METHOD OF AND SYSTEMS FOR PRIVACY PRESERVING MOBILE DEMOGRAPHIC MEASUREMENT OF INDIVIDUALS, GROUPS AND LOCATIONS OVER TIME AND SPACE

Inventors: Andrew J. Broadstone, Watertown, MA (US); Russel K. Jones, Medfield, MA (US); Edward J. Morgan, Needham, MA (US); Farshid Alizadeh-Shabdiz, Wayland, MA (US); Nicolas Brachet, Chestnut Hill, MA (US)

Appl. No.: 13/530,566
Filed: Jun. 22, 2012

Related U.S. Application Data
Provisional application No. 61/499,975, filed on Jun. 22, 2011.

Publication Classification
Int. Cl. G06F 21/24 (2006.01)

Abstract

Disclosed are techniques for privacy preserving mobile demographic measurement of individuals, groups, and locations over time and space. A method of estimating demographic information associated with a user of a mobile device and/or a location while preserving the privacy of the user based at least in part on a location estimate of the mobile device of the user includes receiving an estimated geographical location of the mobile device of the user and receiving a time at which the mobile device was at the estimated geographical location. The method includes assigning substitute identifiers for the geographical location and the time at which the mobile device was at the estimated geographical location. The method includes associating the geographical areas substitute identifiers with demographic information and estimating demographic information associated with the user of the mobile device based on the substitute identifiers and based on the demographic information associated with substitute identifiers.
Fig. 3
Fig. 4

Demographic Server 220

221
Device Demographics Profile Logging

111
Device Demographic Profile Logs

Time substitute, location substitute, device demographic profile
Fig. 5

Demographic Server 220

112

Location Demographics Profile Retrieval 222

501

Retrieve Location Demographic Profile
timestamp, latitude, longitude

502

Demographics profile available for time/location?

Yes

504

Return Location Demographic Profile

No

503

Return Null Demographics
Fig. 6
Fig. 9
Fig. 11A
Fig. 11B
Fig. 13A
Location Demographic Profile Output Record 1401

Geometry: ABAF007
Hour of week: 57
Confidence: 85
Attributes:

<table>
<thead>
<tr>
<th>AttributeName</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>0.375</td>
</tr>
<tr>
<td>High School</td>
<td>0.525</td>
</tr>
<tr>
<td>Bachelors</td>
<td>0.1</td>
</tr>
<tr>
<td>PhD</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 14
METHOD OF AND SYSTEMS FOR PRIVACY PRESERVING MOBILE DEMOGRAPHIC MEASUREMENT OF INDIVIDUALS, GROUPS AND LOCATIONS OVER TIME AND SPACE

CROSS-REFERENCE TO RELATED APPLICATIONS


BACKGROUND OF THE INVENTION

[0001] For the purposes of this disclosure, the term “demographic profile” refers to a set of attributes describing the user of a device (the device demographic profile, or DDP) or the group of people that visit a particular location (the location demographic profile, or LDP). This set of attributes may include, but is not limited to, age, gender, ethnic background, income, years of education, as well as behavioral descriptors, such as “frequent traveler” or “retail shopper.”

[0012] Demographic profile data for where people live (the LDP) has been widely available for decades. Direct marketing firms and others have compiled data from public and private sources to build profiles of neighborhoods. Public data sources include the U.S. Census, public record filings about home purchases, and records on public infrastructure such as water systems. Private sources include shipping and purchase records, magazine and newspaper subscriptions, and voluntary surveys.

[0013] LDPs are available for nearly every residential block in the U.S. Some vendors of demographic profile information further categorize neighborhoods according to the mix of different behavioral types based on the mix of these attributes—labeling combinations of attributes with names like “urban achievers.” Together these data sources have given marketers a way to target households for various offers by direct mail and telephone, providing vendors an easy way to send their catalogs only to highly educated parents under 45 who live in the suburbs, for example.

[0014] While much is known about the LDP of a block of homes, which can be characterized as the static or unchanging demographic profile of a place, it has been discovered that little is known about where the people living there go during the day, which can be characterized as the dynamic or changing LDP. Some companies have surveyed users to determine this information, and extrapolated from these limited samples. This can give some broad general understanding of where some people work during the day, but cannot resolve how the LDP changes during the day, or over the course of a week, for example. Commercial and urban areas in particular, where the demographic mix changes significantly with time of day or day of week, cannot be described by current demographic profile data sources. Companies evaluating these areas for retail expansion or outdoor advertising have very little information on which to base large financial commitments.

[0015] User demographic profiles, not associated with one’s home, are frequently collected using purchasing and shipping records, web browsing histories, and other records related to computer usage. An example of this is the user profile data collected in web browser “cookies”, which are special files many web sites use to store information about users based on pages they have visited. These files are stored on the user’s computer and submitted to the site with every page request or view. Online user demographic profiles are generally computed without considering their location or their current activity.

BRIEF SUMMARY OF THE INVENTION

[0016] In one aspect, the invention features methods of and systems for privacy preserving mobile demographic measurement of individuals, groups and locations over time and space.

[0017] In another aspect of the invention, a method of estimating demographic information associated with a user of a mobile device while preserving the privacy of the user based
at least in part on a location estimate of the mobile device of the user includes receiving an estimated geographical location of the mobile device of the user and receiving a time at which the mobile device was at the estimated geographical location. The method also includes providing a set of substitute identifiers for a corresponding set of at least one geographical area, assigning one of the set of substitute identifiers for the geographical area corresponding to the geographical location of the mobile device, and assigning a substitute identifier for the time at which the mobile device was at the estimated geographical location. The method further includes providing an association between the substitute identifiers for geographical areas and demographic information corresponding to the substituted geographical area and estimating demographic information associated with the user of the mobile device based on the assigned substitute identifiers and based on the demographic information associated with the provided set of substitute identifiers.

[0018] In a further aspect of the invention, the method further includes estimating the received geographical location associated with the mobile device of the user.

[0019] In yet another aspect of the invention, the method includes recording the estimated demographic information associated with the user of the mobile device in a device demographic information log. Optionally, the method also includes estimating demographic information associated with at least one geographical area of the set of geographical areas based on the demographic information recorded in the device demographic information log. The device demographic information log contains a plurality of records of estimated demographic information associated with a plurality of mobile devices.

[0020] In another aspect of the invention, the method includes sending the estimated demographic information to the mobile device.

[0021] In still another aspect of the invention, the substitute identifier for the geographical area corresponding to the geographical location of the mobile device that is assigned identifies a particular set of demographic information.

[0022] In an aspect of the invention, the substitute identifier for the geographical area corresponding to the geographical location of the mobile device that is assigned is reduced in specificity relative to the estimated geographical location of the mobile device.

[0023] In another aspect of the invention, the substitute identifier for the time at which the mobile device was at the estimated geographical location is a measure of time that is reduced in specificity relative to the time that was received. Optionally, the substitute identifier for the time at which the mobile device was at the estimated geographic location is a representation of time lacking date information, a time range, and/or an hour of a week designation.

