的连接方式，权重值和激励函数的不同而不同。而网络自身通常都是对自然界某种算法或者函数的逼近，也可能是对一种逻辑策略的表达。

本发明实施例的确定精确地理位置的方法提高了定位精度，而且不需要铺设大量监测点，降低了成本。具体的，通过 k-means 聚类，减少冗余数据，减少由于天气、信号、周边环境等因素造成的 GPS 定位误差；然后对不同用户（MAC）但是同一 IP 的地理位置，利用加权最小二乘法获取最优地理位置，；随着数据的积累，建立 ANN 神经网络训练模型，对于同一 IP 计算得到的最优解进行训练，排除一段时间内由于某些因素造成的脏数据（一些移动端装置可以模拟 GPS 数据，导致 GPS 数据无效），从而提高定位的准确性。

本发明实施例的方法，还可以根据公式（3）得到方差，为 IP 定位的准确性提供量化指标，方差越小，准确性越高。

图 3 是根据本发明又一实施例的 IP 定位装置的主要模块的示意图。如图 3 所示，该装置 300 包括：获取模块 301，用于获取 IP 以及与所述 IP 关联的多个地理位置；聚类模块 302，用于利用聚类算法，对所述多个地理位置进行聚类以获得所述 IP 的地理位置聚类结果；最优地理位置确定模块 303，用于基于所述地理位置聚类结果，利用优化算法确定所述 IP 对应的最优地理位置；精确地理位置确定模块 304，用于根据所述最优地理位置和预设的人工神经网络模型，确定所述 IP 的精确地理位置。

可选地，所述聚类算法为 k-means 算法，所述优化算法为加权最小二乘法。

可选地，所述聚类模块 302 还用于：从所述 IP 关联的多个地理位置中选取两个地理位置作为第一初始质心和第二初始质心；计算所述多个地理位置中每个地理位置与所述第一初始质心之间的第一球面距离以及与第二初始质心之间的第二球面距离；根据所述第一球面距离和所述第二球面距离，对所述 IP 关联的多个地理位置进行聚类以获得高密度簇，以所述高密度簇作为所述 IP 的地理位置聚类结果。

可选地，所述聚类模块 302 根据如下公式（1）计算每个地理位置与所述第一初始质心之间的第一球面距离以及与第二初始质心之间的第二球面距离：

$$S = R \bullet ar\cos(\cos\beta 1 \bullet \cos\beta 2 \bullet \cos(\alpha 1 - \alpha 2) + \sin\beta 1 \bullet \sin\beta 2) \quad (1)$$

其中，R 表示地球长轴半径，S 表示地理位置 A 与地理位置 B 之间的球面距离，$\beta 1$为地理位置 A 的纬度，$\alpha 1$为地理位置 A 的经度，$\beta 2$为地理位置 B 的纬度，$\alpha 2$为地理位置 B 的经度。

可选地，所述最优地理位置确定模块 303 还用于：对于高密度簇中的每个地理位置，根据每个地理位置与高密度簇质心的球面距离，确定所述每个地理位置的权重；根据所述权重，利用加权最小二乘法确定每个 IP 对应的最优地理位置。

可选地，根据下式（2）确定所述每个地理位置的权重：

$$\lambda_i = \frac{\frac{1}{d_i}}{\sum_{i=1}^{n}\frac{1}{d_i}} \quad (2)$$

其中，λ_i表示第 i 个地理位置的权重，d_i表示第 i 个地理位置与高密度簇质心之间的球面距离，n 为大于或等于 1 的整数；

根据下式（3）确定所述 IP 对应的最优地理位置：

$$Min = \sum_{i=1}^{n}\lambda_i \bullet ((x_i - \hat{x})^2 + (y_i - \hat{y})^2) \quad (3)$$

其中，(x_i, y_i)表示第 i 个地理位置，$(\hat{x}, \hat{y})$表示最优地理位

置。

可选地，所述精确地理位置确定模块 304 还用于：将所述最优地理位置输入所述预设的人工神经网络模型，获取输出结果；若所述输出结果为预设的目标结果，则所述最优地理位置为所述 IP 的精确地理位置。

