February 13, 2017

Analyzing crowdsourced location data to trigger actions

Michael Greene

Analyzing crowdsourced location data to trigger actions

ABSTRACT

Knowledge of the occupancy of campuses or buildings is important in order to deploy and distribute resources. Likewise, knowledge of the directions and types of movement of people and vehicles is important for urban planners, governments, facilities managers, etc. This disclosure describes techniques to derive real-time and long-term occupancy and people-movement data based on information received from devices. Such data can be utilized by a variety of customers, such as insurers, facilities managers, urban planners, governments, property owners, etc.

KEYWORDS

- Location-based services
- Location data
- Building occupancy
- Crowdsourcing

BACKGROUND

Occupancy of a facility such as a gymnasium, office etc. has an important bearing on its value, character, optimum service-intervals and other parameters. For example, knowing the purpose for which a certain building facility is being used is of interest to property insurers. Currently, insurers rely on human agents (e.g., assessors, surveyors, etc.) that inspect various insured properties to obtain information regarding use of the property. In another example, knowing real-time occupancy of a parking lot could facilitate management of incoming traffic.

Presently, parking lot management involves utilization of expensive, purpose-built sensors to determine parking-lot occupancy.

DESCRIPTION

This disclosure describes techniques to obtain occupancy data from devices such as mobile phones, smartwatches, tablets etc., when users of the devices agree to collection and use of such data. It further describes techniques to utilize such data, for example, to infer the purpose for which a certain facility is being used, to trigger service actions, to modify transportation pathways, etc.

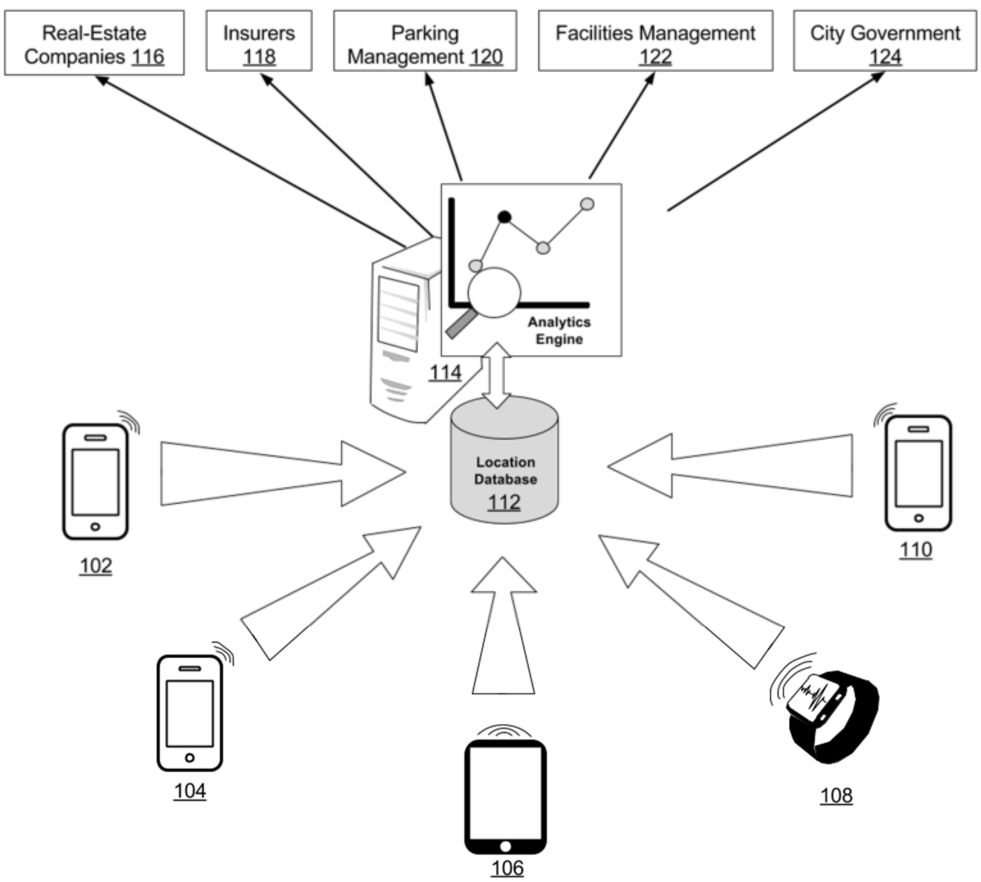

Fig. 1: Example environment

Fig. 1 shows an example environment in which the techniques of this disclosure are implemented. Devices (102-110) such as mobile phones, smartwatches, tablets etc. have on-board hardware or are otherwise capable of determining their locations. For example, a mobile phone is equipped with a global positioning sensor (GPS) that allows it to determine its latitude and longitude. Additionally, or alternatively, signals from cellular telephony base stations, Wi-Fi access points, Bluetooth beacons, ZigBee networks, etc. can also be used to determine the location of a device. User activity such as self-reporting of location, e.g., by checking-in etc., can also specify the location of a device.