[0024] In still another aspect of the invention, the estimating demographic information associated with the user of the mobile device is performed on a separate computer system from a computer system performing any one or a subset of the other steps. Optionally, the separate computer system, relative to the computer system performing any one or a subset of the other steps, is maintained in a separate network, maintained in a separate building, and/or maintained by a separate operational entity.

[0025] In another aspect of the invention, a method of estimating demographic information associated with a geographical area and a time period based on demographic information associated with users of mobile devices within the geographical area includes providing a set of geographical areas and providing a set of time periods. The method also includes receiving an estimated geographical location of a mobile device of the user, receiving a mobile device identifier that is associated with the mobile device, and receiving a time at which the mobile device was at the estimated geographical location. The method further includes determining the geographical area of the set in which the estimated geographical location occurs, determining the time period of the set in which the time at which the mobile device was at the estimated geographical location occurs, and retrieving information representative of demographic information associated with the user of the mobile device based on the mobile device identifier. The method also estimates demographic information associated with the determined geographical area during the determined time period based on the retrieved information representative of demographic information associated with the user of the mobile device.

[0026] In a further aspect of the invention, the method also includes estimating the received geographical location of the mobile device of the user.

[0027] In still another aspect of the invention, the method also includes providing a set of initial demographic information associated with the determined geographical area. The estimating demographic information associated with the determined geographical area is further based on the initial demographic information. Optionally, the method also includes, subsequent to the estimating demographic information associated with the determined geographical area, adjusting the initial demographic information based on the estimated demographic information. Also optionally, the initial demographic information is based on governmental census information, public record information, shipping and purchase records, magazine and newspaper subscriptions, voluntary surveys, and/or records of social media activity.

[0028] In yet another aspect of the invention, the method also includes sending the estimated demographic information to the mobile device.

[0029] In another aspect of the invention, the geographical areas of the set are reduced in specificity relative to the estimated geographical location of the mobile device.

[0030] In a further aspect of the invention, the time periods of the set are reduced in specificity relative to the time that was received. Optionally, the time periods of the set are a representation of time lacking date information, a time range, and/or an hour of a week designation.

[0031] In still another aspect of the invention, the estimating demographic information associated with the determined geographical area during the determined time period is performed on a separate computer system from a computer system performing any one or a subset of the other steps. Optionally, the separate computer system, relative to the computer system performing any one or a subset of the other steps, is maintained in a separate network, maintained in a separate building, and/or maintained by a separate operational entity.

[0032] In another aspect of the invention, the method also includes performing selected steps a plurality of times for different mobile devices of different users and recording sets of information for the determined geographical areas, determined time periods, and retrieved information representative of demographics information associated with the users of the mobile devices in a device demographic information log. The estimating demographic information associated with the
determined geographical area during the determined time period is further based on a plurality of the sets of information in the device demographic information log.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

[0033] FIG. 1 shows a general flow chart of the process of embodiments of the present invention;

[0034] FIG. 2 shows one embodiment of a deployment of the present invention;

[0035] FIG. 3 shows one embodiment of a detailed example of demographic profile retrieval as deployed on a mobile computing device as per FIG. 2;

[0036] FIG. 4 shows a detailed view of one implementation of data storage on the server for embodiments of the present invention;

[0037] FIG. 5 shows a flow chart of the process of retrieving the demographic attributes of a given location for a specific time;

[0038] FIG. 6 shows a flow chart of one embodiment of the process to compute the Location Demographic Profiles;

[0039] FIG. 7 shows a flow chart of one embodiment of the process to compute the Device Demographic Profiles as deployed according to FIG. 2;

[0040] FIG. 8 shows a second embodiment of a deployment of the present invention in which the Mobile Computing Device requires little to no modification;

[0041] FIG. 9 shows a second embodiment of a detailed example of demographic profile retrieval as deployed on a remote location server as per FIG. 8;

[0042] FIG. 10 shows a second embodiment of a flow chart of the process to compute the Device Demographic Profiles as deployed according to FIG. 8;

[0043] FIGS. 11A and 11B depict an illustrative embodiment of the invention showing aspects of the privacy preserving mechanisms. FIG. 13A represents the Demographic Public System and FIG. 13B represents the Demographic Private System.

[0044] FIG. 12 shows an embodiment of the process to convert demographic data into privacy preserving components;

[0045] FIGS. 13A and 13B depict a modified embodiment of FIG. 11 showing additional detail with respect to the data elements used for privacy preserving purposes. FIG. 13A represents the Demographic Public System and FIG. 13B represents the Demographic Private System.

[0046] FIG. 14 depicts an embodiment of a partial output of the Location Demographic Profile for an illustrative geographic geometry.

DETAILED DESCRIPTION OF THE EMBODIMENTS

[0047] As used herein, the term “demographics” means statistical data describing a population. Demographics or demographic data includes, but is not limited to average age, income, education, ethnicity, gender, disabilities, mobility, educational attainment, home ownership, employment status, etc. It also may include psychographic data such as the values, attitudes, moods and interests of a population. It also may include activity, venue, and points of interest details that further describe a population or a location. The data may be time dependent with trending attributes.

[0048] As used herein, the term “demographic attribute” indicates a specific measure of demographics, for example, gender. Each attribute may have many “buckets” which represent the possible categories for a given attribute. For example, the attribute gender could have two buckets, male and female. Each of these buckets could indicate the ratio or percent of the population that falls within the definitions for each bucket.

[0049] As used herein, the term “Location Demographic Profile” or “LDP” means the composite or aggregate statistics of a population based on a specific location or geographical area.

[0050] As used herein, the term “Device Demographic Profile” or “DDP” means the composite or aggregate statistics of a Device and, by proxy the device owner or operator(s), based on a set of inputs including the aggregate set of LDPs in which the device has been observed.

[0051] As used herein, the terms “Canonical Week” and “Hour of the Week” refer to the 168 hours of a week starting with hour 0 at Sunday 12 AM and continuing through Saturday 11 PM.

[0052] Mobile phones equipped with GPS and other location technologies provide a new opportunity for discovering user information. Location trace information can be used to augment observations about what web pages a user visits or what mobile applications a user employs with respect to time and place. Behavioral and demographic attributes of a user can be inferred based on where he or she goes if a complete map trace is stored. Venues and Points of Interest (POIs) and their demographic tendencies can also be used to contribute to understanding an individual’s proclivities. In addition, geotagged social activity (for example, Twitter posts) can provide more social, demographic, activity, and behavior information about a place in time.

[0053] Individual profiling can be used to present relevant advertising and content, both physical and electronic, to individuals. Location profiling can be used to make intelligent decisions about physical places by knowing what kind of people visit a place throughout the day, for example decisions about the best place for retail expansion or outdoor advertising. Such profiling, however, can raise general privacy concerns for the individual, since the location of the individual can tell a lot about the person.

[0054] Tracking of individuals, even without knowledge of their Personally Identifiable Information (e.g. Name, Social Security Number, Address, etc.), presents potential risks to the privacy of individuals. In particular, location information about individuals in which time sequence is maintained and which has high enough spatial precision, can uniquely identify an individual. Linking these unique traces to a specific named person is quite achievable in today’s environment of “Big Data”, data that is easily accessed and searched.