可选地，所述预设的人工神经网络模型的输入层具有 3 个神经元节点，隐含层具有 5 个神经元节点，输出层具有 1 个神经元节点。

本发明实施例的确定精确地理位置的装置提高了定位精度，而且不需要铺设大量监测点，降低了成本。具体的，通过 k-means 聚类，减少冗余数据，减少由于天气、信号、周边环境等因素造成的 GPS 定位误差；然后对不同用户（MAC）但是同一 IP 的地理位置，利用加权最小二乘法获取最优地理位置，；随着数据的积累，建立 ANN 神经网络训练模型，对于同一 IP 计算得到的最优解进行训练，排除一段时间内由于某些因素造成的脏数据（一些移动端装置可以模拟 GPS 数据，导致 GPS 数据无效），从而提高定位的准确性。

图 4 示出了可以应用本发明实施例的 IP-地理位置数据集构建方法或 IP-地理位置数据集构建装置的示例性系统架构 400。

如图 4 所示，系统架构 400 可以包括终端设备 401、402、403，网络 404 和服务器 405。网络 404 用以在终端设备 401、402、403 和服务器 405 之间提供通信链路的介质。网络 404 可以包括各种连接类型，例如有线、无线通信链路或者光纤电缆等等。

用户可以使用终端设备 401、402、403 通过网络 404 与服务器 405 交互，以接收或发送消息等。终端设备 401、402、403 可以是具有显示屏并且支持网页浏览的各种电子设备，包括但不限于智能手机、平

板电脑、膝上型便携计算机和台式计算机等等。

服务器 405 可以是提供各种服务的服务器，例如对用户利用终端设备 401、402、403 所浏览的购物类网站提供支持的后台管理服务器。后台管理服务器可以对接收到的产品信息查询请求等数据进行分析等处理，并将处理结果（例如目标推送信息、产品信息）反馈给终端设备。

需要说明的是，本发明实施例所提供的确定精确地理位置的方法一般由服务器 405 执行，相应地，IP 定位装置一般设置于服务器 405 中。

应该理解，图 4 中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要，可以具有任意数目的终端设备、网络和服务器。

下面参考图 5，其示出了适于用来实现本发明实施例的终端设备的计算机系统 500 的结构示意图。图 5 示出的终端设备仅仅是一个示例，不应对本发明实施例的功能和使用范围带来任何限制。

如图 5 所示，计算机系统 500 包括中央处理单元（CPU）501，其可以根据存储在只读存储器（ROM）502 中的程序或者从存储部分 508 加载到随机访问存储器（RAM）503 中的程序而执行各种适当的动作和处理。在 RAM 503 中，还存储有系统 500 操作所需的各种程序和数据。CPU 501、ROM 502 以及 RAM 503 通过总线 504 彼此相连。输入/输出（I/O）接口 505 也连接至总线 504。

以下部件连接至 I/O 接口 505：包括键盘、鼠标等的输入部分 506；包括诸如阴极射线管（CRT）、液晶显示器（LCD）等以及扬声器等的输出部分 507；包括硬盘等的存储部分 508；以及包括诸如 LAN 卡、调制解调器等的网络接口卡的通信部分 509。通信部分 509 经由诸如因

特网的网络执行通信处理。驱动器 510 也根据需要连接至 I/O 接口 505。可拆卸介质 511，诸如磁盘、光盘、磁光盘、半导体存储器等等，根据需要安装在驱动器 510 上，以便于从其上读出的计算机程序根据需要被安装入存储部分 508。

特别地，根据本发明公开的实施例，上文参考流程图描述的过程可以被实现为计算机软件程序。例如，本发明公开的实施例包括一种计算机程序产品，其包括承载在计算机可读介质上的计算机程序，该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中，该计算机程序可以通过通信部分 509 从网络上被下载和安装，和/或从可拆卸介质 511 被安装。在该计算机程序被中央处理单元（CPU）501 执行时，执行本发明的系统中限定的上述功能。

需要说明的是，本发明所示的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、光纤、便携式紧凑磁盘只读存储器（CD-ROM）、光存储器件、磁存储器件、或者上述的任意合适的组合。在本发明中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本发明中，计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机

可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：无线、电线、光缆、RF 等等，或者上述的任意合适的组合。

附图中的流程图和框图，图示了按照本发明各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，上述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的是，框图或流程图中的每个方框、以及框图或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

描述于本发明实施例中所涉及到的模块可以通过软件的方式实现，也可以通过硬件的方式来实现。所描述的 模块也可以设置在处理器中，例如，可以描述为：一种处理器包括发送模块、获取模块、确定模块和第一处理模块。其中，这些模块的名称在某种情况下并不构成对该单元本身的限定，例如，发送模块还可以被描述为“向所连接的服务端发送图片获取请求的模块”。

