(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 September 2011 (29.09.2011)

PCT

(10) International Publication Number
WO 2011/119575 A1

(51) International Patent Classification:
H04W 24/00 (2009.01)

(21) International Application Number:
PCT/US2011/029379

(22) International Filing Date:
22 March 2011 (22.03.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

61/316,980	24 March 2010 (24.03.2010)	US
12/760,777	15 April 2010 (15.04.2010)	US
12/760,780	15 April 2010 (15.04.2010)	US

(71) Applicant *(for all designated States except US)*: **SKYHOOK WIRELESS, INC.** [US/US]; 34 Farnsworth Street, 4th Floor, Boston, MA 02210 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **ALIZADEH-SHABDIZ, Farshid** [US/US]; 5 Cole Road, Wayland, MA 01778 (US). **IGNATYEV, Oleksiy** [UA/US]; 87 Gordon Street, Apt. 11, Brighton, MA 02135 (US).

(74) Agents: **DICHIARA, Peter, M.** et al.; Wilmer Cutler Pickering Hale And Dorr LLP, 60 State Street, Boston, MA 02109 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, 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) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (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, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

(54) **Title**: SYSTEM AND METHOD FOR ESTIMATING THE PROBABILITY OF MOVEMENT OF ACCESS POINTS IN A WLAN-BASED POSITIONING SYSTEM

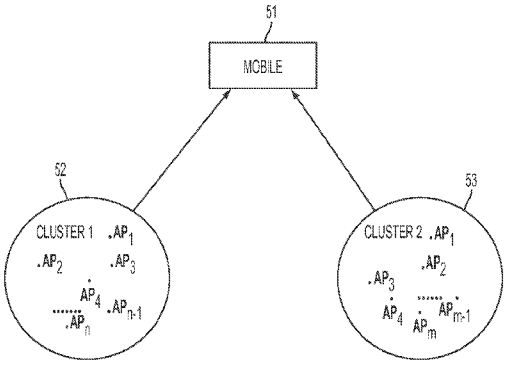

FIG. 6

(57) **Abstract**: Methods of and systems for estimating the probability of movement of access points and/or resolving multiple location estimate conflicts in a WLAN positioning system are provided. Disclosed are methods to quantify the probability that a particular location estimate of a mobile device made by a Wi-Fi based positioning system is correct to within an arbitrary accuracy. Implementations use observed access point cluster size, age information for access point location determination, the probability that one or more access points detected by the mobile device have relocated based on historic information about the movement of a collection of access points, and/or cumulative distribution functions that characterize the conditional probability that one or more access points detected by the mobile device have relocated within a specified time interval to make the probability determinations.

WO 2011/119575 A1

System and Method for Estimating the Probability of Movement of Access Points in a WLAN-based Positioning System

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit under 35 U.S.C. §119(e) of U.S. Provisional Patent Application No. 61/316,980, filed on March 24, 2010, entitled *System And Method For Resolving Multiple Location Estimate Conflicts In A WLAN-Positioning System*, which is herein incorporated by reference in its entirety.

[0002] This application also claims the benefit under 35 U.S.C. §120 of U.S. Patent Application No. 12/760,777, filed on April 15, 2010, entitled *System and Method for Resolving Multiple Location Estimate Conflicts in a WLAN-Positioning System* and U.S. Patent Application No. 12/760,780, filed on April 15, 2010, entitled *System and Method for Estimating the Probability of Movement of Access Points in a WLAN-Based Positioning System*, which are herein incorporated by reference in their entirety.

[0003] This application is related to the following U.S. Patent Applications, the contents of which are hereby incorporated by reference:

> U.S. Patent Application No. 11/261,988, entitled *Location-Based Services That Choose Location Algorithms Based On Number Of Detected Access Points Within Range Of User Device*, filed on October 28, 2005, now U.S. Patent No. 7,305,245;
>
> U.S. Patent Application No. 11/359,144, entitled *Continuous Data Optimization Of New Access Points In Positioning Systems*, filed on February 22, 2006, now U.S. Patent No. 7,493,127;
>
> U.S. Patent Application No. 11/678,301, entitled *Methods And Systems For Estimating A User Position In A WLAN Positioning System Based On User Assigned Access Point Locations*, filed on February 23, 2007, now U.S. Patent No. 7,471,954; and
>
> U.S. Patent Application No. 11/625,450, entitled *System and Method For Estimating Positioning Error Within A WLAN-Based Positioning System*, filed on January 22, 2007.

BACKGROUND OF THE INVENTION

Field of the Invention

[0004] The invention generally relates to position estimates in a WLAN-based positioning system, and, more specifically, to estimating the probability that a position estimate is correct

and managing and resolving conflicting position estimates of a mobile device to improve position estimate performance.

Description of Related Art

[0005] The U.S. Patents and Applications incorporated above and assigned to Skyhook Wireless, Inc. describe a Wi-Fi Positioning System (WPS) that uses the natural properties and widespread deployment of 802.11 access points (APs herein) to deliver precise positioning data to any Wi-Fi enabled device.

[0006] In such a WPS, APs provide a valuable method for determining the location of mobile wireless devices. Accurate knowledge of AP locations is essential to mobile location determination, and the relocation of APs poses a significant challenge to mobile location systems. For example, when a mobile device observes APs that have been relocated, its observations can conflict with stored AP location information and lead to location errors.

BRIEF SUMMARY OF THE INVENTION

[0007] Under one aspect of the invention, a method of and system for measuring and recovering from mobile location errors due to access point relocation is provided.

[0008] Under another aspect of the invention, a method of estimating the likelihood of a Wi-Fi enabled device being located within an estimated geographical area includes the Wi-Fi enabled device receiving signals transmitted by Wi-Fi access points in range of the Wi-Fi enabled device. The method also includes consulting a reference database to determine for each of at least one of the Wi-Fi access points from which signals were received a last-known estimated position of the Wi-Fi access point and time information associated with the last-known position for describing the age of the last-known position relative to other information in the reference database. The method further includes estimating, based on the last-known position, associated time information, and number of Wi-Fi access points from which signals were received the likelihood of the Wi-Fi enabled device being located within an estimated geographical area and displaying on a display device information based on the estimated likelihood of the Wi-Fi enabled device being located within the estimated geographical area.

[0009] Under another aspect of the invention, estimating the likelihood of the Wi-Fi enabled device being located within the estimated geographical area is further based on information that characterizes the probability that a Wi-Fi access points has moved from its

corresponding last-known position. The probability is based on the relative age of said last-known position.

[0010] Under a further aspect of the invention, the method also includes estimating, based on the last-known position and associated time information, a plurality of likelihoods of the Wi-Fi enabled device being located within a corresponding plurality of estimated geographical areas and displaying on a display device information based on the plurality of estimated likelihoods of the Wi-Fi enabled device being located within the corresponding plurality of estimated geographical areas.

[0011] Under an aspect of the invention, the method also includes consulting an historical dataset to determine for at least one of the Wi-Fi access points from which signals were received information describing past relocations for said at least one Wi-Fi access point. The estimating the likelihood of the Wi-Fi enabled device being located within the estimated geographical area is further based on the information describing past relocations for the at least one Wi-Fi access point. Optionally, the information describing past relocations for the at least one Wi-Fi access point includes an average movement frequency for the at least one Wi-Fi access point. Optionally, the information describing past relocations for the at least one Wi-Fi access point includes an aggregate average movement frequency based on a collection of movement data for a plurality of Wi-Fi access points.

[0012] Under yet another aspect of the invention, a method of estimating the likelihood of a Wi-Fi enabled device being located within an estimated geographical area includes the Wi-Fi enabled device receiving signals transmitted by Wi-Fi access points in range of the Wi-Fi enabled device. The method also includes consulting a reference database to determine for each of a plurality of the Wi-Fi access points for which signals were received a last-known estimated position of the Wi-Fi access point and determining that at least a first set of the Wi-Fi access points for which signals were received have moved from their corresponding last-known estimated positions based on the last-known estimated positions for at least a second set of Wi-Fi access points for which signals were received. The method further includes estimating, based on the number of Wi-Fi access points of the first set and the number of Wi-Fi access points of the second set, the likelihood of the Wi-Fi enabled device being located within an estimated geographical area and displaying on a display device information based on the

estimated likelihood of the Wi-Fi enabled device being located within the estimated geographical area.

[0013] Under a further aspect of the invention, the method also includes estimating, based on the number of Wi-Fi access points of the first set and the number of Wi-Fi access points of the second set, a plurality of likelihoods of the Wi-Fi enabled device being located within a corresponding plurality of estimated geographical areas and displaying on a display device information based on the plurality of estimated likelihoods of the Wi-Fi enabled device being located within the corresponding plurality of estimated geographical areas.

[0014] Under still another aspect of the invention, estimating the likelihood of the Wi-Fi enabled device being located within the estimated geographical area is further based on information that characterizes the probability that at least one of the Wi-Fi access points of the first set has moved from its corresponding last-known position, the probability being based on the number of Wi-Fi access points of the first set and the number of Wi-Fi access points of the second set .

[0015] Under another aspect of the invention, the method also includes consulting the reference database to determine for at least a first Wi-Fi access points from which signals were received a last-known estimated position of the first Wi-Fi access point and time information associated with the last-known position for describing the age of the last-known position relative to other information in the reference database. The estimating the likelihood of the Wi-Fi enabled device being located within an estimated geographical area being further based on the last-known position and associated time information of the first Wi-Fi access point.

[0016] Under a further aspect of the invention, a method of estimating the likelihood of a Wi-Fi enabled device being located within an estimated geographical area includes the Wi-Fi enabled device receiving signals transmitted by at least one Wi-Fi access point in range of the Wi-Fi enabled device. The method also includes extracting information from the signals received that identifies each of the at least one Wi-Fi access points from which signals were received and consulting a reference database to determine for at least one of the Wi-Fi access points from which signals were received a set of information identifying a corresponding set of Wi-Fi access points from which signals are expected to be received when the Wi-Fi enabled device receives signals from the at least one Wi-Fi access point of the plurality. The method further includes estimating the likelihood of the Wi-Fi enabled device being located within an

estimated geographical area based on a comparison of the information identifying Wi-Fi access points from which signals were received and the set of information identifying the corresponding set of Wi-Fi access points from which signals are expected to be received and displaying on a display device information based on the estimated likelihood of the Wi-Fi enabled device being located within the estimated geographical area.

[0017] Under yet another aspect of the invention, a method of estimating the likelihood of a Wi-Fi enabled device being located within an estimated geographical area includes the Wi-Fi enabled device receiving signals transmitted by Wi-Fi access points in range of the Wi-Fi enabled device and extracting information from the signals received that identify each of a plurality of the Wi-Fi access points from which signals were received. The method also includes consulting a reference database to determine if the identity of each of the plurality of the identified Wi-Fi access points from which signals were received are present in the reference database. The method further includes estimating the likelihood of the Wi-Fi enabled device being located within an estimated geographical area based on the number of Wi-Fi access points from which signals were received and for which identities are present in the reference database, and displaying on a display device information based on the estimated likelihood of the Wi-Fi enabled device being located within the estimated geographical area.