[0055] Individual location traces provide enough information to identify the home and work place of individuals. It has been shown in Golle & Partridge, On the anonymity of home, work location pair (Golle, P.; Partridge, K. On the anonymity of home, work location pairs. Proceedings of the 7th International Conference on Pervasive Computing, 2009 May 11-14; Nara, Japan. Berlin: Springer; 2009; LNCS 5538: 390-397), that having this information, even at the spatial resolution of a Census Block, uniquely identifies individuals.

[0056] In order to preserve user privacy, a vendor may choose not to collect any location trace information or associate a unique identifier with any individual locations. How-
ever, this makes it very difficult to associate any kind of description or category with a device, since there is no history whatsoever stored about it.

[0057] Embodiments of the present invention disclose methods of and systems for aggregating location and demographic information of individual device users and locations, thereby enabling the various use cases associated with this information, while also preserving the privacy of the individual. Embodiments of the present invention provide a set of methods to preserve the privacy of individuals while computing a Device Demographic Profile based on the Location Demographic Profiles of locations visited by that device. In particular the methods disclosed provide techniques for computing the DDP without storing a location trace of an individual.

[0058] Further, dynamic LDPs can then be determined using a collection of DDPs that visit a location at a particular time of day. These dynamic LDPs can be computed for various scales of time and space.

[0059] In broad terms, embodiments of the present invention include a method and systems to build and maintain demographic estimates of mobile device owners while preserving the anonymity and privacy of the individual. This information is then used to establish dynamic or time-varying demographic information related to a location. Through the use of mobile computing devices in concert with location services, device and location demographic profiles can be computed in a time varying manner without compromising individual privacy.

[0060] Embodiments of the invention determine a Device Demographic Profile based on where that device goes—the interaction with Location Demographic Profiles and potentially other DDPs the device comes near physically—without storing locations, or any trace that would allow discovery of where the device actually went before or after the fact. The only thing stored is the set of demographic attributes for the device (the DDP) and for the location (LDP). No latitude or longitude or description of a location that could uniquely identify a particular point on a map is stored in concert with a unique ID.

[0061] In the first embodiment, the DDP is determined based on the location a device visits at certain times of day. For example, when a phone is used to query location in the evening or at night, for the purpose of looking at a map, or posting a message to a social network, or any other purpose, a simple technique could assume that this is the home demographic of the user, for example. (More sophisticated technique can be created which use statistical measures of time and space to determine demographic clusters and reason as to which of these clusters represent the home demographic profile.) The base Location Demographic Profile for this “home zone” or a set of LDPs chosen for other significance can then be combined to create a DDP. Other LDPs the device encounters can be used to adjust the DDP over time. Embodiments of the present invention determine a DDP without a stored location history thus preserving user anonymity.

[0062] The present illustrative embodiment uses at least three inputs plus an optional additional input. These include the current location as a latitude and longitude point with error probability. The second necessary input includes the available LDPs for the currently or recently visited locations (which may be based on data from the U.S. Census, for example). Third, the current local time, where local is defined relative to the current location. Lastly, the existing DDP as previously computed is an optional input.

[0063] LDPs are initially populated with currently available, static or non-moving place-based demographic data, for example demographic data available from the U.S. Census, but also from other data sources. They may also be extrapolated from neighboring LDPs if not enough data is available for a particular location from other sources. An extrapolation between two areas can be performed in much the same way that a linear interpolation can be performed between two points on a line, taking into account the 2-dimensional nature of the area. The present embodiment adjusts LDPs over time based on the DDPs that visit that location. A feedback mechanism is established that adjusts LDPs and LDPs over time, improving their quality. Individual LDPs are stored for each time of day at a particular place. Additional distinct LDPs can be created and stored which represent various classes or periods of time. This could include such classifications as weekends, morning, lunchtime, afternoon, happy hour, evening, witching hour, etc., and can be defined by any arbitrary block or blocks of time.

[0064] The methods that contribute to the privacy preserving nature of the present invention include A) reducing the specificity of information, B) converting from spatial dimensions to demographic dimensions, C) separating operational domains, D) implementing cryptographic isolation between operational domains, and E) applying probabilistic techniques to further isolate information. These methods contribute to the privacy preserving nature while maintaining the ability to compute DDPs and LDPs while providing the facility to continually add new sources of base demographic or ‘location’ information.

[0065] In one embodiment, the specificity of certain information is reduced. This information includes exchanging the hour of the week in place of the specific local time. This reduces the ability to time sequence a set of observations. This substitution is carried out throughout the entirety of the process.

[0066] In certain embodiments the conversion from spatial dimensions (e.g. latitude and longitude) to demographic dimensions (e.g. age, gender, etc.) occurs at the initial observation. Converting the information early in the process allows the location to never be stored. In other embodiments, the conversion occurs at a later time and before storage of the information.

[0067] In one embodiment, the system is comprised of at least two operational domains. The separation of the operational control provides the ability to reduce information knowledge within each domain. Thus, information that, for example, could link a device with a hashed Device ID is only known in one domain, while information that could link a hashed Device ID to a demographic Profile is only known in the other domain.

[0068] Additionally, the separation of operational domains allows the use of cryptographic techniques to secure private information within one domain from being shared with the other domain. In one embodiment this can be used to create opaque strings which contain very little information for any entity not in possession of the key.

[0069] The use of probabilistic techniques provides another method to increase the privacy of individuals. In one embodiment, probabilistic techniques allow the identification of uniqueness without the ability to determine which of the total set is observed. This technique can be applied, for example, in
counting unique devices that appear within a specific spatial area in a specific amount of time without compromising the privacy of an individual.

[0070] Referring now to embodiments of the invention in more detail, FIG. 1 provides an illustrative example of a dynamic demographic information flow 100. It depicts the overall flow of processing and data flow for an embodiment. A software application (generally running on a mobile computing device) requests the location of the device 101. This request triggers the retrieval of the current Device Demographic Profile (DDP) 102. It then computes the location of the device through any of a number of means (e.g. GPS or Wi-Fi positioning) 102. This information including the timestamp, latitude, longitude and DDP, is logged to the Device Demographic Profile Logs 111 (step 104). In order to ensure privacy of the device and subsequently to the end user of the device, no device identifying information is stored in 111.

[0071] Continuing in FIG. 1, the current Location Demographic Profile (LDP) is retrieved 105 from the Location Demographic Profiles database 112. This information is then logged in the Location Demographic Profile Log 113 (step 106). Information stored in this database includes the timestamp of the request, the location latitude and longitude as well as the current LDP for that location. The location and optionally the LDP and DDP are returned to the calling application by 107.

[0072] Also depicted in FIG. 1 are two processes which do not have to execute in-line or synchronously with the control flows within FIG. 1. These include the computation of the new Device Demographic Profiles 109 and the computation of the new Location Demographic Profiles 108. These will be described further in subsequent figures.

[0073] In addition, the above steps can be initiated by an express request for demographic information determination by an application (step 120).