作为另一方面，本发明还提供了一种计算机可读介质，该计算机可读介质可以是上述实施例中描述的设备中所包含的；也可以是单独存在，而未装配入该设备中。上述计算机可读介质承载有一个或者多个程序，当上述一个或者多个程序被一个该设备执行时，使得该设备包括：获取 IP 以及与所述 IP 关联的多个地理位置；利用聚类算法，对所述多个地理位置进行聚类以获得所述 IP 的地理位置聚类结果；基于所述地理位置聚类结果，利用优化算法确定所述 IP 对应的最优地理位置；根据所述最优地理位置和预设的人工神经网络模型，确定所述 IP

的精确地理位置。

本发明实施例的技术方案因为采用聚类算法，对所述多个地理位置进行聚类以获得每个 IP 的地理位置聚类结果；基于所述地理位置聚类结果，利用优化算法确定所述 IP 对应的最优地理位置；根据所述最优地理位置和预设的人工神经网络模型，确定所述 IP 的精确地理位置的技术手段，所以提高了定位精度，而且不需要铺设大量监测点，降低了成本。具体的，通过 k-means 聚类，减少冗余数据，减少由于天气、信号、周边环境等因素造成的 GPS 定位误差；然后对不同用户（MAC）但是同一 IP 的地理位置，利用加权最小二乘法获取最优地理位置，；随着数据的积累，建立 ANN 神经网络训练模型，对于同一 IP 计算得到的最优解进行训练，排除一段时间内由于某些因素造成的脏数据（一些移动端装置可以模拟 GPS 数据，导致 GPS 数据无效），从而提高定位的准确性。

上述具体实施方式，并不构成对本发明保护范围的限制。本领域技术人员应该明白的是，取决于设计要求和其他因素，可以发生各种各样的修改、组合、子组合和替代。任何在本发明的精神和原则之内所作的修改、等同替换和改进等，均应包含在本发明保护范围之内。

1. 一种确定精确地理位置的方法，其特征在于，包括：

获取 IP 以及与所述 IP 关联的多个地理位置；

利用聚类算法，对所述多个地理位置进行聚类以获得所述 IP 的地理位置聚类结果；

基于所述地理位置聚类结果，利用优化算法确定所述 IP 对应的最优地理位置；

根据所述最优地理位置和预设的人工神经网络模型，确定所述 IP 的精确地理位置。

2. 根据权利要求 1 所述的方法，其特征在于，所述聚类算法为 k-means 算法，所述优化算法为加权最小二乘法。

3. 根据权利要求 2 所述的方法，其特征在于，利用聚类算法，对所述多个地理位置进行聚类以获得所述 IP 的地理位置聚类结果的步骤包括：

从所述 IP 关联的多个地理位置中选取两个地理位置作为第一初始质心和第二初始质心；

计算所述多个地理位置中每个地理位置与所述第一初始质心之间的第一球面距离以及与第二初始质心之间的第二球面距离；

根据所述第一球面距离和所述第二球面距离，对所述 IP 关联的多个地理位置进行聚类以获得高密度簇，以所述高密度簇作为所述 IP 的地理位置聚类结果。

4. 根据权利要求 3 所述的方法，其特征在于，根据如下公式（1）计算每个地理位置与所述第一初始质心之间的第一球面距离以及与第二初始质心之间的第二球面距离：

$$S = R \bullet arcos(\cos\beta 1 \bullet \cos\beta 2 \bullet \cos(\alpha 1 - \alpha 2) + \sin\beta 1 \bullet \sin\beta 2) \quad (1)$$

其中，R 表示地球长轴半径，S 表示地理位置 A 与地理位置 B 之

间的球面距离，$\beta 1$为地理位置 A 的纬度，$\alpha 1$为地理位置 A 的经度，$\beta 2$为地理位置 B 的纬度，$\alpha 2$为地理位置 B 的经度。