[0018] Under another aspect of the invention, a method of estimating the likelihood of a Wi-Fi enabled device being located within an estimated geographical area includes the Wi-Fi enabled device receiving signals transmitted by Wi-Fi access points in range of the Wi-Fi enabled device. The method also includes consulting a reference database to determine for each of at least one of the Wi-Fi access points from which signals were received a last-known estimated position of the Wi-Fi access point and time information associated with the last-known position for describing the age of the last-known position relative to other information in the reference database. The method further includes estimating, based on the last-known position and associated time information, the likelihood of the Wi-Fi enabled device being located within an estimated geographical area and displaying on a display device information based on the estimated likelihood of the Wi-Fi enabled device being located within the estimated geographical area.

[0019] Under another aspect of the invention, estimating the likelihood of the Wi-Fi enabled device being located within the estimated geographical area is further based on

information that characterizes the probability that at least one of the Wi-Fi access points has moved from its corresponding last-known position. The probability is based on the relative age of said last-known position.

[0020] Under yet another aspect of the invention, a method of estimating the likelihood of a Wi-Fi enabled device being located within an estimated geographical area includes the Wi-Fi enabled device receiving signals transmitted by Wi-Fi access points in range of the Wi-Fi enabled device. The method also includes consulting a reference database to determine for each of a plurality of the Wi-Fi access points for which signals were received a last-known estimated position of the Wi-Fi access point and determining that at least a first set of the Wi-Fi access points for which signals were received have moved from their corresponding last-known estimated positions based on the last-known estimated positions for at least a second set of Wi-Fi access points for which signals were received. The method further includes estimating, based on the number of Wi-Fi access points of the first set and the number of Wi-Fi access points of the second set, the likelihood of the Wi-Fi enabled device being located within an estimated geographical area and displaying on a display device information based on the estimated likelihood of the Wi-Fi enabled device being located within the estimated geographical area.

[0021] Under still another aspect of the invention, estimating the likelihood of the Wi-Fi enabled device being located within the estimated geographical area is further based on information that characterizes the conditional probability that at least one of the Wi-Fi access points of the first set has moved from its corresponding last-known position, the conditional probability being based on the number of Wi-Fi access points of the first set and the number of Wi-Fi access points of the second set, and wherein the conditional probability relies on a determination that last-known positions of a first Wi-Fi access point and a second Wi-Fi access point from which signals were received are in conflict. Optionally, the determination that the last-known positions of the first Wi-Fi access point and the second Wi-Fi access point from which signals were received are in conflict is based on the last known positions of the first Wi-Fi access point and second Wi-Fi access point being separated by more than a threshold distance.

[0022] Under an aspect of the invention, the method also includes consulting an historical dataset to determine for at least one of the plurality of Wi-Fi access points from which signals

were received information describing past relocations for the at least one Wi-Fi access point. The estimating the likelihood of the Wi-Fi enabled device being located within the estimated geographical area being further based on the information describing past relocations for the at least one Wi-Fi access point. Optionally, the information describing past relocations for the at least one Wi-Fi access point includes an average movement frequency for the at least one Wi-Fi access point. Optionally, the information describing past relocations for the at least one Wi-Fi access point includes an aggregate average movement frequency based on a collection of movement data for a plurality of Wi-Fi access points.

[0023] Under still another aspect of the invention, the method includes extracting information from the signals received that identifies at least one Wi-Fi access point from which signals were received and consulting a reference database to determine for the at least one Wi-Fi access point from which signals were received a set of information identifying a corresponding set of Wi-Fi access points from which signals are expected to be received when the Wi-Fi enabled device receives signals from the at least one Wi-Fi access point of the plurality. The method also includes the estimating the likelihood of the Wi-Fi enabled device being located within an estimated geographical area further being based on a comparison of the information identifying Wi-Fi access points from which signals were received and the set of information identifying the corresponding set of Wi-Fi access points from which signals are expected to be received.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

[0024] Figures 1A-B illustrate a method for calculating location estimates and center probabilities for a mobile device using a Wi-Fi Positioning System.

[0025] Figure 2 shows a scenario in which a mobile device simultaneously observes APs from two discrete clusters.

[0026] Figure 3 shows an example in which a mobile device detects two clusters of equal size.

[0027] Figure 4 shows an example in which an access point scanning device detects two clusters, one cluster having moved after the last location confirmation scan.

[0028] Figure 5 shows an example in which an access point scanning device detects two clusters, one cluster having moved after the last location confirmation scan.

[0029] Figure 6 shows an example in which a mobile device detects two clusters having differing numbers of access points.

[0030] Figure 7 shows an example in which a mobile device detects a single cluster.

DETAILED DESCRIPTION

[0031] Preferred embodiments of the invention provide methods of and systems for measuring and recovering from mobile location errors due to Wi-Fi access point relocation. The embodiments described herein quantify the probability that a particular location estimate of a mobile device made by a WPS is correct to within an arbitrary accuracy. Implementations of the invention estimate the probability that one or more APs detected by the mobile device have relocated. In one illustrative example, the probability is estimated based on all or a subset of data present in the WPS that describes, in aggregate, the movement of many APs in the system. In another example, a system estimates such a probability upon determining that at least one of the APs observed by a mobile device has moved, such as, when location information associated with the observed APs is in conflict.

[0032] Certain embodiments of the invention build on techniques, systems and methods disclosed in the U.S. Patents and Applications incorporated above. The present techniques, however, are not limited to systems and methods disclosed in the incorporated patents and applications. Thus, while reference to such systems and applications may be helpful, it is not believed necessary to understand the present embodiments or inventions. For example, those applications taught specific ways to gather high quality location data for APs (also called "scan data") so that such data may be used to estimate the geographic location of a Wi-Fi-enabled device utilizing such services and techniques. By collecting location data repeatedly from the same location and recording the date of each scan event in a database, it is possible to observe the appearance and disappearance of APs over time. In some cases, appearing APs will be new devices that have not appeared in the database previously. Similarly, disappearing APs may be removed from service and never again appear in the database. Techniques to mitigate the negative effects of APs appearing and disappearing over time are addressed in other patent applications assigned to the assignee of the present application and are not discussed in detail herein. However, there is a third case in which an AP disappears from one location and appears in another location. The use of repeated scanning and maintenance of a time-indexed database of scan data facilitates the observation of AP movement.

[0033] The ability to track AP movement allows a reference database containing AP location information to be corrected to address obsolete AP location information, and tracking also produces valuable statistics that can be used to infer the movement of APs, even if they have not been scanned in their current locations. Mining the database yields empirical probability distributions on the frequency with which APs relocate and the associations between groups of APs that move together. These probability distributions are discussed in greater detail below.

[0034] Mobile location error is typically quantified using horizontal positioning error (HPE). Given a location estimate and an arbitrary probability threshold, HPE expresses the radius of a circle centered on the location estimate that is large enough to include the mobile's actual location with the desired probability. For instance, if a location technology is extremely accurate, it could be said that there is a 95% probability that the mobile's actual location is within 10 meters of a location estimate. For a less accurate system, the HPE might be 1000 meters for a 95% probability threshold.

[0035] Detecting and recovering from location errors caused by AP movement requires a means of quantifying the probability that stored AP location information is correct. To that end, implementations of the invention employ records of the age of information in the database and the associations between different APs. That is, relationships between APs that have been observed simultaneously to form groups, called clusters herein, are used to quantify the probability values. Further, a family, which is a cluster that relocates as a single group, can be used in the same way. In general, one would expect a family to be made up of APs that are owned by a single entity.

[0036] Figure 2 depicts a scenario in which a mobile device simultaneously observes APs from two discrete clusters. The database lists the 3 APs in the first cluster as residing in Boston, and 2 APs in the second cluster as residing in from Chicago. Since the transmission ranges of the APs is far smaller than the distance from Boston to Chicago, the database must contain incorrect information about at least one of the two clusters. In other words, both clusters are currently located in either Boston or Chicago, but the mobile cannot be absolutely certain which assumption is correct. As a result, the positioning algorithm assigns center probabilities to two separate location estimates, one in Boston and one in Chicago.

[0037] Even when the mobile observes only a single cluster of APs, there is some finite probability that the observed cluster is a family that has been relocated since the last time their location was confirmed. Thus, the center probability, P, applies to scenarios with any number of observed AP clusters.

[0038] In one implementation of the invention, a method is provided for calculating the center probability, P, using empirical data on the sizes of families of moved APs and the frequency with which families are relocated.

[0039] Figures 1A-B show a method for calculating location estimates and center probabilities in a variety of scenarios. First, the mobile device detects surrounding APs (step 71) and counts the total number of observed APs. Next, it consults the database to determine the number of separate clusters and the number of APs in each cluster (step 72). In other words, because the reference database has the coordinates of all scanned APs, it is able to separate the APs (or group of APs) that are far away from each other (when the distance between APs is larger than some large enough threshold). Each cluster is associated with the last time that it was observed by the scanner (*i.e.*, the last time location information on the cluster was updated in the database containing the information about the location coordinates of all APs and last time an AP was observed by the scanner) (step 73). The term "scanner" as used herein describes a device that can detect Wi-Fi APs and also has a trusted and independent source of location information. For example, a scanner device may have both a Wi-Fi radio and a GPS system. Based on the number of clusters, the number of APs in each cluster, and the age of the information for each cluster, the invention determines center probability, P, values and other metrics as described in greater detail below.

[0040] For example, if only a single cluster is detected (at step 74), then the method determines the probability that the single cluster has moved sometime after the last confirmation of the location of the APs of the cluster has been performed. This is also called a "no location" probability herein. Meanwhile, if more than one cluster is detected, then a conditional probability of relocation for each of the APs is determined (step 76). That is, upon the condition of a conflict in location arising from the observation of APs (or clusters) that should not be detected simultaneously based on their last-known location, a probability that one of the APs (or clusters) has relocated is determined. In addition, the relative size of the clusters is compared with an empirical data set to determine the relative likelihood that each cluster of

APs has moved sometime after the last confirmation of the location of the APs of the clusters has been performed (step 76). These techniques are set forth in more detail below.

[0041] Next, the method calls for determining the center probabilities for each detected cluster. As set forth in greater detail below, different formulas are used to determine these values depending on the number of clusters detected.

[0042] The method set forth above and shown in Figures 1A-B is described as having certain steps performed on a mobile device. While the initial search for APs within range of the mobile device is performed by the mobile device, any of the other steps can be performed on the mobile device or a remote computer system.

Probability of AP movement

[0043] Given up-to-date knowledge of AP locations, the actual location of the mobile device is accurately modeled with a radially-symmetric probability distribution centered at the estimated mobile location (location center). However, when the validity of AP location information is uncertain due to AP relocation, the location estimation error can be quantified with two parameters. The first parameter is the radius about the location center within which the mobile resides with a given probability, conditioned on the fact that the location center is accurate to an arbitrary precision. The second parameter is the location center probability, P, which is accurate to an arbitrary precision.

[0044] Given a database of AP locations measured over time, we construct an approximation to the empirical cumulative distribution function (CDF) to describe the conditional probability that a single AP will have relocated within a specified time interval given that the AP has been moved (*e.g.*, a conflict exists between locations of observed APs). Known methods may be used to approximate the CDF from the empirical dataset of AP that have moved. Because the database necessarily spans only a finite time interval, the CDF must saturate to probability 1 at the age of the oldest data. Thus, as more data is collected and the time spanned by the database increases, the CDF will be updated to reflect improved knowledge of the long-term relocation behavior of APs.