[0074] Continuing to describe embodiments of the invention in more detail, in FIG. 2 is one embodiment 200 of the invention in which there is a Mobile Computing Device 201, which contains Applications 202. Many applications on mobile devices require location information to be useful or to enhance the functionality of the application. Mobile devices provide a Location API 203, which allows applications to retrieve an estimate of the current location of the device, which, in some embodiments uses a Client Location Service 204 on the device. Embodiments of the present invention extend the capability of a mobile device by adding Location Demographic Profile Logging 205, Device Demographics Profile Retrieval 206 and Device Demographics Profile Engine 207. Additional functionality are included on a server are grouped within the Demographic Server 220. Components of 220 include the Device Demographics Profile Logging 221, the Location Demographic Profile Retrieval 222 and the Location Demographics Profile Engine 223. The new components will be described in more detail in subsequent figures. Additionally, the invention, in one embodiment, extends existing Location Servers 210 by augmenting the Server Location Service 211 to make a call to the Demographic Server 220.

[0075] FIG. 3 is one embodiment of the Device Demographic Profile Retrieval 206. This figure depicts the process by which a Device Demographic Profile (DDP) is retrieved from the Device Demographics Profiles database 304. This process is activated, e.g., when a location request is made by an application (step 101 of FIG. 4). Referring to FIG. 3, the request 301 is received and a query is performed to determine if a qualified DDP is available 302. The qualification criteria are dependent on several factors including the number of samples that have been collected, the cluster of these samples over time and similar Location Demographic Profiles, and the stability of the sample set with respect to change over time. For example, if a particular sample set shows a rate of change greater than a particular threshold or measure, it can be deemed to be too dynamic to be relied upon. If a qualified DDP is available, it is returned (steps 303 and 305). In the event that a qualified DDP is not available, a null DDP will be returned 305.

[0076] Continuing with FIG. 4, an embodiment of the process by which information is gathered about the demographics of devices (and by proxy the end user of such devices) that visit different locations is described. In particular, each time a location request is made from a device with a qualified DDP, the DDP along with a timestamp or time substitute (such as the hour of the week) and the geographical location or a location substitute (such as a geographic geometry) of the device is captured in the Device Demographic Profile Logs 111 by the Device Demographics Profile Logging 221 process. To preserve device anonymity, no device or personal identifying information is logged thus eliminating the ability to track or even identify which device made any particular request.

[0077] FIG. 5 describes the process 222 of retrieving a specific Location Demographic Profile (LDP). This information is retrieved in order for the techniques to build a history of demographic information about where the device has been observed. A request to retrieve an LDP is received 501. A query is performed 502 to determine if an appropriate LDP is available for the location of the mobile computing device. The LDP may represent a generalized perspective of the demographic makeup of the location or, for example, it may be specific to the hour of the day and the day of the week. If an appropriate LDP is available, it is returned 504, if an appropriate LDP is not present (for example if no devices with known LDPs have visited the location), one may be extrapolated from neighboring LDPs, or a null LDP may be returned 503.

[0078] FIG. 6 represents a process performed according to embodiments of the present invention. The Location Demographic Profile Engine 223 process flow is depicted and describes a method by which various data sources can be combined to create a set of Location Demographic Profiles (LDPs). Each location can have multiple LDPs based on, for example, each hour in the canonical week, the time of day, the day of week, the day of the year, or the season of the year. These demographics are compiled and computed for each location by the depicted process.

[0079] Continuing to refer to FIG. 6, a process to determine the LDPs of one or more locations is initiated. This initiation can happen in real-time based on some event or it can occur in an off-line mode, asynchronous to the operation of the general flow of the mobile device location and demographic requests. For each location that requires a new determination, the location geometry is retrieved from the Location Geometry database 610 (step 601). The location geometry allows the process to retrieve all relevant information related to that location (where location geometry is simply a portioning of areas in the overall space). This information can include many different and varied data from disparate sources such as Demographic Profile Logs 111 from the present embodiment,
Static Demographic Statistics 605 which may come from sources such as the US Census, Other Psychographic Statistics 606 such as interests, activities and opinions of the population that frequents the location, Land Use Data 607 which can help identify the types of activity present at the location as well as the expected influx/efflux of people to the location, Venue and Check-in data 608 which can provide insight into both the type and frequency of activities in the location, Social Media, News, Crime statistics, etc. 609. These alternate sources can provide meaningful measures of activity, inclinations, safety and many other elements that may influence the LDP of the location.

[0080] Continuing to refer to FIG. 6, once the various records related to the location L are gathered (step 602), this information is then processed to compute the new LDPs for the location 108. One embodiment of the technique combines this information using the following technique.

[0081] First, given sufficient Device Demographic Profile samples within the location, for each hour of each day of the week (e.g. Monday at 8 AM), compute the Mobile Demographic Profile for location L by averaging the DDPs within each of the 168 hourly weeks.

[0082] Next, determine other hourly statistics (e.g. for social media, these values may change over time and will need to be recomputed).

[0083] Next, combine the Mobile Demographic Profile information with the hourly statistics which represent the dynamic demographics.

[0084] Next, summarize all other static attributes (e.g. Census Data) as the static demographics.

[0085] Next, determine R, the mobile to static ratio for the location L. For example, this can be determined by combining Venue (608), Land Use (607), and Demographic (605) data to estimate the ratio of residents to visitors at the location at any given time.

[0086] Next, use the ratio R to determine the influence of the dynamic demographics D on the static demographics S. For the set of attributes A that are mutual in these sets, A*=(R*S+1/R*D). For those attributes that are exclusive to any of the sets, use the values of these attributes as the final values for the LDPs of the location L.

[0087] To complete the discussion of FIG. 6, once the new LDPs have been computed for the location L, they are used to update the Location Demographic Profiles database 112 (step 603).

[0088] The process is repeated while there are additional locations L that need to be computed (step 604).

[0089] FIG. 7 depicts a process by which the Device Demographic Profiles (DDPs) are computed by the Device Demographic Profile Engine 207. In one embodiment of this invention, the computation of the DDPs is executed on the Mobile Computing Device 201. In FIG. 7, an event such as a new LDP or an external trigger begins the process of computing a new DDP for the device. The process begins by gathering all relevant LDPs that have been logged for the device 701. These may be filtered by different criteria, such as age of reading, day of the week, etc. The logs for the device are then combined using a technique to compute the new DDP 109. One embodiment of the technique chooses all LDPs for the device that were registered between the hours of 10 PM and 5 AM local time for that device. A rudimentary implementation could then average the demographic information contained in all of the LDPs to compute an average DDP. This would presume that the device was near the home location of the owner a majority of the time that it was observed during this time and thus would be strongly influenced by the demographics of that location.

[0090] Another embodiment of the technique would cluster similar demographics based on similarity measures and grouped by time. Thus, large clusters could be identified that represent a demographic profile of locations in which the device is often observed. These different clusters could then be categorized and based on various factors (e.g. time of day, land use, etc.) to label these DDPs with tags such as ‘home’ or ‘work’.