5. 根据权利要求 3 所述的方法，其特征在于，根据所述地理位置聚类结果，利用优化算法确定每个 IP 对应的最优地理位置包括：

对于高密度簇中的每个地理位置，根据每个地理位置与高密度簇质心的球面距离，确定所述每个地理位置的权重；

根据所述权重，利用加权最小二乘法确定每个 IP 对应的最优地理位置。

6. 根据权利要求 5 所述的方法，其特征在于，根据下式（2）确定所述每个地理位置的权重：

$$\lambda_i = \frac{\frac{1}{d_i}}{\sum_{i=1}^{n}\frac{1}{d_i}} \qquad (2)$$

其中，λ_i表示第 i 个地理位置的权重，d_i表示第 i 个地理位置与高密度簇质心之间的球面距离，n 为大于或等于 1 的整数；

根据下式（3）确定所述 IP 对应的最优地理位置：

$$\mathrm{Min} = \sum_{i=1}^{n} \lambda_i \bullet ((x_i - \hat{x})^2 + (y_i - \hat{y})^2) \qquad (3)$$

其中，(x_i, y_i)表示第 i 个地理位置，$(\hat{x}, \hat{y})$表示最优地理位置。

7. 根据权利要求 1 所述的方法，其特征在于，根据所述最优地理位置和预设的人工神经网络模型，确定所述 IP 的精确地理位置包括：

将所述最优地理位置输入所述预设的人工神经网络模型，获取输出结果；

若所述输出结果为预设的目标结果，则所述最优地理位置为所述 IP 的精确地理位置。

8. 根据权利要求 7 所述的方法，其特征在于，所述预设的人工神经网络模型的输入层具有 3 个神经元节点，隐含层具有 5 个神经元节点，输出层具有 1 个神经元节点。

9. 一种确定精确地理位置的装置，其特征在于，包括：

获取模块，用于获取 IP 以及与所述 IP 关联的多个地理位置；

聚类模块，用于利用聚类算法，对所述多个地理位置进行聚类以获得所述 IP 的地理位置聚类结果；

最优地理位置确定模块，用于基于所述地理位置聚类结果，利用优化算法确定所述 IP 对应的最优地理位置；

精确地理位置确定模块，用于根据所述最优地理位置和预设的人工神经网络模型，确定所述 IP 的精确地理位置。

10. 根据权利要求 9 所述的装置，其特征在于，所述聚类算法为 k-means 算法，所述优化算法为加权最小二乘法。

11. 根据权利要求 10 所述的装置，其特征在于，所述聚类模块还用于：

从所述 IP 关联的多个地理位置中选取两个地理位置作为第一初始质心和第二初始质心；

计算所述多个地理位置中每个地理位置与所述第一初始质心之间的第一球面距离以及与第二初始质心之间的第二球面距离；

根据所述第一球面距离和所述第二球面距离，对所述 IP 关联的多个地理位置进行聚类以获得高密度簇，以所述高密度簇作为所述 IP 的地理位置聚类结果。

12. 根据权利要求 11 所述的装置，其特征在于，所述聚类模块根据如下公式（1）计算每个地理位置与所述第一初始质心之间的第一球面距离以及与第二初始质心之间的第二球面距离：

$$S = R \bullet ar\cos(\cos\beta 1 \bullet \cos\beta 2 \bullet \cos(\alpha 1 - \alpha 2) + \sin\beta 1 \bullet \sin\beta 2) \quad (1)$$

其中，R 表示地球长轴半径，S 表示地理位置 A 与地理位置 B 之间的球面距离，$\beta 1$为地理位置 A 的纬度，$\alpha 1$为地理位置 A 的经度，$\beta 2$为地理位置 B 的纬度，$\alpha 2$为地理位置 B 的经度。

13. 根据权利要求 10 所述的装置，其特征在于，所述最优地理位置确定模块还用于：

对于高密度簇中的每个地理位置，根据每个地理位置与高密度簇质心的球面距离，确定所述每个地理位置的权重；

根据所述权重，利用加权最小二乘法确定每个 IP 对应的最优地理位置。

14. 根据权利要求 13 所述的装置，其特征在于，根据下式（2）确定所述每个地理位置的权重：

$$\lambda_i = \frac{\frac{1}{d_i}}{\sum_{i=1}^{n}\frac{1}{d_i}} \qquad (2)$$

其中，λ_i表示第 i 个地理位置的权重，d_i表示第 i 个地理位置与高密度簇质心之间的球面距离，n 为大于或等于 1 的整数；

根据下式（3）确定所述 IP 对应的最优地理位置：

$$\mathrm{Min} = \sum_{i=1}^{n} \lambda_i \bullet ((x_i - \hat{x})^2 + (y_i - \hat{y})^2) \qquad (3)$$