[0045] In one implementation of the invention, the CDF that a given AP will have relocated as a function of the time since it was last observed by the scanner is approximated. Time, t, is measured in units of months, and the approximation of the CDF is defined as $Y(t)$, where

$$\begin{aligned} Y(t) &= a, && \text{if} \quad 0 < t \le h_0, \\ Y(t) &= bt + c, && \text{if} \quad h_0 < t \le h_1, \\ Y(t) &= 1, && \text{if} \quad h_1 < t. \end{aligned} \tag{1a}$$

and a, b, c, h_0, h_1 are some constants, which are found empirically by analyzing the database of moved APs. One example of $Y(t)$ is given by the formula

$$\begin{aligned} Y(t) &= 0.02, && \text{if} \quad 0 < t \le 3, \\ Y(t) &= (2.51t - 5.54)/100, && \text{if} \quad 3 < t \le 42, \\ Y(t) &= 1, && \text{if} \quad 42 < t, \end{aligned} \tag{1}$$

which corresponds to step 76 of Figure 1B.

[0046] The values of $Y(t)$ represent a probability weight that the AP has been moved during the last t months, given that the last time it was scanned was t months ago, and given that the system has determined the AP is in a set of observed APs, one or more of which have moved.

[0047] Function (1) was determined by analyzing the database of moved APs. This database has all the information about moved APs, including the history of all locations of all detected moved APs and time of each AP being observed at each location.

[0048] As the database ages and contains a growing time series of scanned data, Y(t) is updated as follows:

$$\begin{aligned} Y(t) &= \frac{84}{100(42+n)}, && \text{if} \quad 0 < t \le 3, \\ Y(t) &= \frac{41.16+n}{(39+n)(42+n)}t - \frac{2.16}{39+n}, && \text{if} \quad 3 < t \le 42+n, \end{aligned} \tag{2}$$

$$Y(t) = 1, \quad \text{if} \quad 42 + n < t.$$

where $n = 1,2,3,...$ represents the number of 6 month intervals since the first use of the function.

[0049] The conventional definition of HPE works well when the distribution of error in a location estimate is unimodal. Errors arising from noisy range estimation can be considered unimodal, but errors due to AP relocation are multimodal and not well quantified by HPE. For example, in a scenario in which a mobile observes two APs, one AP has correct location information in the database, but the other AP has been relocated from 100 kilometers away. The database has not been updated to reflect the relocation of the second AP, so it supplies the mobile with incorrect location information. If a single location estimate is constructed and a single circle drawn around it to express HPE, the radius will be on the order of 50 kilometers. However, if the underlying structure of the scenario is examined, it is seen that the location error is bimodal, and is better described with two circles, corresponding to each of the two APs. Each circle has its own HPE and center probability, P, which expresses the relative confidence in the information about each of the two APs in the database.

Clusters of equal size

[0050] Figure 3 shows an example in which the mobile device 21 detects two equal size clusters of APs: one cluster of 3 APs from New York 22 and one cluster of 3 APs from Miami 23. Note that a cluster might be as small as a single AP and has no upper limit on the number of members.

[0051] When a mobile device observes multiple APs whose locations as recorded in the database should preclude them from being observed simultaneously, the mobile device recognizes that one or more APs must have relocated. More generally, the mobile device may observe multiple clusters of APs, and one or more clusters may have relocated as a family unit. Because location estimation using any of the individual clusters would lead to disparate center locations, the mobile device must resolve the conflicting information in order to decide on a single location center.

[0052] Under one implementation, the conflicting location information is resolved by using the age of the newest information on each cluster. Function Y(t) shows that the probability of obsolescence of measurement information increases with the age of the measurement, so newer

measurements are considered to be more reliable. Thus, the most effective method of choosing a location center is to assume that the cluster with the most recent measurement data has not relocated. Based on the difference in ages between the measurements of each cluster, the mobile device can then express its confidence in the chosen location center.

[0053] For example, consider the case where the mobile has observed two clusters of n APs. Each AP in each cluster is assigned an arbitrary index from 1 to n. The time since AP_i of cluster 1 was last observed is denoted t_i, and the time since AP_j of cluster 2 was last observed is T_j. The ratio, r, can be calculated as follows:

$\min(t_i)$ = the smallest number among $t_1, t_2, \ldots, t_n$;

$\min(T_j)$ = the smallest number among $T_1, T_2, \ldots, T_n$.

$$r = \frac{Y(\min(t_i))}{Y(\min(T_j))}$$

[0054] The parameter P_1 is assigned to the location estimate based on cluster 1 and probability parameter P_2 is assigned to the location estimate based on cluster 2. P_1 and P_2 are calculated according to the following equations (this corresponds to step 78 of Figure 1B):

$$P_1 = \frac{1}{1+r}, \qquad P_2 = \frac{r}{1+r} \tag{3}$$

[0055] In practice, there are several ways to apply P_1 and P_2. For example, only the location corresponding to the maximum center probability could be reported. Alternatively, a probability threshold could be set and a location only reported if one of the two center probabilities exceeds that threshold. If neither center probability exceeds the threshold (note that they are complementary and cannot both exceed 0.5 simultaneously), then the location estimate could be considered to be too unreliable, and a "no location" result reported. As a third possibility, both location estimates along with their associated center probabilities could be reported.

Cluster movement

[0056] The sizes of observed clusters also affect the probability that the clusters have relocated. It has been discovered that small families are much more common than large families because there are a relatively small number of organizations that relocate with large numbers of APs. Thus, a large cluster is unlikely to be made up of a single family, and larger clusters are less likely to have been relocated.

[0057] Under another implementation of the invention, statistics based on the ratio of cluster sizes, a method for updating cluster information in the database, and a method for identifying and tracking families of APs as they relocate is described. As stated above, it is assumed that the scanning equipment used to compile the database possesses accurate location information independent of observed APs, so the database has high precision information regarding where and when each AP was observed. For example, the scanning equipment used to compile the database may use a GPS system to determine its position when detecting APs in a given geographical area.

[0058] For each scan in the database, any APs that have been observed simultaneously as belonging to the same cluster are identified during a scanning event to update and/or create the reference database. Then, whether any of the APs in the cluster have been relocated since the last time they were observed is also identified. If one or more APs have been relocated, previous observations are consulted to see if any of the relocated APs were moved as a family from their previous location(s). Taking the general definition that a family can be as small as one AP, each instance of family relocation is examined as follows. Each time a family relocates, the number of APs in the family and the number of APs not in the family is counted that combined to form the new cluster. That is, the number of relocated APs is compared to the number of APs that were not relocated. The notation

$$i \rightarrow j$$

is used to represent the situation in which a family of i relocated APs is observed at the same time as j APs that have not been relocated since the last scan. In other words, the scanning device detected i new APs and j old APs at the same time and place.

[0059] Figure 4 shows an example in which a scanner located in Boston 31 observes 1 AP that has been moved from Dallas 32 and 2 APs that have remained, unmoved, in Boston 33 since the last time they were scanned. In this case, one pair $1 \rightarrow 2$ would be recorded.

[0060] Referring to Figure 5, if the same APs had been observed by a scanner in Dallas 41, then the scanner would conclude that the two APs from Boston 43 had been moved, and 1 AP from Dallas 42 remained, unmoved, and record $2 \rightarrow 1$.

[0061] After recording the total number of different pairs $i \rightarrow j$, parameter K_{ij} is defined to describe the empirical probability that, given simultaneously observed clusters of size i and j with conflicting location information, the cluster of size i was relocated (for the case when $i \neq j$) (this corresponds to step 76 of Figure 1B). Thus,

$$K_{ij} = \frac{number\_of\_pairs\_i \rightarrow j}{total\_number\_of\_pairs\_i \rightarrow j\_and\_j \rightarrow i} \qquad (4)$$

and vice versa,

$$K_{ji} = \frac{number\_of\_pairs\_j \rightarrow i}{total\_number\_of\_pairs\_i \rightarrow j\_and\_j \rightarrow i}. \qquad (5)$$

[0062] When i=j, $K_{ij} = \frac{1}{2}$. Therefore, $K_{ij} + K_{ji} = 1$.

[0063] In order to keep the values of K_{ij} as accurate as possible, it is preferred to recalculate them periodically to reflect new scan data.

Clusters of equal age

[0064] In another implementation, a method is described for assigning center probabilities to location estimates based on two clusters when one of the clusters has been relocated and database information on both clusters is of the same age. In this case, the cluster sizes are used to determine the probability that either of the two clusters was relocated as a family.

[0065] Figure 6 shows a scenario in which a first cluster 52 and a second cluster 53 are observed simultaneously by the mobile device 51, but the location information in the database indicates that the clusters are separated by a distance much greater than the transmission range

of the APs. The first cluster consists of n APs and the second cluster consists of m APs. The positioning algorithm assigns center probabilities, P_1 and P_2, based on the most recent observation of each cluster.

[0066] It this illustrative example, it is assumed that the database information is of equal age for both clusters. Therefore, time does not play any role in the computation of the probabilities of the center location. The following definition for r is used:

$$r = \frac{K_{nm}}{(1 - K_{nm})}.$$

[0067] Center probability P_1 is associated with the location of the first cluster and center probability P_2 is associated with the location of the second cluster. The following formulas define P_1 and P_2 (this corresponds to step 78 of Figure 1B):

$$P_1 = \frac{1}{1+r}, \qquad P_2 = \frac{r}{1+r} \tag{6}$$

Clusters of differing sizes and ages

[0068] In another embodiment of the invention, both cluster size and cluster age (the age of the most recent update on the cluster's location in the database) are used to assign center probabilities (this corresponds to step 78 of Figure 1B).

[0069] Referring again to Figure 6, the first cluster 52 is of size n, and the second cluster 53 is of size m. The time since AP j of the first cluster 52 was last observed is denoted t_j, and the time since AP k of the second cluster 53 was last observed is T_k. The values of $\min(t_i)$ and $\min(T_j)$ are determined as set forth above.

[0070] Next, the ratio, r, is determined using parameters based on both cluster size ratio, K_{nm}, and cluster age, Y, as follows:

$$r = \frac{K_{nm} Y(\min(t_i))}{(1 - K_{nm}) Y(\min(T_j))}$$

Once again, the center probabilities are as follows:

$$P_1 = \frac{1}{1+r}, \qquad P_2 = \frac{r}{1+r} \tag{7}$$

Movement of a single cluster

[0071] Under another implementation, a center probability is determined when the mobile device observes only a single cluster. In other words, if the mobile device finds only a single cluster of APs, there is some probability that the cluster has been relocated as a single family, and that possibility is quantified with a center probability.

[0072] Figure 7 shows a situation in which the mobile device 61 has observed a single cluster 62 of n APs, and the last time the first AP in this cluster was updated was t_1 months ago. The second AP in this cluster was updated t_2 months ago, and the n^{th} AP was updated t_n months ago. As above,

$\min(t_i)$ = the smallest number among $t_1, t_2, \ldots, t_n$.

[0073] When a mobile device observes only a single cluster (*i.e.*, there is no cluster conflict), the center probability is calculated as a function of following items: (1) the number of APs detected by the mobile device which are known in the database, (2) the number of APs detected by the mobile device which are not known and are not in the database, (3) other APs expected to neighbor a particular AP and the number of expected neighbors based on the collection of the history of observation of scanned results including that AP, (4) the minimum time elapsed from the last time that location of known APs is confirmed by a scanner, $\min(t_i)$, and (5) the confidence in the known location of individual APs that are in the database via a scanner detection. We can find, empirically, from the dataset of moved APs the probability of movement of an AP as a function of these parameters.