[0091] The clustering techniques are novel in that they cluster demographics within the demographic dimension and time dimension. This differs from methods which relied on the spatial attributes (e.g. location) to provide clustering and distance measures. Computations must include a mechanism to determine ‘Demographic Distance’ between any two Demographic Profiles. For example, the distance measure could be the normalized sum of differences in ratios across all buckets in all of the attribute classes. A more sophisticated embodiment computes weighted difference based on the number of buckets contained in each attribute class.

[0092] Completing the discussion of FIG. 7, once the new DDPs have been computed by 109, this information is stored in the Device Demographic Profiles database 110 (step 702).

[0093] FIG. 8, FIG. 9, and FIG. 10 depict an alternate embodiment of the invention in which the Device Demographic Profiles (DDPs) are computed and stored on the Location Server 210 rather than the Mobile Computing Device 201. In these instances FIG. 8 is a modified version of FIG. 2, FIG. 9, 900 is a modified version of FIG. 3, and FIG. 10 is a modified version of FIG. 7. One key distinction between these sets of figures is the inclusion of a hashed ID in the process and data flows that identify a device as distinct from all other devices.

[0094] FIG. 8 illustrates an embodiment of the invention in which there is a Mobile Computing Device 201 which contains Applications 202. Many applications on mobile devices require location information to be useful or to enhance the functionality of the application. Mobile devices provide a Location API 203 which allows applications to retrieve an estimate of the current location of the device. Additionally, the invention, in one embodiment, extends Demographic Server 220 by adding the Location Demographics Profile Logging 205, the Device Demographics Profile Retrieval 806 and Device Demographics Profile Engine 807.

[0095] Continuing with FIG. 9, the modified Device Demographic Profile Retrieval 806 is described in an embodiment deployed on the Demographic Server 220. In addition to the request for a DDP, the calling process must include a hashed ID to identify the particular Device that a DDP is being requested for 901. This same information is subsequently required to retrieve the DDP 903. All other steps remain analogous as those described for Device Demographic Profile Retrieval 206.

[0096] Continuing with FIG. 10, 1000 depicts the Device Demographic Profile Engine 807 is depicted in a second embodiment whereby the process is located on the Demographic Server 220 rather than the Mobile Computing Device 201. Similar to FIG. 9, the processing required to compute a DDP on the server is a hashed ID which identifies the particular device for which to compute the DDP. This hashed ID will also be used to retrieve the logged LDPs for that device...
1001 as well as to update the DDP 1002 upon completion of the computation as performed by step 109 previously described.

[0097] FIG. 11A and FIG. 11B combine to depict an illustrative embodiment of the invention highlighting aspects of the privacy preserving mechanisms. This simplified model shows the data and control flow of one embodiment of the system, highlighting those components which contribute to the privacy preserving capabilities. FIG. 11A represents activity occurring on the Demographic Public System 1150A. FIG. 11B represents activity occurring on the Demographic Private System 1160B. In one embodiment, all actions and data occur on systems independent of Mobile Device 1110A. In particular, the Mobile Device 1101A requests either a Location or Device Demographic Profile or a combination of both from a Location and Demographic Service. The Mobile Device 1101A passes either the location of the device or enough information for the service to compute its location in conjunction with a DeviceID that uniquely identifies the device.

[0098] The Demographic Service receives the request and routes it to the Demographic Public System 1150A which determines if location information is present 1102A. If location information is present, the request is routed to the Demographic Service 1104A, otherwise the request is routed to the Location Service 1103A. If the location is not present, the Location Service 1103A will attempt to compute the location of the device. If it cannot determine the location of the device, an empty (or equivalent) response is returned to the Mobile Device 1101A. Otherwise, the request is forwarded to the Demographic Service 1104A along with the newly computed location. Henceforth, all activities continue as if the request had come from the Mobile Device 1101A with location included.

[0099] Continuing with FIG. 11A, the information is passed to three independent processes: the Device Demographic Logger 1105A, the Location Demographic Counter 1114A, and the Location Demographic Logger 1118A.

[0100] The Device Demographic Logger 1105A provides a number of privacy preserving actions. It converts the DeviceID into a DeviceToken. It also converts the time of request to the local time based on the location of the device. This local time, in certain embodiments, may then be further reduced in specificity, for example by reducing it to indicate only hour of the week information. Other options include a representation of time lacking date information (such as month, day, and/or year) and/or a time range. The Device Demographic Logger 1105A then retrieves the demographic keys for the given location from the Location Demographic Keys 1106A. These keys, the converted request time, and the DeviceToken are then passed over an Administrative Boundary 1127 to the Device Demographic Loader 1107B in FIG. 11B.

[0101] The Administrative Boundary 1127 provides additional privacy preserving capabilities. It represents the separation of knowledge such that information available on the Demographic Public System 1150A is not discoverable on the Demographic Private System 1160B. Only the information that is explicitly shared from one side of the boundary to the other is available. For example, the relationship between the DeviceID and the DeviceToken is not discoverable on the Demographic Private System 1160B in FIG. 11B.

[0102] Continuing with FIG. 11B, the Device Demographic Loader 1107B generates log records from the input DeviceToken, request time, and Demographic Keys. The output from this process is the Device Dominographic Profile Logs 1108B.

[0103] The Device Demographic Processor 1110B reads records from the Device Demographic Profile Logs 1108B as well as specific values from the Location Demographic Attributes 1109B based on the Demographic Keys previously logged. The information in the Location Demographics Attributes 1109B represent the values for the demographic attributes, for example, the ratio of male people in the given area which are males based on previously known information. All Device Demographic Profile Logs records are gathered for each DeviceToken. Using this set of information, the Device Demographic Processor 1110B will perform various clustering algorithms to compute the most likely set of demographics that represent the mobile device for the specified criteria. For example, if the ‘Home’ demographic for the device is requested, the cluster algorithm may choose all demographic log records that occurred between 10 PM and 5 AM to represent the most likely times that a device would be located within the home demographics of the device owner.

[0104] Clustering is a common data mining method by which groups of 'like' items are assigned to a set by computing a difference, or distance between items. Cluster analysis can be accomplished with a number of different algorithms based on the specific need. In particular, a cluster of similar demographic records can be found by measuring the difference between two demographic records. Records with similar values can be clustered together. An example would be to measure the difference between all attribute’s buckets. If the average value between buckets is less than 0.1, these records would be considered as part of the cluster.

[0105] Upon successful clustering by the Device Demographic Processor 1110B, a set of records that represent the computed demographics for the device will be written to the Device Demographic Profiles 1111B database and a relationship between the DeviceToken and the DemographicsProfileToken will be stored in the DeviceTokenTags 1126B database.

[0106] Continuing in FIG. 11B, a copy of the Demographics Token 1126B database is transferred to the Demographic Public System 1150 in FIG. 11A (at 1126A). The Device Demographic Profiles 1111B are optionally copied to the Demographic Public System 1150A (at 1111A) for use in computing a public version the Device Demographic Summary.

