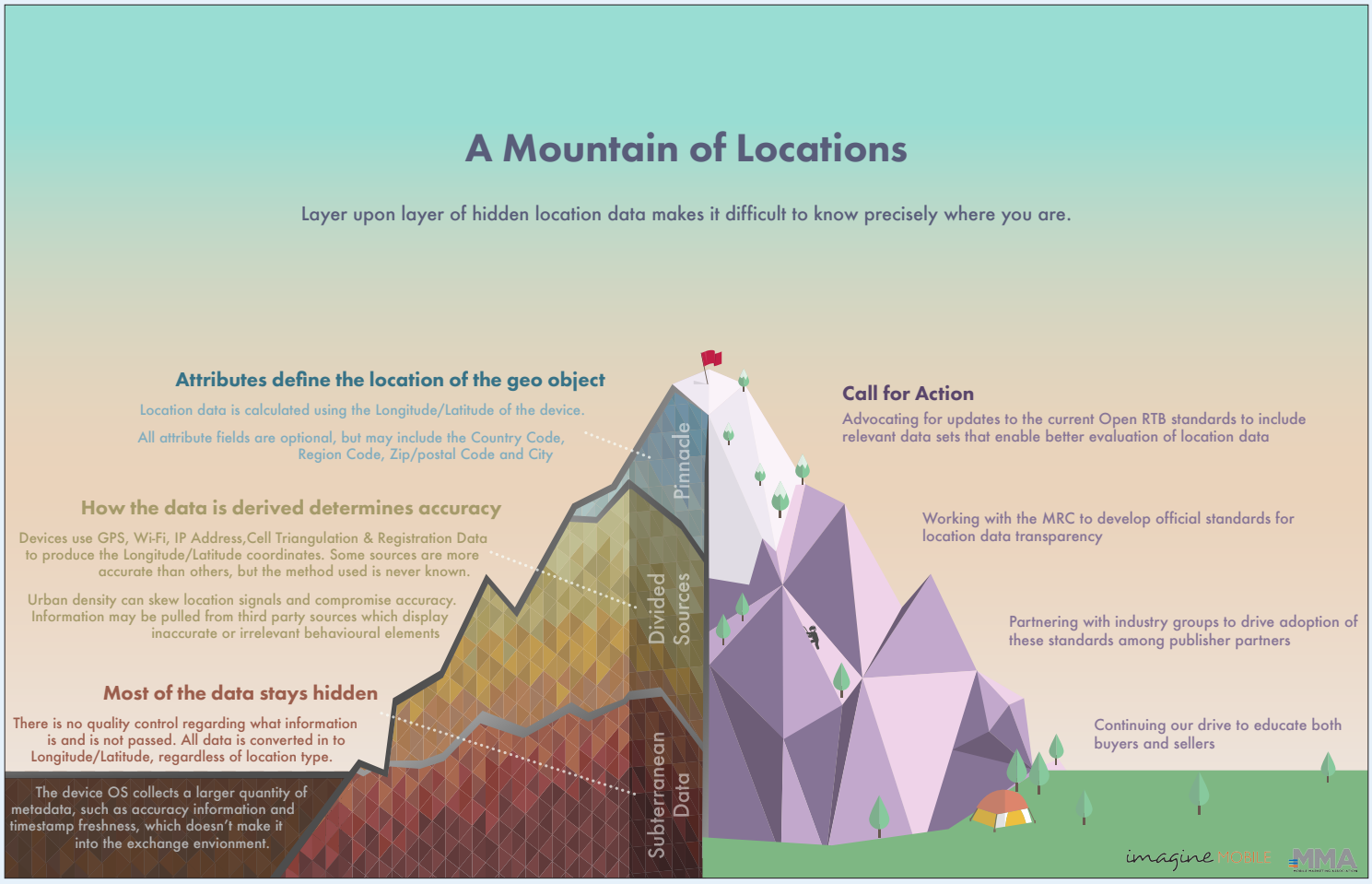
Demystifying Location Data Accuracy

The new frontier and biggest mobile opportunity

MMA
MOBILE MARKETING ASSOCIATION

CONTENTS

MOBILE MARKETING ASSOCIATION



Produced by: imagine MOBILE Contact UK USA Singapore

MMA
MOBILE MARKETING ASSOCIATION

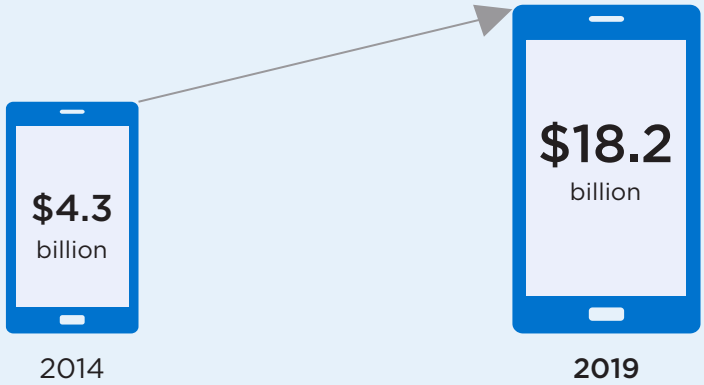
Introduction

The always-there, always-on nature of mobile devices combined with the ability to determine precise location gives marketers the unprecedented ability to deliver relevant messages to consumers. Location data can enhance the entire marketing cycle — from developing an understanding of who consumers are, to understanding their context at the moment the message is displayed, and then measuring the impact of marketing programs.

Location data, like all data, has its own nuances and issues. For a variety of reasons, the location data available in the mobile ad ecosystem is of varying quality. This paper explains some of the drivers of location data quality, how data quality impacts various mobile marketing use cases, and how the industry is working together to improve location data quality.

The simplest place to start is understanding why accurate location data is critical to the entire mobile ecosystem: brands, agencies, publishers, the end user and everyone in between. Accurate location data is proven to increase relevance to consumers, which in turn leads to improved ad performance. The quality of location data directly impacts one's ability to build audience understanding from location history. Audience segments based upon faulty data provide no meaningful value to consumers or advertisers. Better location data results in more targeted audiences. Location data also becomes critical when dealing with attribution, for example, measuring how foot traffic in a retail outlet will be affected by a mobile ad campaign. With location data's ability to show the direct impact of a mobile ad campaign on physical behavior, it offers the media industry a crucial connector of a person's online and offline activities.