[0074] The above parameters are now discussed in more detail. A set of APs detected by a mobile device can be divided into two groups. The first group consists of APs that are present in the database by virtue of a scanner or derived from the locations of other known APs, whose locations are known with some degree of certainty. For example, when an unknown AP is

scanned by the mobile device along with other known APs, then the position of the unknown AP can be estimated using a triangulation based on all known APs that are observed by the mobile device. The second group consists of APs that are new and have neither been located by a scanner nor derived from other known APs. Thus, some of the APs present in the database may be associated with a location that has been derived from other known APs. Therefore, center probability can be calculated as a function of number of known APs having locations.

[0075] Center probability can also be a function of the second group, which contains new APs that have been detected along-side of known APs that were not found by the scanner at the time of scanning. For example, assume a mobile device observes 2 known APs and 10 unknown APs, and the $\min(t_i)$ parameter of the cluster of the 2 known APs is equal to 1 year. Then, empirically, the probability that the cluster of 2 known APs has been relocated, as a family of APs, can be found from the data set of moved APs, given that the mobile device observed 2 known APs and 10 unknown APs, and the $\min(t_i)$ parameter of the cluster of the 2 known APs is equal to 1 year. Therefore, we can also find the empirical probability that cluster of 2 known APs has not moved.

[0076] Center probability can also be a function of expected neighbors of collective set of scanned APs. Based on the past history of observing a given AP with another set of APs at the time of systematic scanning and/or by detection of other mobile devices a set of expected neighbors is established for any known AP along with an expected number of neighbors. These measures are based on past observations and the density of APs in the neighborhood of the given AP. When a mobile device detects a set of APs, the probability that the observed cluster of APs has been moved can be determined based on consistency between the expected neighbors and what the mobile device observes. Probability that an AP has been moved, given a set of expected neighbors, expected number of neighbors, APs currently detected by the mobile, and number presently detected is determined empirically based on historical data gathered by, for example, scanners.

[0077] Center location can also be a function of time. In other words, the probability that an AP has been relocated can be determined as a function of time. For example, if an AP has been confirmed at a given location, the probability that the AP is still located in the same location after one day is higher than after one year.

[0078] Also, the certainty of an estimated location for a particular AP can be different. In other words, upon conducting a scanner survey of an area, the estimated location of an AP may be determined with a high confidence (*e.g.*, 100%) or low confidence. Each AP that is found during a scanner pass is associated with a corresponding confidence. Center location probability can be a function of confidence in location of known APs. In one implementation, the confidence of all known APs detected by a mobile device can be estimated by taking the maximum confidence of all the detected APs.

[0079] The probability $P_{recent\_location}$ that a given cluster of n APs did not change its location can be estimated based on a collection of information on moved APs and moved families of APs (*e.g.*, a "moved APs" database can contain information including previous and current location coordinates, the date a move was detected by a scanner, the size of families of APs which have been moved, etc.). Probability parameter $P_{recent\_location}$ depends on the cluster size n and time $\min(t_i)$. In other words, one can find from the moved APs database the following probability depending on two parameters:

$$P_{recent\_location} = P_{recent\_location}(n, \min(t_i)). \quad (8)$$

[0080] Probability parameter $P_{no\ location}$ represents the probability that a given cluster of n APs has been moved to some other location, which has yet to be detected yet by a scanner. Thus, the "no location" subscript as used in this context represents the fact that an affirmative location cannot be provided until the cluster is detected by a scanner at some later time. Obviously,

$$P_{no\ location} = 1 - P_{recent\ location} \quad (9)$$

[0081] The potential measures of probabilities of center location as a function of the set of parameters given above (*e.g.*, number of scanned APs that are known in the database, number of scanned APs that are not known and are not in the database, expected neighbors and expected number of neighbors based on scanned AP data, minimum time elapsed after the last time that the location of known scanned APs is confirmed, and confidence in the location of individual known scanned APs) can be found empirically from the moved APs information.

More than two clusters

[0082] In another implementation of the disclosed embodiment, the method for determining center probabilities is generalized to scenarios in which the mobile device detects X clusters of APs, where X can be greater than 2. The variable w_n is the size of the n^{th} cluster (n=1,2,3,…X), and T_n is the age (measured in months) of the most recent scan of cluster n in the database. Ratio r is determined as follows:

$$r_{ij} = \frac{k_{w_i, w_j} Y(T_i)}{(1 - k_{w_i, w_j}) Y(T_j)}, \quad (i, j = 1,2,3,..., X; i \neq j)$$

where k_{w_i, w_j} are K-values, Y is the function given by the formula (1). Next, Q_n is defined as follows:

$$Q_n = \frac{1}{1+r_{n,1}} \cdot \frac{1}{1+r_{n,2}} \cdots \frac{1}{1+r_{n,n-1}} \cdot \frac{1}{1+r_{n,n+1}} \cdots \frac{1}{1+r_{n,X}}$$

[0083] The following probabilities are associated with the location of n^{th} cluster (this corresponds to step 79 of Figure 1B):

$$P_n = \frac{Q_n}{\sum_{i=1}^{X} Q_i} \qquad (10)$$

Influence of environment and history of particular AP movement on probabilities of center location.

[0084] Above, techniques for defining and estimating the probabilities of center locations of a mobile device were provided. These probabilities were based on an analysis of aggregate data, that is, a relatively large collection of location and movement data for relatively large number of APs. However, improvements to those estimates can be made using the APs

environment and history of a given set of APs. For example, if it is known from the history of APs (which can be extracted from the moved APs database), that a particular AP relocates once per 6 months on average, and a second AP relocates once per 2 years on average, then this information is taken into account to improve the probabilities estimate. In other words, the variable "freq" representing the average movement frequency of a particular AP or cluster of APs may be taken into account. In such a case, instead of empirically estimated CDF approximation function Y(t) (formulas 1 and 2), an empirical estimate of the CDF approximation function Y(t, freq) is provided, where the joint probability distribution of random variables t and "freq" is considered, given that a particular AP (or cluster of APs) is determined to be within a set of detected APs, one or more of which have moved.

[0085] Another parameter that may affect the accuracy of the center probability estimate is the APs environment, or surrounding APs. For example, assume a first AP was detected in the recent past by mobile devices in conjunction with 10 other surrounding known APs on average. A mobile device now reports the observation of the first AP along with the detection of a second known AP, the location of which should preclude the two APs from being observed simultaneously. Assume also that the second known AP was previously scanned or observed without any other surrounding APs. Thus, at least one of the two APs has moved from its previously recorded location. This circumstance makes it is very likely that the first AP has been moved because, in addition to the conflicting location situation, the surrounding AP density changed substantially.

[0086] Considering another example, a third AP has been reported in the recent past by mobile devices with 5 known surrounding APs on average, and now a particular mobile device reports the third AP with 4 unknown surrounding APs. In this case, it is very likely that the third AP has been moved, because although the surrounding APs density has not changed substantially (4 versus 5), and no location conflict exists, the 4 presently detected APs are not the same, or a subset of the same, APs that are expected to be surrounding the third AP. The same would hold true were a location conflict to exist.

[0087] As one example of the use of an AP's environment parameter in adjusting its corresponding movement probability, a parameter, *E*, is provided that describes a change in APs surrounding density environment:

$$E = \frac{E_1 + 1}{E_2 + 1},$$

where E_1 = "recent average number of known APs detected along with given a given AP (or a given cluster of APs) by mobile devices" and E_2 = "current number of known APs detected along with a given AP (or a given cluster of APs) by mobile devices". Furthermore, a "density environment" function *DE* is as follows:

$$DE = \max(E, E^{-1}).$$

[0088] When $DE > C_1$, where C_1 is a threshold constant determined empirically (*e.g.*, C_1 = 4), then the given AP (or given cluster of APs) is designed as having moved from its previous known location.

[0089] In general, all 3 parameters t, "freq" and *DE* can be combined in the estimation of center probability. Thus, instead of empirically estimated CDF approximation function Y(t) (formulas 1 and 2), an empirically estimated CDF approximation function Y(t, freq, *DE*) is provided where the joint probability distribution of random variables t, "freq" and *DE* are considered, given that an AP (or cluster of APs) is determined to be within a set of detected APs, one or more of which have moved.

[0090] In the case of single cluster observed by the mobile device, the empirically estimated parameter $P_{recent\_location}$ is provided as a function of cluster size *n*, time $\min(t_i)$, "freq", and *DE* as follows:

$$P_{recent\_location} = P_{recent\_location}(n, \min(t_i), \text{“freq”}, DE),$$

and

$$P_{no\ location} = 1 - P_{recent\ location}.$$

Known-in-advance location bias correction

[0091] Above, the meaning and techniques for determining K_{ij} parameters, or K-values were described. Those techniques relied on knowing the ground truth location during scanning. In other words, while scanning and recording all possible pairs $i \rightarrow j$, a highly accurate geographical location was known. Typically, an accurate geographical location is only known when using GPS or some other reliable positioning technology.

[0092] However, there are some locations (*e.g.*, indoors locations) when GPS or any other, typically accurate, positioning technology does not work or is not reliable. This means that it is possible that the location information for some APs in the reference database may be biased to certain known-in-advance locations, where a "known-in-advance" location is one that is known with a relatively high degree of accuracy. In certain implementations, it is desirable to ensure this bias does not affect the accuracy of the K-values for the entire distribution of all geographical locations of all possible AP observations. Note that the entire distribution consists of a first portion of data for which the geographical location was known with high precision during data collection and a second portion of the data for which the geographical location was not known or was inaccurate (*e.g.*, locations where GPS does not work).

[0093] In such a case, a sufficiently large (*e.g.*, on the order of 10,000 samples) number of AP observation samples containing the information about $i \rightarrow j$ pairs from the geographical locations where GPS location did not work can be selected. For each of the selected observations, an accurate geographical location can be determined using other means, and that location replaces the missing or inaccurate location information in the reference database. At this point, the updated data can then be reprocessed, as set forth in detail above. This will help to increase the diversity of the samples and reduce any potentially existing bias in the K-value estimations.

[0094] There may be instances in which there exist an insufficient number of samples having manually corrected location data to reduce potentially biased K-values as described above. However, it may still be desirable to reduce such bias. In such a case, unbiased K-values can be estimated from the known-in-advance scan locations statistics as set forth below.

[0095] The variable N is the total number of existing different families of APs in the known-in-advance possible scan population, and M is the total number of APs in the known-in-

advance possible scan population. Proportion (weight) w_1 represents the proportion of all 1-size families of APs among N families, w_2 the proportion (weight) of all 2-size families of APs among N families, …, and w_i the proportion (weight) of all i-size families of APs among N families. The variable L is the largest possible size of a family of APs that exists. The following holds true:

$$\sum_{i=1}^{L} w_i = 1 \qquad (11)$$

and

$$N = \frac{M}{\sum_{i=1}^{L} i w_i}. \qquad (12)$$

[0096] The variable p_i is the rate of movement of i-size families (i=1,2,3,…). In other words,

$$p_i = \frac{number\_of\_moved\_during\_the\_year\_i-size\_families}{total\_number\_of\_i-size\_families} \qquad (13)$$

[0097] The variable a_{ij} is the total collected number of i-size clusters vs. j-size clusters situations among the known-in-advance scan location data. For example, a_{13} represents the total collected number of 1-size clusters vs. 3-size clusters, or, as defined above, the total number of $1 \to 3$ pairs plus the total number of $3 \to 1$ pairs. Given this, the total number of 1-size families that have been moved can be found. This number is determined as follows:

$$a_{11} + k_{12}a_{12} + k_{13}a_{13} + ... + k_{1L}a_{1L} = a_{11} + \sum_{i=2}^{L} k_{1i}a_{1i}$$

[0098] This value is also equal to $w_1 Np_1$. Thus, the following system of equations results:

$$a_{11} + k_{12}a_{12} + k_{13}a_{13} + ... + k_{1L}a_{1L} ... = w_1 Np_1 \quad (e_1)$$

$$a_{22} + (1-k_{12})a_{12} + k_{23}a_{23} + ... + k_{2L}a_{2L} ... = w_2 Np_2 \quad (e_2)$$

$$a_{33} + (1-k_{13})a_{13} + (1-k_{23})a_{23} + ... + k_{3L}a_{3L} ... = w_3 Np_3 \quad (e_3)$$

.....