[0107] Returning to FIG. 11A, a request by the Mobile Device 1101 will also instigate a process within the Location Demographic Counter 1114A. This process counts the presence of unique devices within a given geographic boundary during a specific time window. For example, it may count the number of devices that appeared in a city block on Monday between 3 PM and 4 PM. Utilizing, for example, a Probabilistic Counter 1115A such as a Bloom filter, the unique number of devices can be counted without identifying the specific Devices that were counted, thus preserving the privacy of the individual devices that are found to be present at various locations. These counts are then passed to the Device Count Loader 1116B located within the administrative domain of the Demographic Private System 1160B shown in FIG. 11B.
[0108] The Device Count Loader 1116B generates records that are stored in the Device Count Logs 1117B database in FIG. 11B.

[0109] Returning to FIG. 11A, a request by the Mobile Device 1101 will also instigate a process within the Location Demographic Logger 1118A. This process will convert the DeviceID to a DeviceToken. It will then use this information to find a matching record, if available, within the Demographic Tokens database. If present, the record will contain the DeviceToken for the given DeviceID. The specific location is converted to a geometric shape representing some area larger than the identified point. This process reduces the specificity of the location, providing additional privacy preserving properties. Additionally, the specific time of the request is reduced in a manner similar to that described within the Device Demographic Logger 1105A process to provide additional privacy preserving qualities. The DeviceToken, the location geometry, and the modified time value are then sent to the Location Demographic Loader 1119B within the Demographic Private System 1160B shown in FIG. 11B.

[0110] Continuing in FIG. 11B, the Location Demographic Loader 1119B generates records which are stored in the Location Demographic Logs 1120B database.

[0111] The Location Demographic Processor 1121B periodically computes the Location Demographic Profiles for each location geometry for which the Location Demographic Logs 1120B database contains information. For each location geometry, the process retrieves records for a specific time frame (e.g. each hour). These records are combined with the Device Demographic Profiles 1111B database to compute the aggregate Location Demographic Profile for that location during that time duration.

[0112] For each Location Demographic Logs 1120B record, the associated Device Demographic Profiles 1111B record is retrieved based on the DeviceToken. To combine the demographic information, a number of techniques can be applied. For example, the median of the demographics for each attribute can be chosen. Using median removes outliers and considers the best representative of temporal/spatial demographics. The median can be modified to select a bin with the most number of like records, if they exist. In this case, the most frequently occurring values could be chosen as the result for each attribute and bucket.

[0113] An alternative method would be to use the average of demographics, which is calculated by averaging each bucket of each attribute. Other statistical means to compute the combined demographics could be applied (e.g. weighted average based on number of like samples).

[0114] The Device Count Logs 1117B are consulted to assist in computing the confidence in the profile based on the number of devices observed within the given location geometry during the time in question. The confidence can be computed based on the total number of samples observed for that location within the time window. This number can be compared to historical sample counts to determine if the sample count is statistically significant. Alternatively, the expected population of the area can be used to compare the sample count to determine the statistical significance of the samples. In one embodiment, the statistical significance would be used as the confidence factor. Other factors, such as the accuracy of the location, or externally available information about the number of mobile device users in the area could be used to compute the confidence factors.

[0115] This process is repeated for each demographic attribute and for each desired time span. This results in record for each computation that includes the location geometry, the time span, the confidence factor(s) and values for each computed bucket within the given attribute (e.g. the computed ratio of males and females). These records are then written to the Location Demographic 1112B database.

[0116] The Location Demographic database 1112B is copied to the Demographic Public System 1150A depicted in FIG. 11A.

[0117] Returning to FIG. 11A, if the Mobile Device 1101A has requested a Location Demographic Profile, the Location Demographic Summary 1123A process is executed for the location that is requested. This process combines information from the Demographic Bucket Reference 1125A database, the Demographic Attribute Reference 1124A database, and the Location Demographics 1112A database to generate a record indicating the computed demographics for a given geographic geometry for the desired time window. The reference tables allow the generic information contained in the Location Demographics 1112A database to be converted back into semantically meaningful demographic attribute and bucket names (e.g. convert from “Table A” to “Education”, and “Buckell” to “Bachelors”). This record is then returned to the Mobile Device 1101A.

[0118] Optionally, the Mobile Device 1101A may request its own Device Demographic Summary. If requested, the Device Demographic Summary 1113A process will be executed. This process will convert the DeviceID to a DeviceToken. Using this information, it will retrieve the Device Demographic Profiles 1111A for this device. It will use this information as well as information obtained from the Demographic Tokens 1126A database, the Demographic Bucket Reference 1125A database and the Demographic Attribute Reference 1124A database to generate a record that summarizes the computed demographics for the given device. This record is then returned to the Mobile Device 1101A. Note that providing this optional service reduces the privacy preserving nature of the system. In order to produce the Device Demographic Summary, the Demographic Public System 1150A must have access to the Device Demographic Profiles 1111A database, thus providing a potential for the exposure of this private information.

[0119] FIG. 12 shows an embodiment of the process 1200 to convert demographic data into privacy preserving subsets. This process occurs prior to the system being deployed to the Operational Data 1230, and can be processed once for each demographic attribute that should be included in the demographic computations. FIG. 12 shows a specific demographic attribute, the Education Demographics 1215, being converted from the Original Data 1210 into its constituent parts in the Location Demographic Attributes 1240 database and the Location Demographic Buckets 1250 databases.

[0120] Continuing with FIG. 12, the Demographic Encoder 1220 reads the Education Demographics 1215 database. For each record, it generates a reference ID, called the Recno. This Recno will serve as the link between the G1 Demographics 1241 database and the A1 Demographics database 1252. The Demographic Encoder 1220 then stores the Location from the Education Demographics 1215 record along with the generated Recno in the G1 Demographics database. It also stores the generated Recno along with the values from the columns in the Education Demographics 1215 database into their respective columns in the A1 Demographics 1252 database.
base. For example, the values in the High School column from the Education Demographics 1215 database will be stored in the Bucket 1 column of the A1 Demographics 1252 database for their respective rows. The extent of the Locations in the Education Demographics 1215 database can be any size and shape that provides a meaningful designation of likely common demographics shared by those within the extent. For example, the Locations can be a neighborhood, a city block, a venue, a small town, or anything on the order thereof.

[0121] In addition to splitting the data values between two distinct tables, the relationship between these tables and new column names must be maintained in order to reverse the mapping when processed data is to be returned to the Mobile Device. To encode this mapping, two additional databases are required. The Demographic Attribute Reference 1243 database encodes the mapping from the new table name (e.g. “A1”) to the original data table name (e.g. “Education”). The second database, Demographic Bucket Reference 1242, provides the mapping for each bucket or column within an attribute. For example, this allows the system to convert from “Bucket 1” to the semantically meaningful value of “High School”.

[0122] Continuing with FIG. 12, note that the division of the data is also maintained across the Administrative Boundary 1127. The G1 Demographics 1241 database is contained within the Demographic Public System 1150 while the A1 Demographics 1252 is contained on the Demographic Private System 1160. This division is an example that provides additional privacy preserving properties both by obfuscating the information (e.g. converting recognizable information such as “High School” to generic labels such as “Bucket 1”) as well as separation of knowledge (e.g. the relationship of Recno to Area is only known by the Demographic Public System 1150, while the relationship of Recno to attribute buckets is only known by the Demographic Private System 1160).