其中，（x_i, y_i）表示第 i 个地理位置，（$\hat{x}$，$\hat{y}$）表示最优地理位置。

15. 根据权利要求 8 所述的装置，其特征在于，所述精确地理位置确定模块还用于：

将所述最优地理位置输入所述预设的人工神经网络模型，获取输出结果；

若所述输出结果为预设的目标结果，则所述最优地理位置为所述 IP 的精确地理位置。

16. 根据权利要求 15 所述的装置，其特征在于，所述预设的人工神经网络模型的输入层具有 3 个神经元节点，隐含层具有 5 个神经元节点，输出层具有 1 个神经元节点。

17. 一种电子设备，其特征在于，包括：

一个或多个处理器；

存储装置，用于存储一个或多个程序，

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如权利要求 1-8 中任一所述的方法。

18. 一种计算机可读介质，其上存储有计算机程序，其特征在于，所述程序被处理器执行时实现如权利要求 1-8 中任一所述的方法。

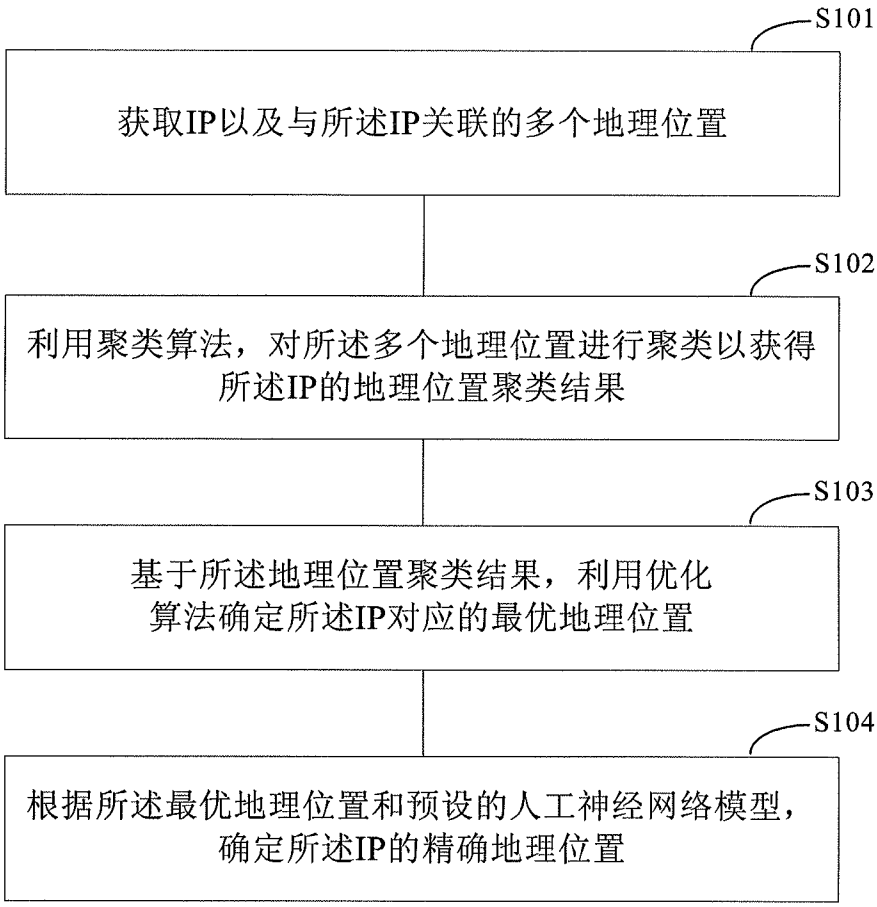
S101
获取IP以及与所述IP关联的多个地理位置
S102
利用聚类算法，对所述多个地理位置进行聚类以获得所述IP的地理位置聚类结果
S103
基于所述地理位置聚类结果，利用优化算法确定所述IP对应的最优地理位置
S104
根据所述最优地理位置和预设的人工神经网络模型，确定所述IP的精确地理位置