.....

$$a_{LL} + (1-k_{1L})a_{1L} + (1-k_{2L})a_{2L} + ... + (1-k_{L-1,L})a_{L-1,L} ... = w_L Np_L \quad (e_{l1})$$

[0099] Everything on the left-hand side of the system $(e_1) - (e_L)$ is known. Thus, the values for all of the ratios $\frac{w_i p_i}{w_j p_j}$ can be found. It is important to note that parameters w_i and p_i do not depend on whether only the known-in-advance locations scan population is taken into account, or whether the entire population (*i.e.*, known location scans and unknown location observations) are taken into account.

[0100] Consider now the whole population of data from both known-in-advance and unknown locations. The variable b_{ij} is the total collected number of i-size clusters vs. j-size clusters situations among the entire population of scan and observation data. For the sake of consistency, the notation k_{ij} for the K-values, as used above for only locations known-in-advance, is retained in the description below for the entire population of data. However, it is

noted that the K-values of the two scenarios are not necessarily equal because there may be bias of estimation of K-values to the known-in-advance scan population.

[0101] The variable N_1 is the total number of existing different families of APs if the entire population (*i.e.*, known location scans and unknown location observations) is taken into account. Using logic similar to the description immediately above, the following system of equations results:

$$b_{11} + k_{12}b_{12} + k_{13}b_{13} + ... + k_{1L}b_{1L} ... = w_1 N_1 p_1 \qquad (f_1)$$

$$b_{22} + (1 - k_{12})b_{12} + k_{23}b_{23} + ... + k_{2L}b_{2L} ... = w_2 N_1 p_2 \qquad (f_2)$$

$$b_{33} + (1 - k_{13})b_{13} + (1 - k_{23})b_{23} + ... + k_{3L}b_{3L} ... = w_3 N_1 p_3 \qquad (f_3)$$

.....

.....

$$b_{LL} + (1 - k_{1L})b_{1L} + (1 - k_{2L})b_{2L} + ... + (1 - k_{L-1,L})b_{L-1,L} ... = w_L N_1 p_L \qquad (f_L)$$

[0102] In the above system of equations (f_1) - (f_L), the parameters k_{ij} must be estimated. The values of b_{ij} are known because the information is derived from scan data, and all of the ratios $\frac{w_i p_i}{w_j p_j}$ are known, as provide above. Therefore, L independent equations exist and $\frac{L(L-1)}{2}$ variables k_{ij} exist. The value of $\frac{L(L-1)}{2} - L = \frac{L(L-3)}{2}$ different k_{ij} parameters, described immediately above, are obtained during scan data collection. These parameters are substituted in the system of equations (f_1) - (f_L), in order to find the remaining L different k_{ij} parameters.

[0103] Thus, the k_{ij} parameters of the entire population of data (*i.e.*, known and unknown locations) are estimated using a "smart guess" based on knowledge of some statistics of entire population (*e.g.*, b_{ij} values) and knowledge of some statistics (*e.g.*, K-values) of known-in-advance scan location information. In other words, the $\frac{L(L-3)}{2}$ estimates of the k_{ij} parameters, found as a solution to the system of equations $(e_1) - (e_L)$, are substituted into the system of equations $(f_1) - (f_L)$ to find the remaining L k_{ij} parameters.

[0104] The techniques and systems disclosed herein may be implemented as a computer program product for use with a computer system or computerized electronic device. Such implementations may include a series of computer instructions, or logic, fixed either on a tangible medium, such as a computer readable medium (*e.g.*, a diskette, CD-ROM, ROM, flash memory or other memory or fixed disk) or transmittable to a computer system or a device, via a modem or other interface device, such as a communications adapter connected to a network over a medium.

[0105] The medium may be either a tangible medium (*e.g.*, optical or analog communications lines) or a medium implemented with wireless techniques (*e.g.*, Wi-Fi, cellular, microwave, infrared or other transmission techniques). The series of computer instructions embodies at least part of the functionality described herein with respect to the system. Those skilled in the art should appreciate that such computer instructions can be written in a number of programming languages for use with many computer architectures or operating systems.

[0106] Furthermore, such instructions may be stored in any tangible memory device, such as semiconductor, magnetic, optical or other memory devices, and may be transmitted using any communications technology, such as optical, infrared, microwave, or other transmission technologies.

[0107] It is expected that such a computer program product may be distributed as a removable medium with accompanying printed or electronic documentation (*e.g.*, shrink wrapped software), preloaded with a computer system (*e.g.*, on system ROM or fixed disk), or distributed from a server or electronic bulletin board over the network (*e.g.*, the Internet or World Wide Web). Of course, some embodiments of the invention may be implemented as a

combination of both software (*e.g.*, a computer program product) and hardware. Still other embodiments of the invention are implemented as entirely hardware, or entirely software (*e.g.*, a computer program product).

[0108] Moreover, the techniques and systems disclosed herein can be used with a variety of mobile devices. For example, mobile telephones, smart phones, personal digital assistants, satellite positioning units (*e.g.*, GPS devices), and/or mobile computing devices capable of receiving the signals discussed herein can be used in implementations of the invention. The location estimate, corresponding expected error of the position estimate, and/or the probability values can be displayed on the mobile device and/or transmitted to other devices and/or computer systems. Further, it will be appreciated that the scope of the present invention is not limited to the above-described embodiments, but rather is defined by the appended claims; and that these claims will encompass modifications of and improvements to what has been described.

[0109] What is claimed is:

1. A method of estimating the likelihood of a Wi-Fi enabled device being located within an estimated geographical area, the method comprising:

the Wi-Fi enabled device receiving signals transmitted by at least one Wi-Fi access point in range of the Wi-Fi enabled device;

consulting a reference database to determine for the at least one Wi-Fi access point from which signals were received a last-known estimated position of each of the at least one Wi-Fi access point and time information associated with each of said last-known positions for describing the age of each said last-known position relative to other information in the reference database;

estimating, based on at least one of the last-known positions, associated time information, and number of Wi-Fi access points from which signals were received, the likelihood of the Wi-Fi enabled device being located within an estimated geographical area; and

displaying on a display device information based on the estimated likelihood of the Wi-Fi enabled device being located within the estimated geographical area.

2. The method of claim 1, wherein estimating the likelihood of the Wi-Fi enabled device being located within the estimated geographical area is further based on information that characterizes the probability that a Wi-Fi access point has moved from its corresponding last-known position, the probability being based on the relative age of said last-known position.

3. The method of claim 2, further comprising determining the information that characterizes the probability that a Wi-Fi access point has moved from its corresponding last-known position.

4. The method of claim 3, wherein the determining the information that characterizes the probability that a Wi-Fi access points has moved from its corresponding last-known position based on the relative age of said last-known position comprises:

determining a set of Wi-Fi access points, each Wi-Fi access point of the set located at a first geographic position for the corresponding Wi-Fi access point at a first point in time, and having moved to a second geographic position for the corresponding Wi-Fi access point at a second point in time; and

based on the set of Wi-Fi access points that moved and based on the amount of time between the first and second points in time, determining information that characterizes the probability that a Wi-Fi access point has moved from its last-known position based on the relative age of said last-known position.

5. The method of claim 1, further comprising:

consulting an historical dataset to determine for at least one of the Wi-Fi access points from which signals were received information describing past relocations for said at least one Wi-Fi access point;

wherein estimating the likelihood of the Wi-Fi enabled device being located within the estimated geographical area is further based on the information describing past relocations for the at least one Wi-Fi access point.

6. The method of claim 5, wherein the information describing past relocations for the at least one Wi-Fi access point includes an average movement frequency for the at least one Wi-Fi access point.

7. The method of claim 5, wherein the information describing past relocations for the at least one Wi-Fi access point includes an aggregate average movement frequency based on a collection of movement data for a plurality of Wi-Fi access points.

8. A method of estimating the likelihood of a Wi-Fi enabled device being located within an estimated geographical area, the method comprising:

the Wi-Fi enabled device receiving signals transmitted by at least one Wi-Fi access point in range of the Wi-Fi enabled device;

extracting information from the signals received that identifies each of the at least one Wi-Fi access points from which signals were received;

consulting a reference database to determine for at least one of the Wi-Fi access points from which signals were received a set of information identifying a corresponding set of Wi-Fi access points from which signals are expected to be received when the Wi-Fi enabled device receives signals from the at least one Wi-Fi access point of the plurality;

estimating the likelihood of the Wi-Fi enabled device being located within an estimated geographical area based on a comparison of the information identifying Wi-Fi access points from which signals were received and the set of information identifying the corresponding set of Wi-Fi access points from which signals are expected to be received; and

displaying on a display device information based on the estimated likelihood of the Wi-Fi enabled device being located within the estimated geographical area.

9. The method of claim 8, wherein the estimating the likelihood of the Wi-Fi enabled device being located within an estimated geographical area further comprises comparing a first number of Wi-Fi access points from which signals were received and for which information was found in the reference database to a second number of Wi-Fi access points from which signals are expected to be received when the Wi-Fi enabled device receives signals from the at least one Wi-Fi access point of the plurality.

10. The method of claim 8, wherein the estimating the likelihood of the Wi-Fi enabled device being located within an estimated geographical area further comprises comparing a first number of Wi-Fi access points from which signals were received and for which no identification information was found in the reference database to a second number of Wi-Fi access points from which signals are expected to be received when the Wi-Fi enabled device receives signals from the at least one Wi-Fi access point of the plurality.

11. The method of claim 8, wherein the Wi-Fi access points from which signals are expected to be received when the Wi-Fi enabled device receives signals from the at least one Wi-Fi access point of the plurality are identified based on a history of one or more Wi-Fi enabled devices receiving signals from the expected Wi-Fi access points when receiving signals from the at least one Wi-Fi access point.

12. The method of claim 8, further comprising:

consulting an historical dataset to determine for at least one of the of Wi-Fi access points from which signals were received information describing past relocations for the at least one Wi-Fi access point;

wherein estimating the likelihood of the Wi-Fi enabled device being located within the estimated geographical area is further based on the information describing past relocations for the at least one Wi-Fi access point.

13. The method of claim 12, wherein the information describing past relocations for the at least one Wi-Fi access point includes an average movement frequency for the at least one Wi-Fi access point.

14. The method of claim 12, wherein the information describing past relocations for the at least one Wi-Fi access point includes an aggregate average movement frequency based on a collection of movement data for a plurality of Wi-Fi access points.