[0123] FIG. 13A and FIG. 13B show an instance of FIG. 11A and FIG. 11B containing additional detail with respect to the data elements used for privacy preserving purposes. In particular, FIG. 13A and FIG. 13B depict an embodiment of the system in which the abstract database icons and names from FIG. 11A and FIG. 11B have been replaced with specific instances of tables and data to provide more concrete examples of the privacy preserving features. In particular, it depicts a single demographic attribute, “Education” and the attribute buckets of “High School”, “Bachelors”, and “PhD”.

[0124] In FIG. 13A, the Mobile Device 1301A requests either a Location or Device Demographic Profile or a combination of both from a Location and Demographic Service 1300A. The Mobile Device 1301A passes either the location of the device or enough information for the service to compute its location in conjunction with a DeviceID.

[0125] The Demographic Service 1300A receives the request and routes it to the Demographic Public System 1350A which determines if location information is present via 1302A. If location information is present, the request is routed to the Demographic Service 1304A, otherwise the request is routed to the Location Service 1303A. If the location is not present, the Location Service 1303A will attempt to compute the location of the device. If it cannot determine the location of the device, an empty (or equivalent) response is returned to the Mobile Device 1301A. Otherwise, the request is forwarded to the Demographic Service 1304A along with the newly computed location. Henceforth, all activities continue as if the request had come from the Mobile Device 1301A with location included.

[0126] Continuing with FIG. 13A, the information is passed to three independent processes: the Device Demographic Logger 1305A, the Location Demographic Counter 1314A, and the Location Demographic Logger 1318A.

[0127] The Device Demographic Logger 1305A provides a number of privacy preserving actions. It converts the DeviceID into a DeviceToken. It also converts the time of request to the local time based on the location of the device. This local time, in certain embodiments, may then be further reduced in specificity, for example by reducing it to indicate only hour of the week information. The Device Demographic Logger 1305A then retrieves the demographic keys for the given location from the G1 Demographics 1306A database. For example, assume the request was received by a location within “Area1”, the resulting key (e.g. Recno) would be “8230”. This Recno, the converted request time, and the DeviceToken are then passed over an Administrative Boundary 1327A to the Device Demographic Logger 1307B shown in FIG. 13B.

[0128] The Administrative Boundary 1327A provides additional privacy preserving capabilities. It represents the separation of knowledge such that information available on the Demographic Public System 1350A is not discoverable on the Demographic Private System 1360B in FIG. 13B. Only the information that is explicitly shared from one side of the boundary to the other is available. For example, the relationship between the DeviceID and the DeviceToken is not discoverable on the Demographic Private System 1360B.

[0129] Referring to FIG. 13B, the Device Demographic Loader 1307B generates log records from the input DeviceToken, request time, and Demographic Keys. The output from this process is the A1 Device Logs 1308B. Each record contains the DeviceToken, the HourOfWeek, and the Recno as sent by the Device Demographic Logger 1305A from FIG. 13A.

[0130] Continuing with FIG. 13B, the Device Demographic Processor 1310B reads records from the A1 Device Logs 1308B as well as specific values from the A1 Demographics 1309B based on the Demographic Keys previously logged. The information in the A1 Demographics 1309B represent the values for the buckets within the “A1” demographic attribute. All Device Demographic Profile Logs records are gathered for each DeviceToken. Using this set of information, the Device Demographic Processor 1310B will perform various clustering algorithms to compute the most likely set of demographics that represent the Mobile Device for the specified criteria. For example if the ‘Home’ demographic for the device is requested, the cluster algorithm may choose all demographic log records that occurred between 10 PM and 5 AM to represent the most likely times that a device would be located within the home demographics of the device owner.

[0131] Upon successful clustering by the Device Demographic Processor 1310B, a set of records that represent the computed demographics for the device will be written to the Device Demographic Profiles 1311B database and a relationship between the DeviceToken and the DemographicProfileToken will be stored in the Demographic Token 1326B database. Examples of this output are depicted in the respective databases in FIG. 13B. In particular, it is noted that DeviceToken “ax8778as02” is connected to the Recono “8230” via this relationship as indicated in FIG. 13B.
Continuing in FIG. 13B, a copy of the Demographic Token 1326B database is transferred to the Demographic Public System 1350A depicted on FIG. 13A (as 1326A). The Device Demographic Profiles 1311B database is optionally copied to the Demographic Public System 1350A, shown in FIG. 13A, for use in computing a public version the Device Demographic Summary (as 1311A).

Returning to FIG. 13A, a request by the Mobile Device 1301A will also instigate a process within the Location Demographic Counter 1314A. This process counts the presence of unique devices within a given geographic boundary during a specific time window. For example, it may count the number of devices that appeared in a city block on Monday between 3 PM and 4 PM. Utilizing, for example, a Probabilistic Counter 1315A such as a Bloom Filter, the unique number of devices can be counted without identifying the specific devices that were counted, thus preserving the privacy of the individual devices that are found to be present at various locations. These counts are then passed to the Device Count Loader 1316B located within the administrative domain of the Demographic Private System 1360B depicted in FIG. 13B.

Continuing with FIG. 13B, the Device Count Loader 1316B generates records that are stored in the Device Count Logs 1317B database. For example, referring to FIG. 13B that for Geometry “ABAF007” and hour 57 of the week, there were 8 distinct devices counted.

Returning to FIG. 13A, a request by the Mobile Device 1301A will also instigate a process within the Location Demographic Logger 1318A. This process will convert the DeviceID to a DeviceToken. It will then use this information to find a matching record, if available, within the DemographicToken database 1326B. If present, the record will contain the DemographicToken for the given DeviceToken. The specific location is converted to a geometric shape representing some area larger than the identified point. This process reduces the specificity of the location, providing additional privacy preserving properties. Additionally, the specific time of the request is reduced in a manner similar to that described within the Device Demographic Logger 1305A process to provide additional privacy preserving qualities. The Demographic Token, the location geometry, and the modified time value are then sent to the Location Demographic Loader 1319B within the Demographic Private System 1360B shown in FIG. 13B.

Continuing in FIG. 13B, the Location Demographic Loader 1319B generates records which are stored in the Location Demographic Logs 1320B database.

The Location Demographic Processor 1321B periodically computes the Location Demographic Profiles for each location geometry for which the Location Demographic Logs 1320B database contains information. For each location geometry, the process retrieves records for a specific time frame (e.g. each hour). These records are combined with the Device Demographic Profiles 1311B database to compute the aggregate Location Demographic Profile for that location during that time duration. The Device Count Logs 1317B are consulted to assist in computing the confidence in the profile based on the number of devices observed within the given location geometry during the time in question. This process is repeated for each demographic attribute and for each desired time span. This results in a record for each computation that includes the location geometry, the time span, the confidence factor(s) and values for each computed bucket within the given attribute. These records are then written to the A1 Location Demographics 1312B database. Referring to FIG. 13B, there is a single entry in A1 Location Demographics 1312B database representing the computed values for demographic attribute “A1” at location geometry “ABAF007” for the 57th hour of the week.

The A1 Location Demographics database 1312B is copied to the Demographic Public System 1350B (as 1312A) shown in FIG. 13A.