图 1

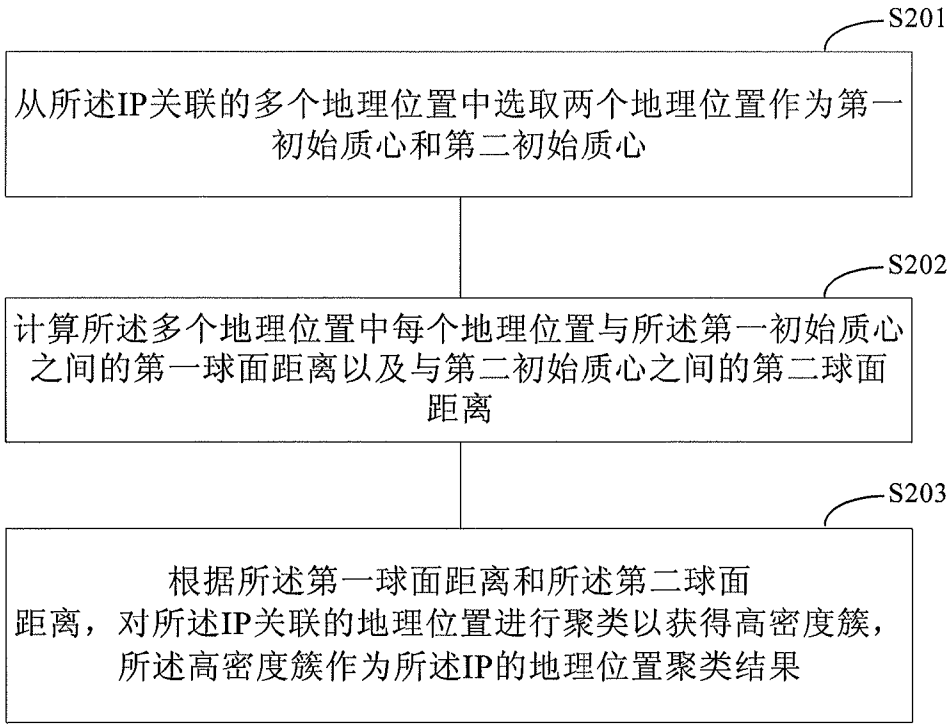
S201
从所述IP关联的多个地理位置中选取两个地理位置作为第一初始质心和第二初始质心
S202
计算所述多个地理位置中每个地理位置与所述第一初始质心之间的第一球面距离以及与第二初始质心之间的第二球面距离
S203
根据所述第一球面距离和所述第二球面距离，对所述IP关联的地理位置进行聚类以获得高密度簇，所述高密度簇作为所述IP的地理位置聚类结果

图 2

5

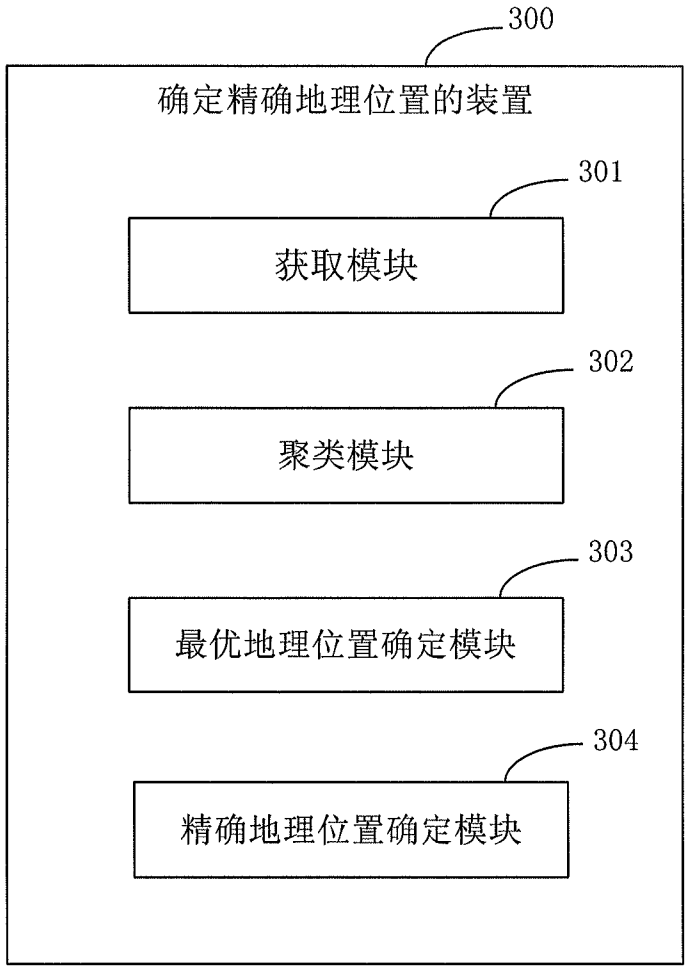
300
确定精确地理位置的装置
301
获取模块
302
聚类模块
303
最优地理位置确定模块
304
精确地理位置确定模块