In the U.S., location-targeted mobile ad revenues are projected to grow from $4.3 billion in 2014 to $18.2 billion in 2019, a 33.5 CAGR.



This is defined as mobile advertising targeted based on a user's location, or including proximity-relevant content to trigger local offline conversions.

source: BIA Kelsey Local Media Forecast, 2015

Location Data Uses

In the past few years location data has transformed mobile marketing, introducing both new targeting technologies and improving the ability to define and track audiences and behaviors.

Here are a few examples of how location is driving marketing innovation on mobile, going far beyond the standard geo-fencing methodologies:

- **Proximity Targeting:** Delivering ads based on users' real-time location, typically defined as proximity to specific place(s). For example, "all users within 100 meters of a Ford Dealership" or "users in California who are currently at a Taco Bell".
- **Location-based Audiences:** Deriving an understanding of a user, based on analyzing that specific user's location history. Analysis of geo-behavioral patterns can be used to infer demographics, behavioral attributes, and geographic attributes (i.e. where the user lives or works). 3rd party data (e.g. purchase data, TV viewership, car ownership etc.) can be joined to provide an understanding of home location. (MMA Location Committee Guidance Report: Location Audience Targeting)
- **Attribution:** Measuring store visitation or attaching purchase data (as described above) from location data. These metrics can be used to measure the ROI of ad campaigns.
- **Location-based Creatives:** Dynamically modifying ad creative based on the user's location. Typically this breaks down into two categories: modifying the creative itself (e.g. inserting the address or phone number of a location in the creative or inserting a localized offer) and modifying the click action on the creative (e.g. click-to-call and click-to-map/navigation actions.)