15. The method of claim 8, further comprising:

consulting a reference database to determine for the at least one of the plurality of Wi-Fi access points from which signals were received a last-known estimated position of each of the at least one of the Wi-Fi access points and time information associated with each of said last-known positions for describing the age of each said last-known position relative to other information in the reference database;

wherein the estimating the likelihood of the Wi-Fi enabled device being located within an estimated geographical area is further based on at least one of the last-known positions and associated time information.

16. A method of estimating the likelihood of a Wi-Fi enabled device being located within an estimated geographical area, the method comprising:

the Wi-Fi enabled device receiving signals transmitted by Wi-Fi access points in range of the Wi-Fi enabled device;

extracting information from the signals received that identify each of a plurality of the Wi-Fi access points from which signals were received;

consulting a reference database to determine if the identity of each of the plurality of the identified Wi-Fi access points from which signals were received are present in the reference database;

estimating the likelihood of the Wi-Fi enabled device being located within an estimated geographical area based on the number of Wi-Fi access points from which signals were received and for which identities are present in the reference database; and

displaying on a display device information based on the estimated likelihood of the Wi-Fi enabled device being located within the estimated geographical area.

17. The method of claim 16, wherein the estimating the likelihood of a Wi-Fi enabled device being located within an estimated geographical area further comprises comparing the number of Wi-Fi access points from which signals were received and for which identities are present in the reference database to a number of Wi-Fi access points from which signals were received and for which identities are not present in the reference database.

18. The method of claim 16, further comprising:

consulting an historical dataset to determine for at least one of the plurality of Wi-Fi access points from which signals were received information describing past relocations for the at least one Wi-Fi access point;

wherein estimating the likelihood of the Wi-Fi enabled device being located within the estimated geographical area is further based on the information describing past relocations for the at least one Wi-Fi access point.

19. The method of claim 18, wherein the information describing past relocations for the at least one Wi-Fi access point includes an average movement frequency for the at least one Wi-Fi access point.

20. The method of claim 18, wherein the information describing past relocations for the at least one Wi-Fi access point includes an aggregate average movement frequency based on a collection of movement data for a plurality of Wi-Fi access points.

21. A method of estimating the likelihood of a Wi-Fi enabled device being located within an estimated geographical area, the method comprising:

the Wi-Fi enabled device receiving signals transmitted by Wi-Fi access points in range of the Wi-Fi enabled device;

consulting a reference database to determine for each of a plurality of the Wi-Fi access points for which signals were received a last-known estimated position of the Wi-Fi access point;

determining that at least a first set of the Wi-Fi access points for which signals were received have moved from their corresponding last-known estimated positions based on the last-known estimated positions for at least a second set of Wi-Fi access points for which signals were received;

estimating, based on the number of Wi-Fi access points of the first set and the number of Wi-Fi access points of the second set, the likelihood of the Wi-Fi enabled device being located within an estimated geographical area; and

displaying on a display device information based on the estimated likelihood of the Wi-Fi enabled device being located within the estimated geographical area.

22. The method of claim 21, further comprising:

estimating, based on the number of Wi-Fi access points of the first set and the number of Wi-Fi access points of the second set, a plurality of likelihoods of the Wi-Fi enabled device being located within a corresponding plurality of estimated geographical areas; and

displaying on a display device information based on the plurality of estimated likelihoods of the Wi-Fi enabled device being located within the corresponding plurality of estimated geographical areas.

23. The method of claim 21, wherein each of the Wi-Fi access points of the first set has a last-known estimated position that is within a first predetermined distance from the last-known estimated positions of other Wi-Fi access points of the first set, and each of the Wi-Fi access points of the second set has a last-known estimated position that is within a second predetermined distance from the last-known estimated positions of the other Wi-Fi access points of the second set.

24. The method of claim 23, wherein the Wi-Fi access points of the first set are different Wi-Fi access points from the Wi-Fi access points of the second set.

25. The method of claim 21, wherein the number of Wi-Fi access points of the first set is less than the number of Wi-Fi access points of the second set.

26. The method of claim 21, wherein estimating the likelihood of the Wi-Fi enabled device being located within the estimated geographical area is further based on information that characterizes the conditional probability that the Wi-Fi access points of the first set have moved from their corresponding last-known position, the conditional probability being based on the number of Wi-Fi access points of the first set and the number of Wi-Fi access points of the second set, and wherein the conditional probability relies on a determination that last-known positions of a first Wi-Fi access point and a second Wi-Fi access point from which signals were received are in conflict.

27. The method of claim 26, wherein the determination that the last-known positions of the first Wi-Fi access point and the second Wi-Fi access point from which signals were

received are in conflict is based on the last known positions of the first Wi-Fi access point and second Wi-Fi access point being separated by more than a threshold distance.

28. The method of claim 26, further comprising determining the conditional probability that at least one of the Wi-Fi access points of the first set has moved from its corresponding last-known position.

29. The method of claim 28, wherein the determining the information that characterizes the conditional probability that at least one of the Wi-Fi access points has moved from its corresponding last-known position comprises:

determining a first collection of Wi-Fi access points that moved from a first geographic area to a second geographic area;

determining a second collection of Wi-Fi access points in the second geographic area, the Wi-Fi access points of the second collection being different Wi-Fi access points than the Wi-Fi access points of the first collection; and

based on the number of Wi-Fi access points of the first collection and the number of access points of the second collection, determining information that characterizes the conditional probability that a Wi-Fi access point has moved from its last-known position based on the number of Wi-Fi access points of the first set and the number of Wi-Fi access points of the second set.

30. The method of claim 29, further comprising:

determining a third collection of Wi-Fi access points, each Wi-Fi access point of the third collection having a corresponding actual position that permits radio signals transmitted from said Wi-Fi access points to be received by a device capable of receiving Wi-Fi radio signals; and

determining that each Wi-Fi access point of a subset of the Wi-Fi access points of the third collection has moved from its corresponding last-known position based on a comparison of the last-known positions corresponding to each of the Wi-Fi access points of the third collection;

the determining information that characterizes the conditional probability that a Wi-Fi access point has moved from its last-known position being further based on the number of Wi-Fi access points of the subset and the number of Wi-Fi access points of the third collection not included in the subset.

31. The method of claim 21, further comprising:

consulting the reference database to determine for at least a first Wi-Fi access points from which signals were received a last-known estimated position of the first Wi-Fi access point and time information associated with the last-known position for describing the age of the last-known position relative to other information in the reference database; and

the estimating the likelihood of the Wi-Fi enabled device being located within an estimated geographical area being further based on the last-known position and associated time information of the first Wi-Fi access point.

32. The method of claim 31, wherein the estimating the likelihood of the Wi-Fi enabled device being located within the estimated geographical area is further based on information that characterizes the conditional probability that a Wi-Fi access point has moved from its corresponding last-known position, the conditional probability being based on the relative age of said last-known position, and wherein the conditional probability relies on a determination that last-known positions of a first Wi-Fi access point and a second Wi-Fi access point from which signals were received are in conflict.

33. The method of claim 32, wherein the determination that the last-known positions of the first Wi-Fi access point and the second Wi-Fi access point from which signals were received are in conflict is based on the last known positions of the first Wi-Fi access point and second Wi-Fi access point being separated by more than a threshold distance.

34. The method of claim 32, further comprising determining the information that characterizes the conditional probability that a Wi-Fi access point has moved from its corresponding last-known position.

35. The method of claim 34, wherein the determining the information that characterizes the conditional probability that a Wi-Fi access point has moved from its corresponding last-known position based on the relative age of said last-known position comprises:

determining a set of Wi-Fi access points, each Wi-Fi access point of the set located at a first geographic position for the corresponding Wi-Fi access point at a first point in time, and having moved to a second geographic position for the corresponding Wi-Fi access point at a second point in time; and

based on the set of Wi-Fi access points that moved and based on the amount of time between the first and second points in time, determining information that characterizes the conditional probability that a Wi-Fi access point has moved from its last-known position based on the relative age of said last-known position.

36. The method of claim 31, further comprising:

estimating, based on the last-known position and associated time information, a plurality of likelihoods of the Wi-Fi enabled device being located within a corresponding plurality of estimated geographical areas; and

displaying on a display device information based on the plurality of estimated likelihoods of the Wi-Fi enabled device being located within the corresponding plurality of estimated geographical areas.

37. The method of claim 21, further comprising:

consulting an historical dataset to determine for at least one of the plurality of Wi-Fi access points from which signals were received information describing past relocations for the at least one Wi-Fi access point;

wherein estimating the likelihood of the Wi-Fi enabled device being located within the estimated geographical area is further based on the information describing past relocations for the at least one Wi-Fi access point.

38. The method of claim 37, wherein the information describing past relocations for the at least one Wi-Fi access point includes an average movement frequency for the at least one Wi-Fi access point.

39. The method of claim 37, wherein the information describing past relocations for the at least one Wi-Fi access point includes an aggregate average movement frequency based on a collection of movement data for a plurality of Wi-Fi access points.

40. The method of claim 21, further comprising:

extracting information from the signals received that identifies at least one Wi-Fi access point from which signals were received;

consulting a reference database to determine for the at least one Wi-Fi access point from which signals were received a set of information identifying a corresponding set of Wi-Fi access points from which signals are expected to be received when the Wi-Fi enabled device receives signals from the at least one Wi-Fi access point of the plurality; and

the estimating the likelihood of the Wi-Fi enabled device being located within an estimated geographical area being further based on a comparison of the information identifying Wi-Fi access points from which signals were received and the set of information identifying the corresponding set of Wi-Fi access points from which signals are expected to be received.

41. The method of claim 40, wherein the estimating the likelihood of the Wi-Fi enabled device being located within an estimated geographical area further comprises comparing a first number of Wi-Fi access points from which signals were received and for which information was found in the reference database to a second number of Wi-Fi access points from which signals are expected to be received when the Wi-Fi enabled device receives signals from the at least one Wi-Fi access point of the plurality.

42. The method of claim 40, wherein the estimating the likelihood of the Wi-Fi enabled device being located within an estimated geographical area further comprises

comparing a first number of Wi-Fi access points from which signals were received and for which no identification information was found in the reference database to a second number of Wi-Fi access points from which signals are expected to be received when the Wi-Fi enabled device receives signals from the at least one Wi-Fi access point of the plurality.

43. The method of claim 40, wherein the Wi-Fi access points from which signals are expected to be received when the Wi-Fi enabled device receives signals from the at least one Wi-Fi access point of the plurality are identified based on a history of one or more Wi-Fi enabled devices receiving signals from the expected Wi-Fi access points when receiving signals from the at least one Wi-Fi access point.