图 3

5

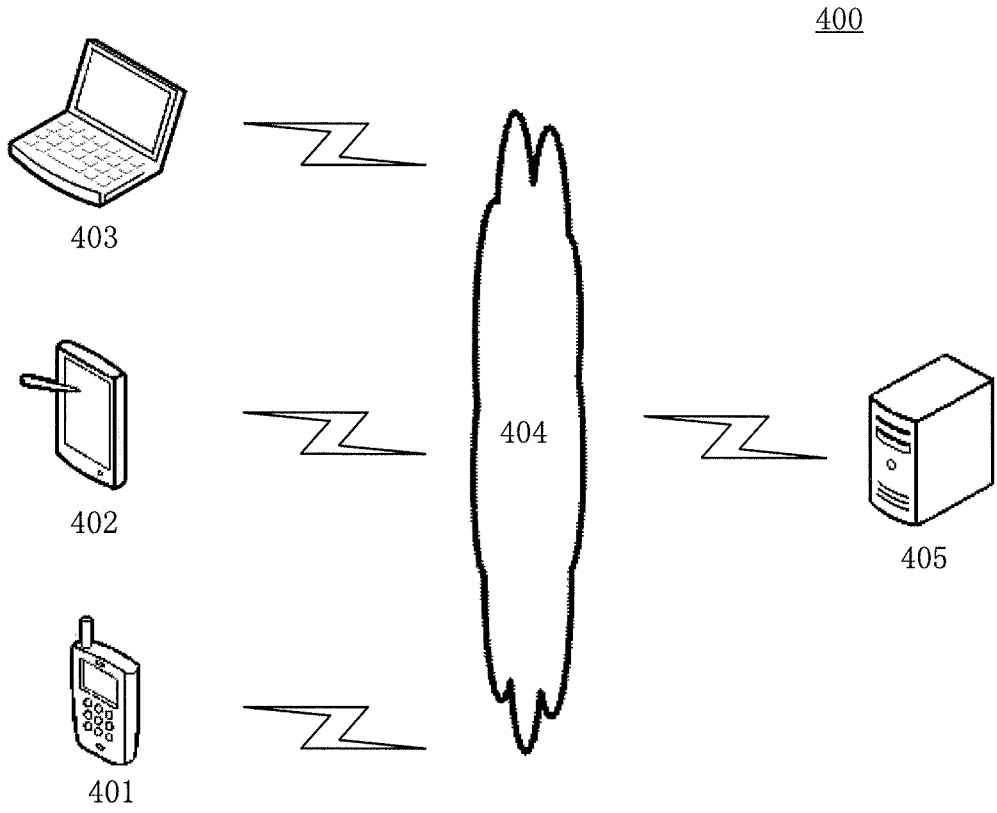
400
403
402
401
404
405

图 4

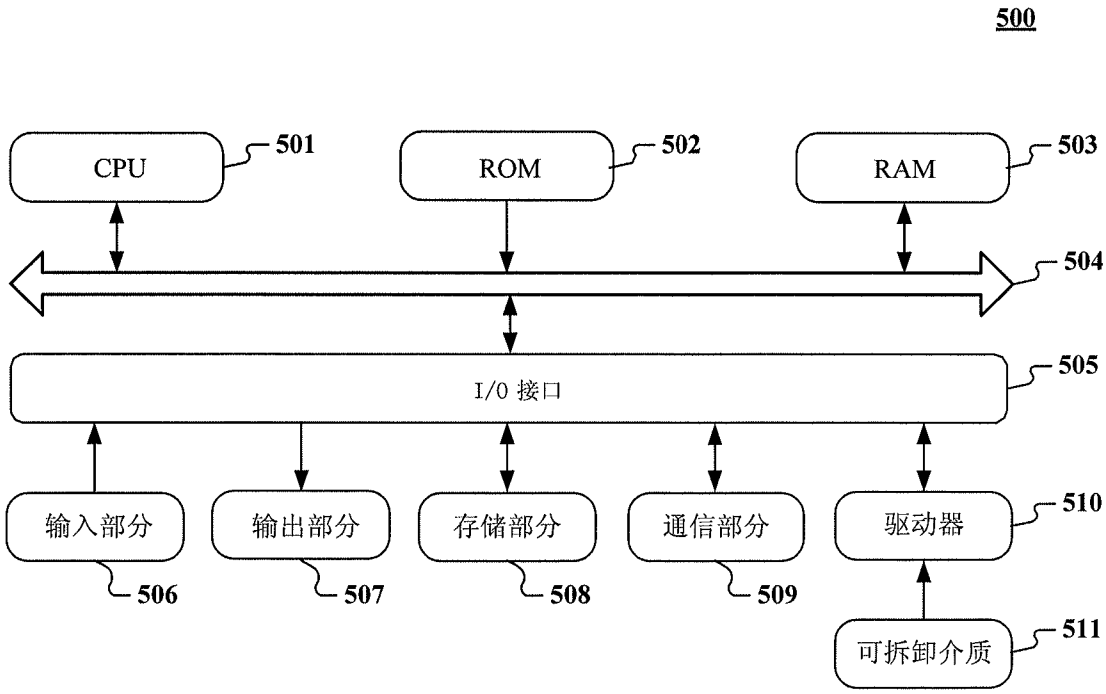
500
CPU
501
ROM
502
RAM
503
504
I/O 接口
505
输入部分
506
输出部分
507
存储部分
508
通信部分
509
驱动器
510
可拆卸介质
511