When users consent to providing their location data, determined automatically or based on user activity, a user device sends its location(s) to a location database (112). The location database stores the received location data, if the user consents to such storage, with associated time of reporting for the locations. An analytics engine (114) accesses the location data and produces various inferences. Analysis of the data are performed using any of various techniques including, for example, machine-learning, pattern recognition, neural networks, etc. When users permit, inferences produced by the analytics engine are provided to various customers, such as real-estate companies (116), insurers (118), parking management (120), facilities managers (122), city governments (124), etc.

Transmission of location data by a user device to the location database is at all times controlled by the user and is performed only upon the user’s consent. Personally identifiable information is removed from the data prior to transmission, e.g., by anonymizing the data, aggregating data, etc. Further, in providing consent, the user can choose the granularity of data collection, for example, the precision e.g., within 1 meter, within 50 meters, within a postal code, within a city, etc. at which the location data is reported. Still further, the user can specify a frequency at which location data is reported, place restrictions on reports from certain locations, limit the total amount of data collected within a given time period, etc. Analysis of location data and drawing inferences is performed based on the user's permission for such analysis. The user can choose the types of inferences that are drawn from the data. Storage and analysis of data is performed in such a manner that no personally identifiable information is collected or used, without a user's explicit approval. The user can restrict the use of the data or inferences, e.g., make it available only to certain third parties, disallow third party usage, etc. The user can specify the retention time for their data, e.g., do not retain, retain for one hour, retain for one day, retain last *n* observations, etc. A user can decline authorization for collection of their data, in which case no location data is collected or used. The user is permitted to modify their consent and parameters of data collection.

Examples of use

Example 1: Property insurers

Calculation of insurance premium for any facility, e.g., a building, are sometimes based on the purpose for which the facility is used. For example, a bar, a restaurant, and an office each present different risks to an insurer, and therefore have different insurance premiums. Insurers utilize the actual use of a facility, and verify such use periodically.

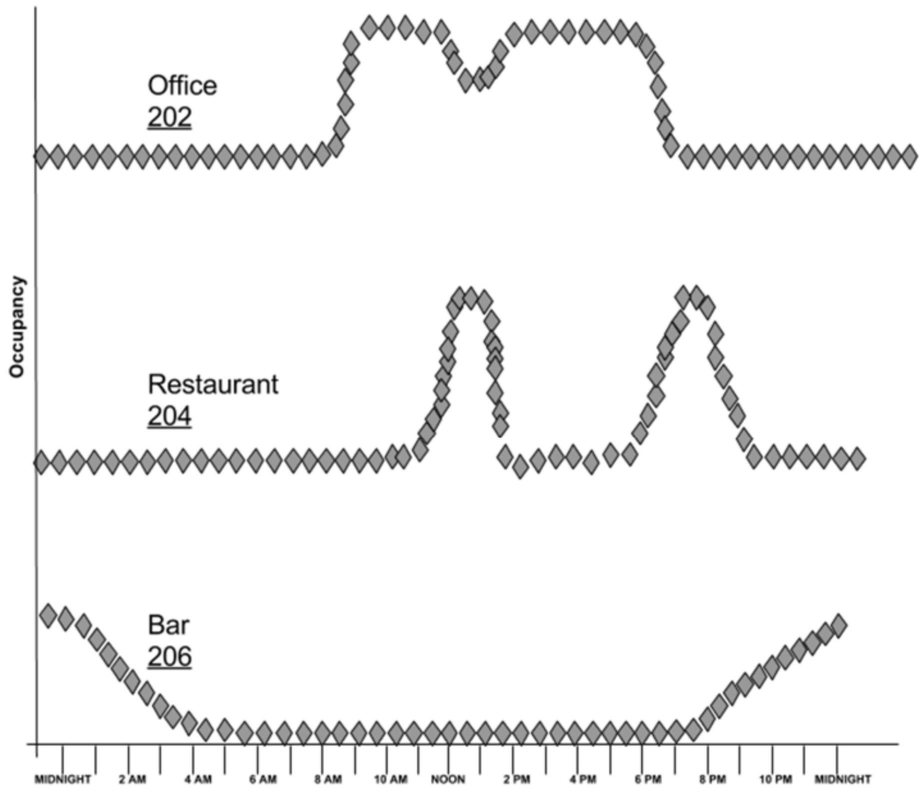

Fig. 2: Examples of occupancy patterns

Fig. 2 shows as a function of time-of-day occupancy levels of different example facilities, namely office (202), restaurant (204) and bar (206). In this example, occupancy of the office is more-or-less a plateau during working hours of 8 AM and 5 PM, with a one-hour lunch-time dip roughly between noon and 1 PM. The occupancy of a restaurant shows two peaks. The first peak is between noon and 1 PM, e.g., corresponding to lunch hour, and the other peak starts around 530 PM and ends around 830 PM, corresponding to dinner. The occupancy of a bar is low throughout daytime hours, and picks up after 6 PM and stays high until 2 AM or so.