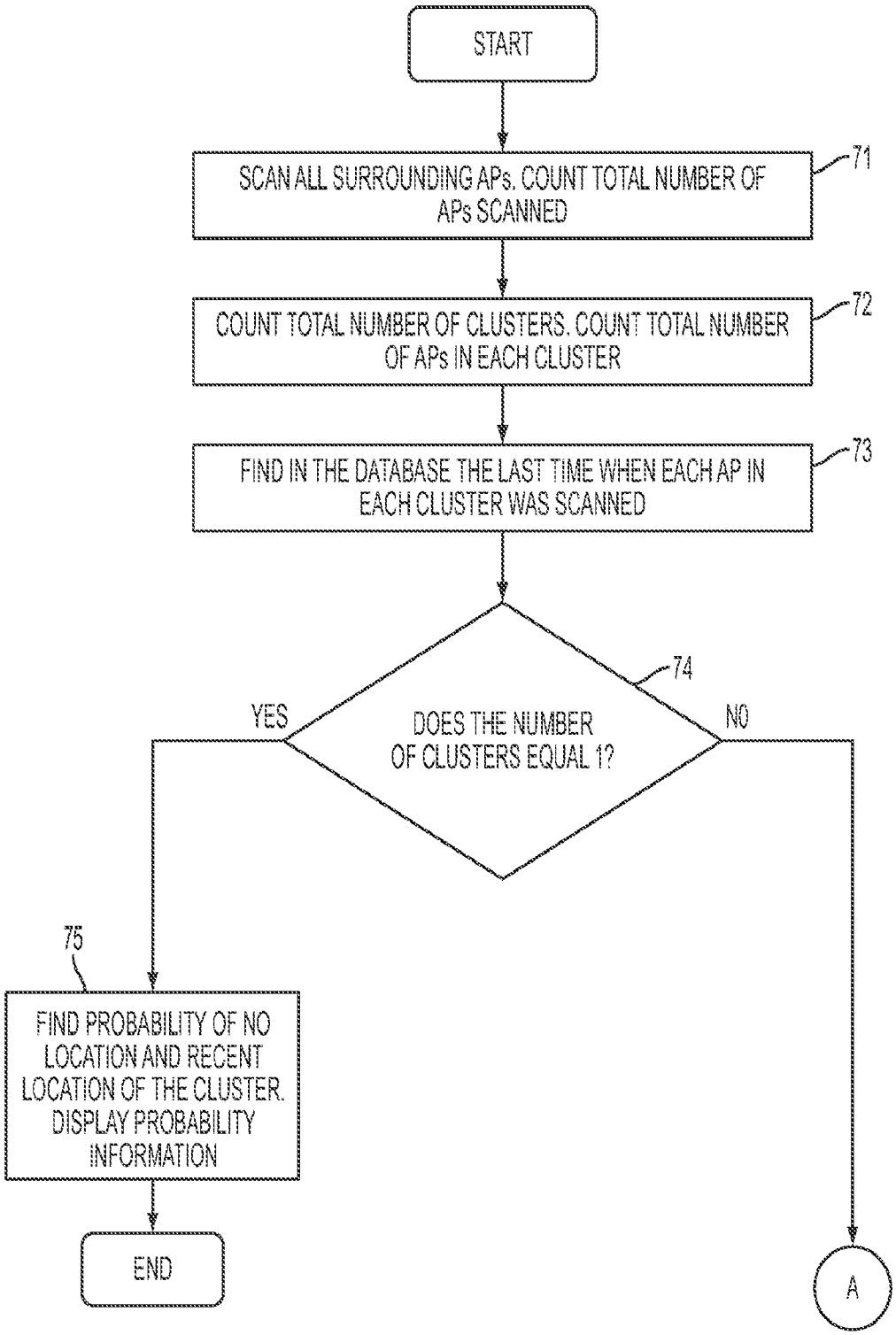
START
SCAN ALL SURROUNDING APs. COUNT TOTAL NUMBER OF APs SCANNED
71
COUNT TOTAL NUMBER OF CLUSTERS. COUNT TOTAL NUMBER OF APs IN EACH CLUSTER
72
FIND IN THE DATABASE THE LAST TIME WHEN EACH AP IN EACH CLUSTER WAS SCANNED
73
DOES THE NUMBER OF CLUSTERS EQUAL 1?
74
YES
NO
75
FIND PROBABILITY OF NO LOCATION AND RECENT LOCATION OF THE CLUSTER. DISPLAY PROBABILITY INFORMATION
END
A

FIG. 1A

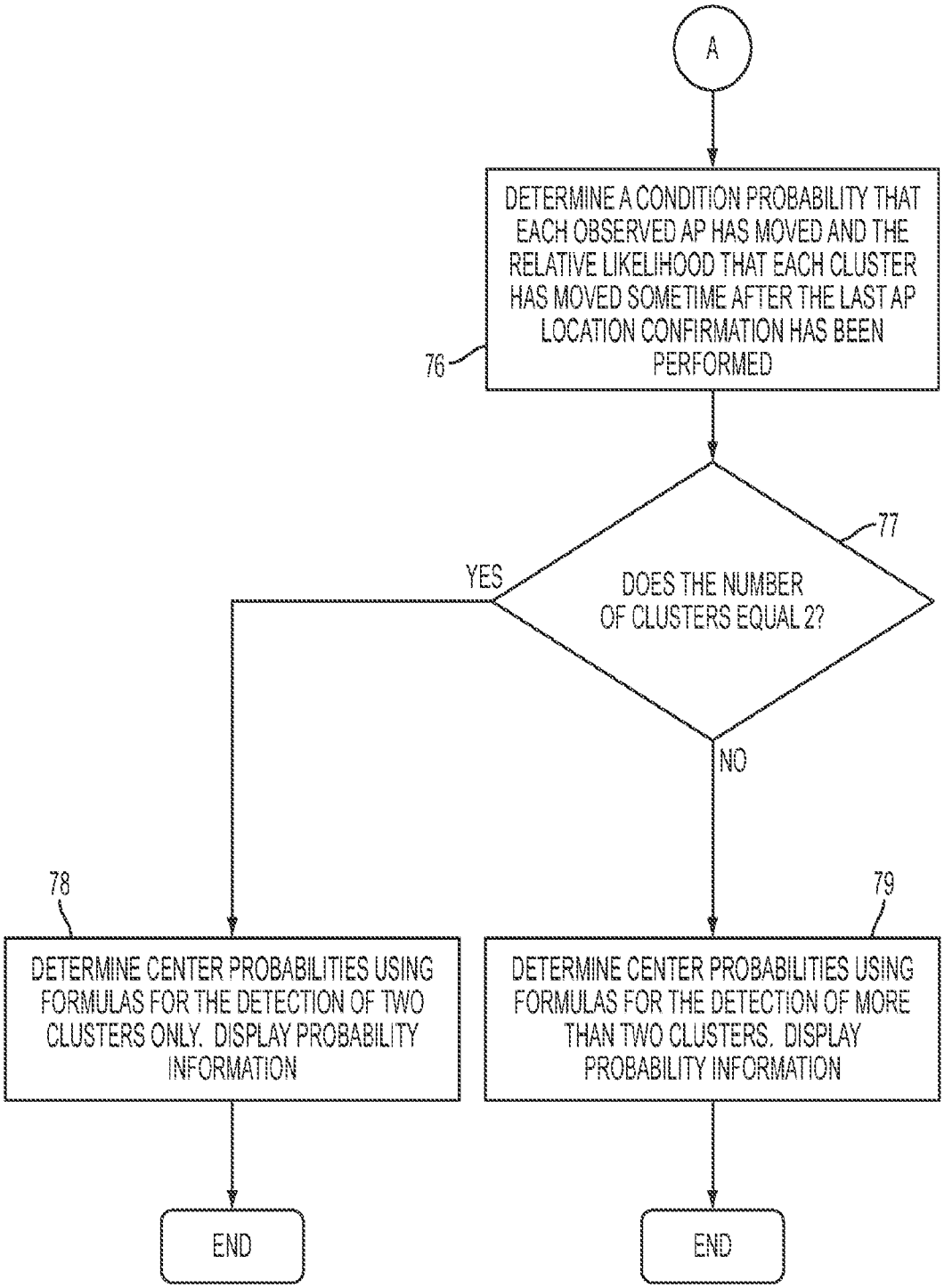
A
DETERMINE A CONDITION PROBABILITY THAT EACH OBSERVED AP HAS MOVED AND THE RELATIVE LIKELIHOOD THAT EACH CLUSTER HAS MOVED SOMETIME AFTER THE LAST AP LOCATION CONFIRMATION HAS BEEN PERFORMED
76
77
DOES THE NUMBER OF CLUSTERS EQUAL 2?
YES
NO
78
DETERMINE CENTER PROBABILITIES USING FORMULAS FOR THE DETECTION OF TWO CLUSTERS ONLY. DISPLAY PROBABILITY INFORMATION
79
DETERMINE CENTER PROBABILITIES USING FORMULAS FOR THE DETECTION OF MORE THAN TWO CLUSTERS. DISPLAY PROBABILITY INFORMATION
END
END