How Location Data Is Collected

Data Passed in an Ad Request

When discussing data passed by the publisher, either through a direct inventory relationship or through a media exchange, there are often misconceptions regarding data and quality. Direct, open/private exchanges are just a few examples of data transfer protocols, and they do not automatically imply what data is passed or what efforts are made to verify it.

Since there is no common API used for direct relationships, we will focus in this analysis on the exchanges protocol, as the same concepts and data fields are applicable to the direct case, and often serve as a baseline regardless of whether or not the inventory is purchased programmatically.

Tech leaders on both the demand and supply side have recognized the challenge of identifying the datasets to be used as standard units for the valuation of inventory. With the growth of automated trading of digital media across a broader range of platforms, the various industry stakeholders came together to form the OpenRTB project.

Ad Request: when a user's browser or device requests information from an ad server, an 'ad request' occurs.

API: an abbreviation of application program interface, is a set of routines, protocols, and tools for building software applications. The API specifies how software components should interact.

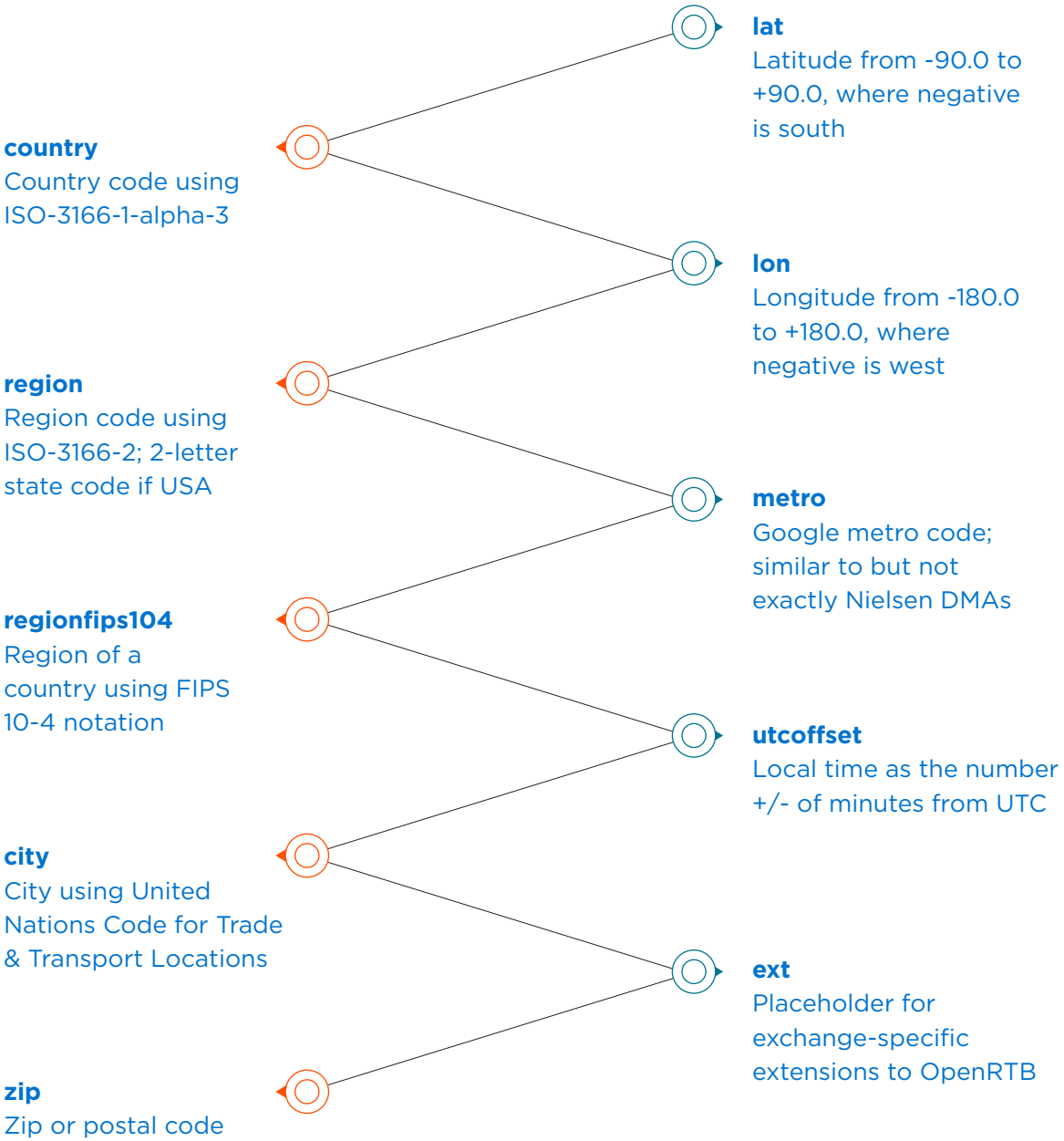
RTB: an abbreviation for Real-Time Bidding, is a means by which advertising inventory is bought and sold on a per-impression basis, via programmatic instantaneous auction, similar to financial markets.