图 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/108635

A. CLASSIFICATION OF SUBJECT MATTER

H04L 29/06(2006.01)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)

H04L

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: 定位, 地理, 地址, 位置, 精确, IP, 聚类, 最优, 神经网络, 模型, locate, geography, address, position, accuracy, clustering, best, neural, network, model

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 105933294 A (ZAMPLUS ADVERTISING (SHANGHAI) CO., LTD.) 07 September 2016 (2016-09-07) description, paragraphs [0037]-[0062], and claims 1-19	1-18
Y	CN 106469205 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 01 March 2017 (2017-03-01) description, paragraphs [0042]-[0100], and claims 1-11	1-18
A	CN 105718465 A (ALIBABA GROUP HOLDING LIMITED) 29 June 2016 (2016-06-29) entire document	1-18
A	CN 101267374 A (TSINGHUA UNIVERSITY) 17 September 2008 (2008-09-17) entire document	1-18
A	US 7543045 B1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 02 June 2009 (2009-06-02) entire 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
05 December 2018	**07 January 2019**

Name and mailing address of the ISA/CN	Authorized officer
National Intellectual Property Administration, PRC (ISA/CN) **No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088** **China**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2018/108635

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 105933294 A	07 September 2016	None	
CN 106469205 A	01 March 2017	None	
CN 105718465 A	29 June 2016	WO 2016086786 A1	09 June 2016
CN 101267374 A	17 September 2008	None	
US 7543045 B1	02 June 2009	None	

国际检索报告

国际申请号
PCT/CN2018/108635

A. 主题的分类

H04L 29/06(2006.01)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)

H04L

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，地理，地址，位置，精确，IP，聚类，最优，神经网络，模型，locate，geography，address，position，accuracy，clustering，best，neural，network，model

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
Y	CN 105933294 A (晶赞广告上海有限公司) 2016年 9月 7日 (2016-09-07) 说明书第[0037]-[0062]段，权利要求1-19	1-18
Y	CN 106469205 A (百度在线网络技术北京有限公司) 2017年 3月 1日 (2017-03-01) 说明书第[0042]-[0100]段，权利要求1-11	1-18
A	CN 105718465 A (阿里巴巴集团控股有限公司) 2016年 6月 29日 (2016-06-29) 全文	1-18
A	CN 101267374 A (清华大学) 2008年 9月 17日 (2008-09-17) 全文	1-18
A	US 7543045 B1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2009年 6月 2日 (2009-06-02) 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
2018年 12月 5日	2019年 1月 7日
ISA/CN的名称和邮寄地址 中国国家知识产权局(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 王伦杰 电话号码 (86-10)53961776

表 PCT/ISA/210 (第2页) (2015年1月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2018/108635

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）
CN 105933294 A	2016年 9月 7日	无	
CN 106469205 A	2017年 3月 1日	无	
CN 105718465 A	2016年 6月 29日	WO 2016086786 A1	2016年 6月 9日
CN 101267374 A	2008年 9月 17日	无	
US 7543045 B1	2009年 6月 2日	无	