Occupancy charts such as shown in Fig. 2 are generated by the analytics engine based on location data in the database. The occupancy charts are further analyzed, e.g., by a machine learner that is trained to recognize and label patterns. Once trained, the machine learner monitors occupancy of various facilities of interest to the insurer, infers the likely current use of such facilities, and reports such use to the insurer. For example, the machine learner reports deviations, e.g., unusually high occupancy of an office facility beyond office hours, or periodic occurrences of occupancy of a restaurant exceeding a threshold for fire safety, etc.

Example 2: Real estate

Value of commercial real-estate, such as a storefront or other facility, often depends on the amount of foot traffic and passers-by it experiences. The analytics engine aggregates and counts location reports received from within or close to a storefront. The count serves as potentially valuable data for market players, such as real-estate companies or store owners, to arrive at an appropriate value, e.g., for rent for the storefront.

Example 3: Optimal service intervals

Maintenance and janitorial services needed in facilities depends on the volume of visitors, rather than on a time schedule. For example, a gymnasium or restroom needs service such as cleaning, equipment maintenance, refurbishment of consumables such as soap, etc. on average after a certain number of visitors. The analytics engine counts the number of location reports received from facilities that require frequent service. Upon reaching a threshold, the analytics engine provides input that triggers a service request.

In this manner, service can be triggered based on real-time occupancy data, which is more optimal than service that is based on a fixed schedule. A facilities manager could maintain a central pool of staff across several facilities and provision staff to those locations that are flagged by the analytics engine as having met a threshold to trigger service. This approach reduces instances of service inefficiency such as staff not attending to a facility in need of service. Further, the analytics engine can be trained to recognize patterns in complaints and incidents in order to produce an optimal cleaning schedule.

Example 4: Urban planning

Cities are designing and putting in place shared-use and/or variable-width transportation pathways. For example, a sidewalk is partitioned between cyclists and pedestrians depending on the relative numbers of users of each kind. As a further example, a four-lane street is configured to have three northbound lanes and one southbound lane in the morning, and one northbound lane and three southbound lanes in the evening. Knowledge of real-time data of commuters that use the street, e.g., walk, bicycle, ride a bus, or drive, is useful to city governments to manage and optimize such reconfiguration. The analytics engine classifies various location reports received from devices into various categories such as pedestrians, bicyclists, public transport users, and car users, e.g., based on speed and direction of motion. Relative numbers of each category are provided by the analytics engine and are available for utilization by city governments, e.g., traffic management authorities, to optimize transport.

Example 5: Resource allocation

It is of interest to various entities, such as property managers, landlords, architects, security agencies, and urban planners to know the numbers and directions of movement of people in various public places. For example, such information is useful in installing or redeploying resources such as security personnel, escalators, elevators, etc. The analytics engine utilizes location data to determine trends and provide predictions of the movements of people in the aggregate. Further, the analytics engine accounts for factors like day-of-week, sales, major events etc., e.g., to predict the possibility of underserving at different resource levels.

Example 6: Parking management

It is of interest to a parking facility manager to know the numbers and locations of empty parking spaces. Such information may be determined by location reports transmitted by devices in the following manner. The analytics engine detects that an automobile has entered the parking structure and is attempting to occupy a particular parking space when it observes rapid (vehicular speed) movement of a device towards the particular space, followed by slow (walking-speed) movement of the same device away from the space. When a such sequence of events, such as, a vehicle being driven, then stopped, then occupants walking out, is detected by the analytics engine, it decrements the number of available spaces and marks the particular parking space as occupied. Conversely, when the analytics engine observes walking-speed movement of a device towards a particular parking space followed by vehicular-speed movement away from the parking space, it infers that a just-occupied space has been vacated. In this case, the number of available spaces is incremented, and the particular parking space is marked as empty. The analytics engine compensates for multi-passenger vehicles by identifying user devices with statistically significant coordination in space and time as co-passengers.

In situations in which certain implementations discussed herein may collect or use personal information about users (e.g., user data, information about a user’s social network, user's location and time at the location, user's biometric information, user's activities and demographic information), users are provided with one or more opportunities to control whether information is collected, whether the personal information is stored, whether the personal information is used, and how the information is collected about the user, stored and used. That is, the systems and methods discussed herein collect, store and/or use user personal information specifically upon receiving explicit authorization from the relevant users to do so. For example, a user is provided with control over whether programs or features collect user information about that particular user or other users relevant to the program or feature. Each user for which personal information is to be collected is presented with one or more options to allow control over the information collection relevant to that user, to provide permission or authorization as to whether the information is collected and as to which portions of the information are to be collected. For example, users can be provided with one or more such control options over a communication network. In addition, certain data may be treated in one or more ways before it is stored or used so that personally identifiable information is removed. As one example, a user’s identity may be treated so that no personally identifiable information can be determined. As another example, a user’s geographic location may be generalized to a larger region so that the user's particular location cannot be determined.

CONCLUSION

This disclosure describes techniques to collect location data from devices and utilize such data to draw various inferences to optimize use of resources such as streets, building facilities, service staff, etc. Collection, analysis and third-party usage of such data is performed specifically upon a user’s explicit authorization, and using parameters that are agreed to by the user.