The OpenRTB protocol serves as a "contract" between sellers and buyers of online media and defines the type of information that should and can be passed in an ad request, such as information about the **Impression**, the **User** and the **Device**.

The OpenRTB spec defines **Geo** objects which represent physical locations, and can be attached to either the **Device** or the **User** objects.

It is important to note that both the Device object and the User object can contain location information in **the same** ad request but for each object, yet the location data will have different meaning. For example, a Geo object that is attached to a Device object is intended to represent the "current" location of the device at the moment of the request. A Geo object attached to a User object will represent the user's "home base" location which likely to be different from the user's current location.

Information contained in the Geo object:



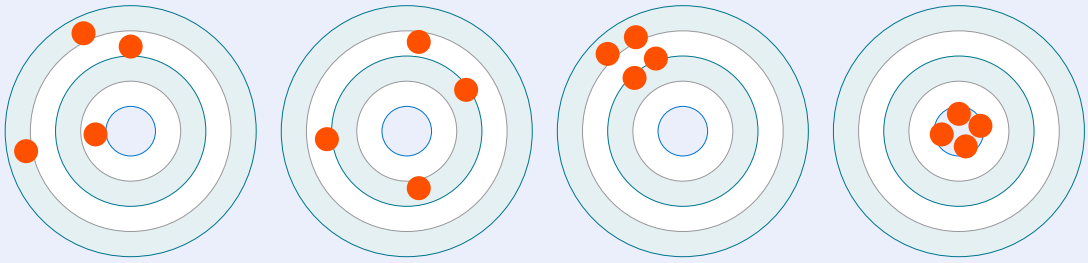
All fields in the Geo object are optional, so some requests may appear with only a subset of attributes.

An important field in the Geo object in relation to the Lat/Lon is the **type** attribute, which indicates the source of the location data, and can be either:

- **GPS/Location Services**
- **IP Address**
- **User provided** (e.g., registration data)

The **type** information contained in the Geo object provides an important indicator to assess the accuracy level of the location data. It should be noted, this information does not guarantee any specific accuracy level and might be misleading if taken "as is" without additional data or verification methods. Some advanced buy-side platforms will use big data methodologies and algorithms to detect patterns of inaccurate/fraudulent location data and filter it out.

Location Accuracy and Precision — though they sound similar there is a subtle difference in the meaning. Accuracy refers to the closeness of a measured location to the real location of the device at the time of the measurement. Precision refers to the closeness of two or more measurements to each other.