FIG. 1B

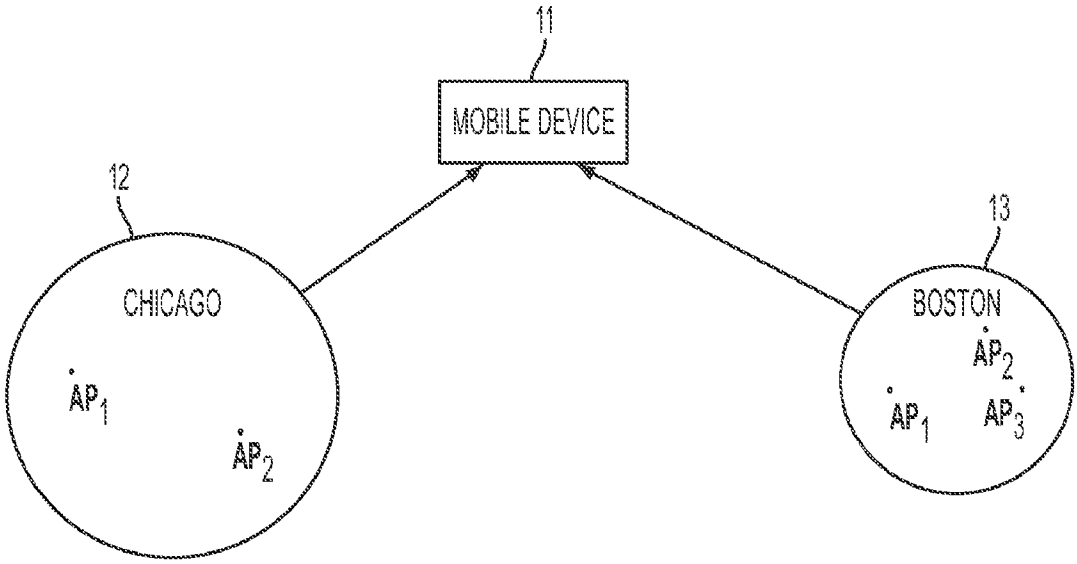

FIG. 2

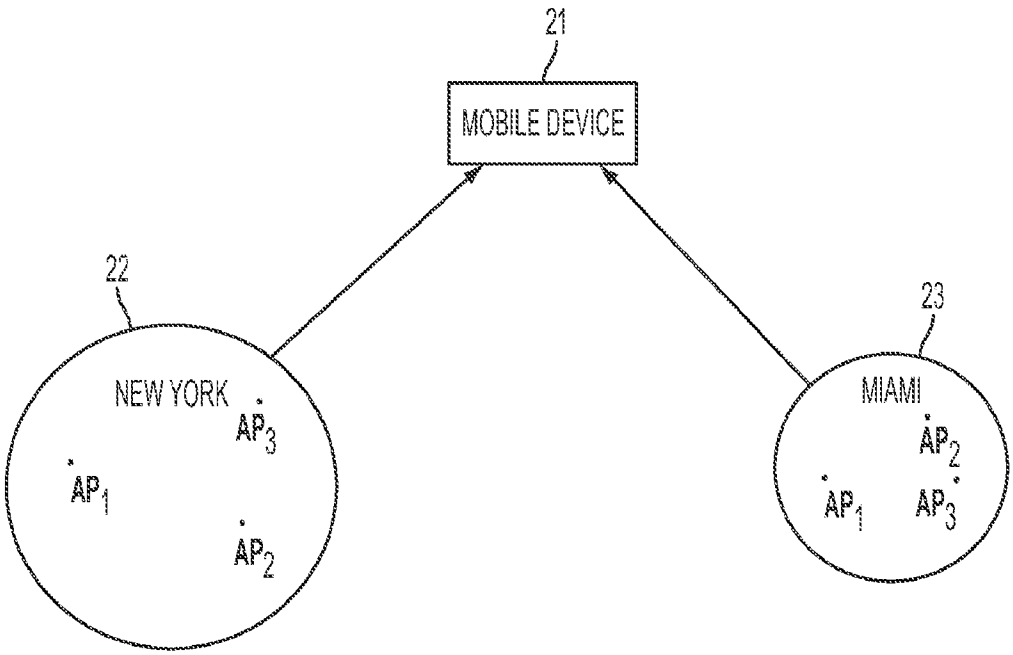

FIG. 3

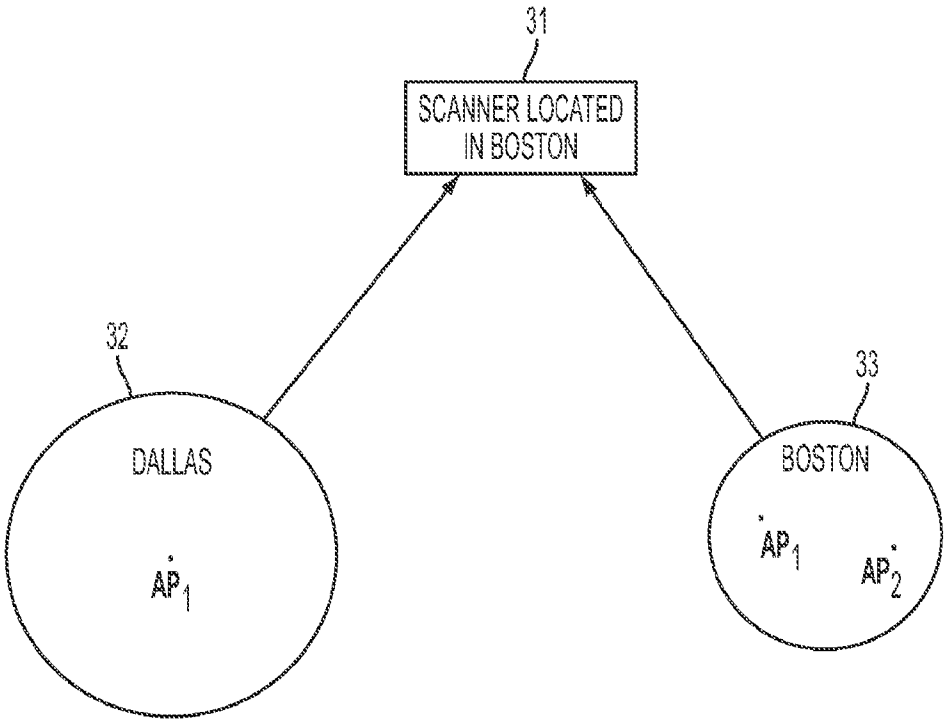

FIG. 4

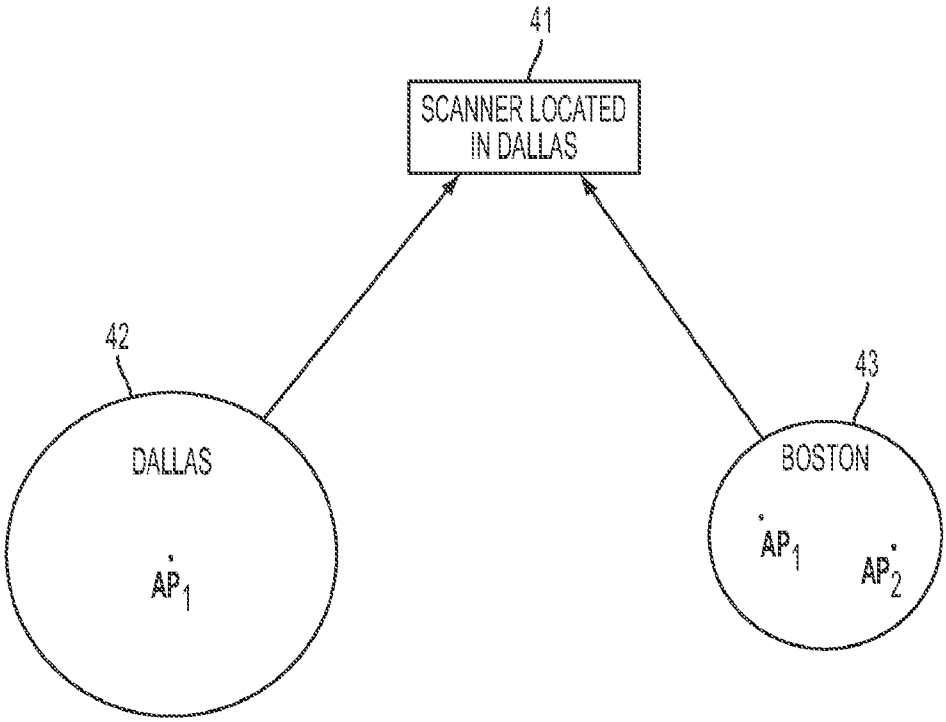

FIG. 5

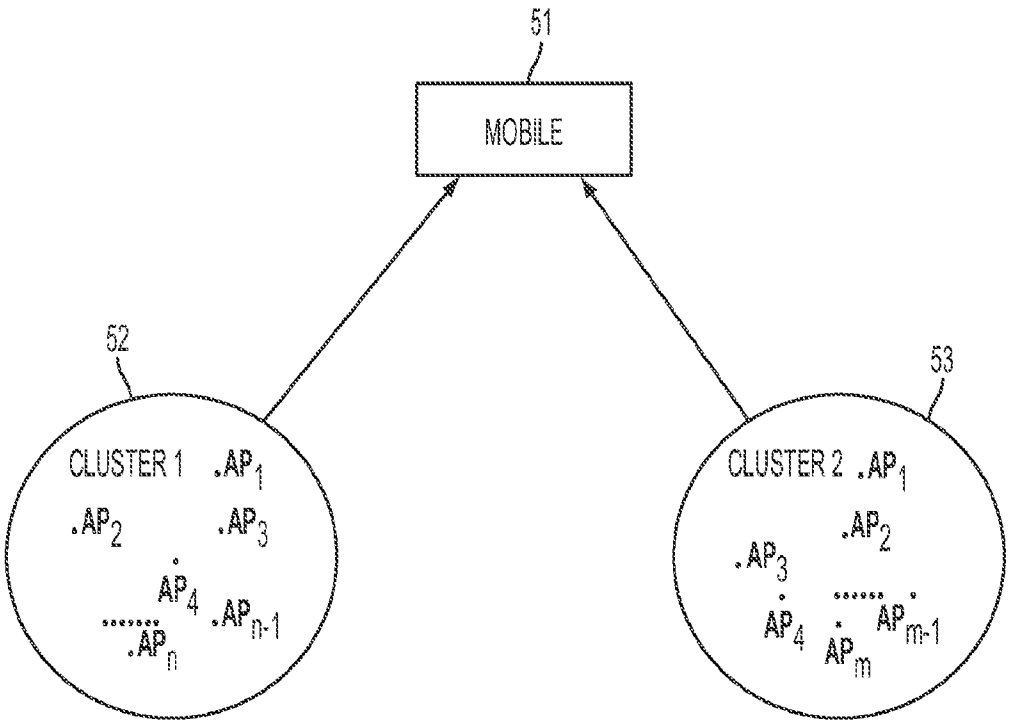
51
MOBILE
52
CLUSTER 1
.AP1
.AP2
.AP3
.AP4
.APn-1
.APn
53
CLUSTER 2
.AP1
.AP2
.AP3
.AP4
.APm
.APm-1

FIG. 6

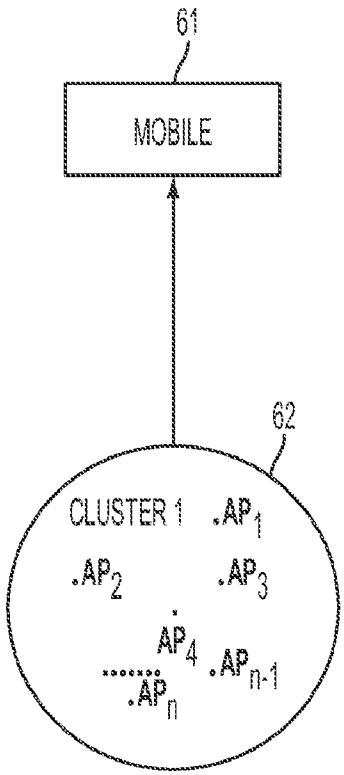

FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2011/029379

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - H04W 24/00 (2011.01)
USPC - 455/456.1
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)
IPC(8) - H04M 11/04; H04W 04/00, 24/00, 72/00 (2011.01)
USPC - 370/338; 455/404.2, 452.1, 456.1, 457

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)
MicroPatent, GooglePatent, Google

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2008/0132170 A1 (ALIZADEH-SHABDIZ et al) 05 June 2008 (05.06.2008) entire document	1-43
Y	Zhou, " Wireless Indoor Tracking System (WITS), Tagungsband zum dolT Software-Forschungtag, 13 July 2006, Mannheim, Germany, , [retrieved on 2011-05-11, Retrieved from the Internet <URLhttp://www.ks.uni-freiburg.de/assist/rui/index.php?page=publications>, entire document	1-43
Y	US 6,674,403 B2 (GRAY et al) 06 January 2004 (06.01.2004) entire document	1-43
A	Krumm et al., "LOCADIO: Inferring motion and location from Wi-Fi signal strengths", First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services, Mobiquitous, 2004, [retrieved on 2011-05-11], Retrieved from the Internet <URL http://scholar.google.com/scholar?hl=en&lr=&q=related:capHcK2acsoJ:scholar.google.com/&um=1&ie=UTF-8&ei=twXLTcL1A8PJgQelruHsBQ&sa=X&oi=science_links&ct=sl-related&resnum=1&ved=0CBkQzwIwAA>, entire document	1-43
A	Curran et al., "Pinpointing Users with Location Estimation Techniques and Wi-Fi Hotspot Technology", International Journal of Network Management, 2008 [retrieved on 2011-05-11], Retrieved from the Internet <URL http://www.google.com/#sclient=psy&hl=en&safe=off&site=&source=hp&q=Pinpointing+users+with+location+estimation+techniques+and+Wi-Fi+hotspot+technology&aq=&aqi=&aql=&oq=&pbx=1&fp=5068b48774b6de5d>, entire document	1-43

☒ Further documents are listed in the continuation of Box C. ☐

* 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
11 May 2011	01 JUN 2011
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2011/029379

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2009/0042557 A1 (VARDI et al) 12 February 2009 (12.02.2009) entire document	1-43
A	US 2008/0176583 A1 (BRACHE T et al) 24 July 2008 (12.07.2008) entire document	1-43