Returning to FIG. 13A, if the Mobile Device 1301A has requested a Location Demographic Profile, the Location Demographic Summary 1323A process is executed for the location that was requested. This process combines information from the Demographic Bucket Reference 1325A database, the Demographic Attribute Reference 1324A database, and the Location Demographics 1312A database to generate a record indicating the computed demographics for a given geographic geometry for the desired time window. The reference tables allow the generic information contained in the A1 Location Demographics 1312A database to be converted back into semantically meaningful demographic and bucket names. This record is then returned to the Mobile Device 1301A.

FIG. 14 depicts an example output 1400 based for an Location Demographic Profile produced by the system shown in FIG. 13A and FIG. 13B. This represents a partial demographic profile 1401 based on the limited example data contained in the figure.

Optionally, the Mobile Device 1301A may request its own Device Demographic Summary. If requested, the Device Demographic Summary 1313A process will be executed. This process will convert the DeviceID to a DeviceToken. Using this information, it will retrieve the Device Demographic Profiles 1312A for this device. It will use this information as well as information obtained from the Demographic Token 1326A database, the Demographic Bucket Reference 1325A database and the Demographic Attribute Reference 1324A database to generate a record that summarizes the computed demographics for the given device. This record is then returned to the Mobile Device 1301A. Note that providing this optional service reduces the privacy preserving nature of the system. In order to produce the Device Demographic Summary, the Demographic Public System 1350A must have access to the Device Demographic Profiles 1311A database, thus providing a potential for the exposure of private information.

In the foregoing description, certain steps or processes were described as being performed on particular servers or as part of a particular engine. These descriptions are merely illustrative, as the specific steps can be performed on various hardware devices, including, but not limited to, server systems and/or mobile devices. Similarly, the division of where the particular steps are performed in the above description illustrates certain embodiments, if being understood that no division or a different division is within the scope of the invention.

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.

[0144] 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.

[0145] 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.

[0146] 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).

[0147] 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.

1. A method of estimating demographic information associated with a user of a mobile device while preserving the privacy of the user based at least in part on a location estimate of the mobile device of the user, the method comprising:
   (a) receiving an estimated geographical location of the mobile device of the user;
   (b) receiving a time at which the mobile device was at the estimated geographical location;
   (c) providing a set of substitute identifiers for a corresponding set of at least one geographical area;
   (d) assigning one of the set of substitute identifiers for the geographical area corresponding to the geographical location of the mobile device;
   (e) assigning a substitute identifier for the time at which the mobile device was at the estimated geographical location;
   (f) providing an association between the substitute identifiers for geographical areas and demographic information corresponding to the substituted geographical area; and
   (g) estimating demographic information associated with the user of the mobile device based on the assigned substitute identifiers and based on the demographic information associated with the provided set of substitute identifiers.

2. The method of claim 1, further comprising (h) estimating the received geographical location associated with the mobile device of the user.

3. The method of claim 1, further comprising (h) recording the estimated demographic information associated with the user of the mobile device in a device demographic information log.

4. The method of claim 3, further comprising (i) estimating demographic information associated with at least one geographical area of the set of geographical areas based on the demographic information recorded in the device demographic information log, wherein the device demographic information log contains a plurality of records of estimated demographic information associated with a plurality of mobile devices.

5. The method of claim 1, further comprising (h) sending the estimated demographic information to the mobile device.

6. The method of claim 1, wherein the substitute identifier for the geographical area corresponding to the geographical location of the mobile device that is assigned identifies a particular set of demographic information.

7. The method of claim 1, wherein the substitute identifier for the geographical area corresponding to the geographical location of the mobile device that is assigned is reduced in specificity relative to the estimated geographical location of the mobile device.

8. The method of claim 1, wherein the substitute identifier for the time at which the mobile device was at the estimated geographical location is a measure of time that is reduced in specificity relative to the time that was received.

9. The method of claim 8, wherein the substitute identifier for the time at which the mobile device was at the estimated geographical location is at least one of (A) a representation of time lacking date information, (B) a time range, and (C) an hour of a week designation.

10. The method of claim 1, wherein the estimating demographic information associated with the user of the mobile device is performed on a separate computer system from a computer system performing any one or more of steps (a) through (e).

11. The method of claim 10, wherein the separate computer system, relative to the computer system performing any one or more of steps (a) through (e), is as least one of (A) maintained in a separate network, (B) maintained in a separate building, and (C) maintained by a separate operational entity.

12. A method of estimating demographic information associated with a geographical area and a time period based on demographic information associated with users of mobile devices within the geographical area, the method comprising:
   (a) providing a set of geographical areas;
   (b) providing a set of time periods;
   (c) receiving an estimated geographical location of a mobile device of the user;
   (d) receiving a mobile device identifier that is associated with the mobile device;
   (e) receiving a time at which the mobile device was at the estimated geographical location;
   (f) determining the geographical area of the set in which the estimated geographical location occurs;
   (g) determining the time period of the set in which the time at which the mobile device was at the estimated geographical location occurs;
   (h) retrieving information representative of demographic information associated with the user of the mobile device based on the mobile device identifier;
   (i) estimating demographic information associated with the determined geographical area during the determined
13. The method of claim 12, further comprising (j) estimating the received geographical location of the mobile device of the user.

14. The method of claim 12, further comprising (j) providing a set of initial demographic information associated with the determined geographical area, wherein the estimating demographic information associated with the determined geographical area is further based on the initial demographic information.

15. The method of claim 14, further comprising, (k) subsequent to the estimating demographic information associated with the determined geographical area, adjusting the initial demographic information based on the estimated demographic information.

16. The method of claim 14, wherein the initial demographic information is based on at least one of governmental census information, public record information, shipping and purchase records, magazine and newspaper subscriptions, voluntary surveys, and records of social media activity.

17. The method of claim 12, further comprising (j) sending the estimated demographic information to the mobile device.

18. The method of claim 12, wherein the geographical areas of the set are reduced in specificity relative to the estimated geographical location of the mobile device.

19. The method of claim 12, wherein the time periods of the set are reduced in specificity relative to the time that was received.

20. The method of claim 19, wherein the time periods of the set are at least one of (A) a representation of time lacking date information, (B) a time range, and (C) an hour of a week designation.

21. The method of claim 12, wherein the estimating demographic information associated with the determined geographical area during the determined time period is performed on a separate computer system from a computer system performing any one or more of steps (c) through (g).

22. The method of claim 21, wherein the separate computer system, relative to the computer system performing any one or more of steps (c) through (g), is at least one of (A) maintained in a separate network, (B) maintained in a separate building, and (C) maintained by a separate operational entity.

23. The method of claim 12, further comprising: (j) performing steps (c) through (h) a plurality of times for different mobile devices of different users; and (k) recording sets of information for (A) the determined geographical areas, (B) determined time periods, and (C) retrieved information representative of demographics information associated with the users of the mobile devices in a device demographic information log; wherein the estimating demographic information associated with the determined geographical area during the determined time period is further based on a plurality of the sets of information in the device demographic information log.

* * * * *