NOT ACCURATE
NOT PRECISE

ACCURATE
NOT PRECISE

PRECISE
NOT ACCURATE

ACCURATE
AND PRECISE

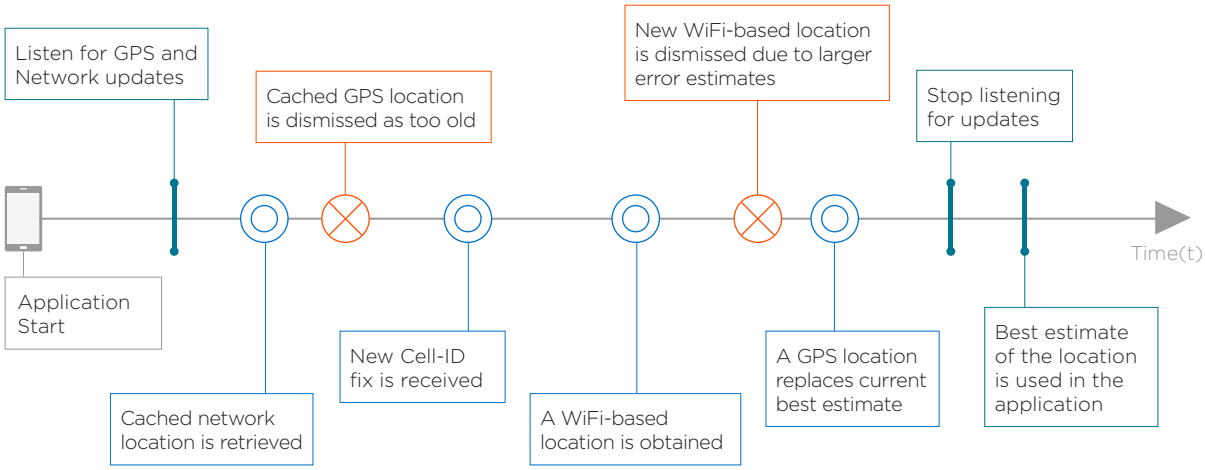
Device Location Services

In both iOs and Android environments the developer relies on the operating system to get a location fix, and uses the Location Services API to query the device's current location.

The Location Service is a system that uses various sources of information with varying levels of accuracy in parallel to acquire a location. Regardless of the sources used, location is always returned as Lat/Lon coordinates. These sources include, amongst others:

Source	Location Data Traits	Use Cases
Beacons	Very accurate and precise data, capable of identifying user location to within a matter of feet. Currently lacks scale for broad marketing opportunities	In-store targeting by department or aisle, verification of store visits
GPS	Very accurate and precise data under the right conditions, capable of identifying user location to within 10-100 meters (in most ad platform scenarios). Large scale but dependent on publisher and user enablement of GPS The quality of a GPS signal degrades significantly indoors or in locations that do not have an unobstructed view of multiple GPS satellites	Numerous targeting opportunities based on user location & context
Wifi	Very accurate and precise data, particularly in signal-dense environments (e.g. buildings and malls) Capable of identifying user location to within 10-100 meters when wifi signals are present	Numerous targeting opportunities based on user location & context
Cell Triangulation	Accurate but less precise data. Typically capable of identifying user location within a zip or neighborhood area	Targeting over a broader geo-area such as a DMA, city or zip code
IP Addresses	Not reliably accurate or precise. Typically based on IP address of app server which can vary significantly from true location of mobile user	Not relevant for real-time or behavioral targeting
Registration data	Not relevant for real-time or behavioral targeting Typically relevant only for a mobile user's home zip code at time of registration	Targeting at the zip code or DMA level

Since the operating system uses multiple sources concurrently, it is important to note that the device location is constantly being updated as the service continues to operate. The below illustration from the Android developer guide provides an example of a possible scenario:



In addition to the time effect on accuracy, each developer can use different strategies and direct the OS to use more coarse or granular methods. For example, the developer may elect to conserve a user's battery and work only with city level accuracy. The developer can also select how often to update the location data and whether to be notified on significant changes.

Therefore, in many cases, even when the Location Service is used and a Lat/Lon fix is retrieved, it may not represent an accurate or recent location. However, additional information returned by the OS can be used to determine the data quality.

As seen in both the Android and iOS Location Service APIs, in addition to the latitude and longitude included in the Location object returned by the operating system, there are additional parameters such as accuracy, time of reading, speed, altitude etc.

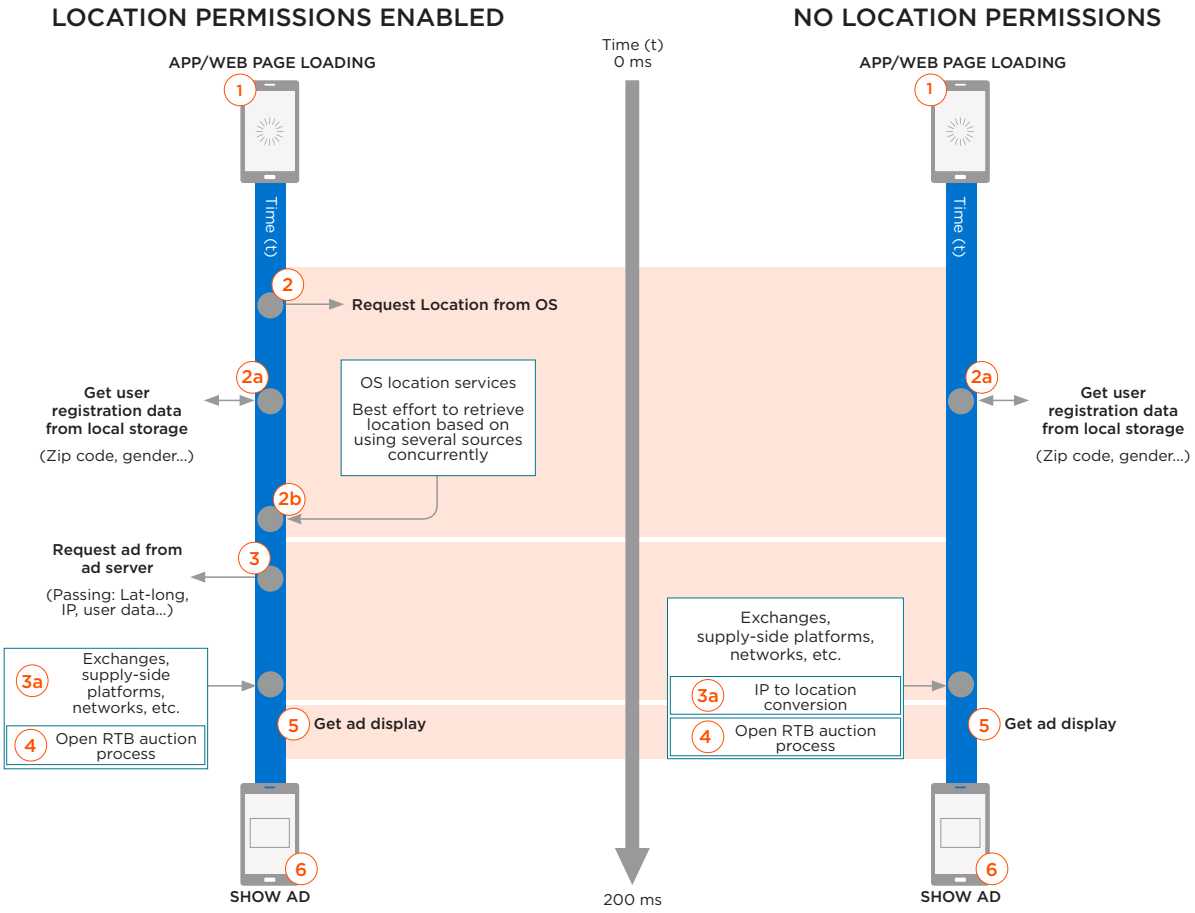
Out of the various parameters returned by the operating system the most interesting and relevant to location data accuracy are:

- **Accuracy** — Estimated accuracy of location in meters
- **Timestamp** — The time at which this location was determined

However, since these parameters are not included in the OpenRTB protocol, they are often ignored by developers and not passed in addition to the Lat/Lon data retrieved.

Location Data Flow

Once it's clear what data is passed from the publisher to the buyer and the process of obtaining that data from the device, the next step in understanding the issues surrounding location accuracy is to review the overall process from the moment the mobile application/web page loads to the point where an ad is being displayed.



1. A user opens an app or opens a new page / area of an app
2. As the page loads, if OS location services is enabled, a request is made for the current location. A combination of GPS, wifi and cell-tower triangulation is used to provide the most accurate Lat/Lon
   - 2a. User provided data is retrieved from local storage
   - 2b. Location data is returned as Lat/Lon regardless of location source used
3. The ad request is sent to the exchanges, supply side platforms, networks etc
   - 3a. In an effort to enhance the bid request, additional data may be inserted by different parties at this point
4. The bid won by the party who provided the highest bid in the time allotted (milliseconds)
5. Request for ad from ad server
6. Ad is displayed to user

Factors Impacting Accuracy

Due to the increasing demand for location data, combined with the lack of standards around how this data is collected and used, there is a risk of data becoming compromised (intentionally or unintentionally) in the process of being obtained by the publisher and delivered to the ad property.

> Up to 60% of ad requests contain some form of location data. Of these requests, less than ⅓ are accurate within 50-100 meters of the stated location.

(source: Foursquare, Thinknear, Ubimo, xAd, Inc)

Publishers are focused on building apps and websites and often do not invest in resources to better understand the data which they send to programmatic exchanges. Additionally, there are some inherent technical factors including variance in the quality of the methods used to derive location for the specific data fields which are passed on to the exchanges.

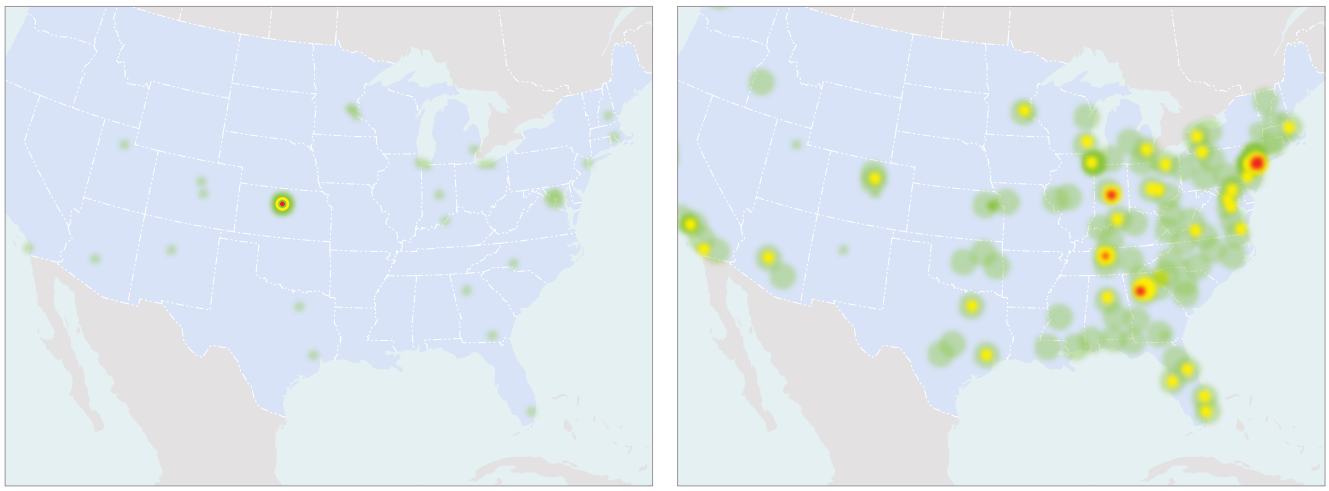
Some of the main issues encountered in this process include:

- **Lack of accuracy standards and market education**
  - » Lack of education in the market on how certain data should be pulled and importance of including certain data elements
  - » No requirements around what fields are mandatory
- **Inaccurate interpretations**
  - » Location data is translated into a Lat/Lon coordinate. This data could come from a variety of sources as granular as GPS or Wifi data, or can be as broad as IP / User Registration data
  - » The number of digits in a Lat/Lon coordinate are not necessarily a reliable measure of it's accuracy:

Lat/Lon Truncation - The green points represent single ad requests sent from devices that have not been identified as fraudulent. The red boxes represent Lat/Lon coordinates that have resulted from a bad publisher truncating all digits after the fifth decimal place.

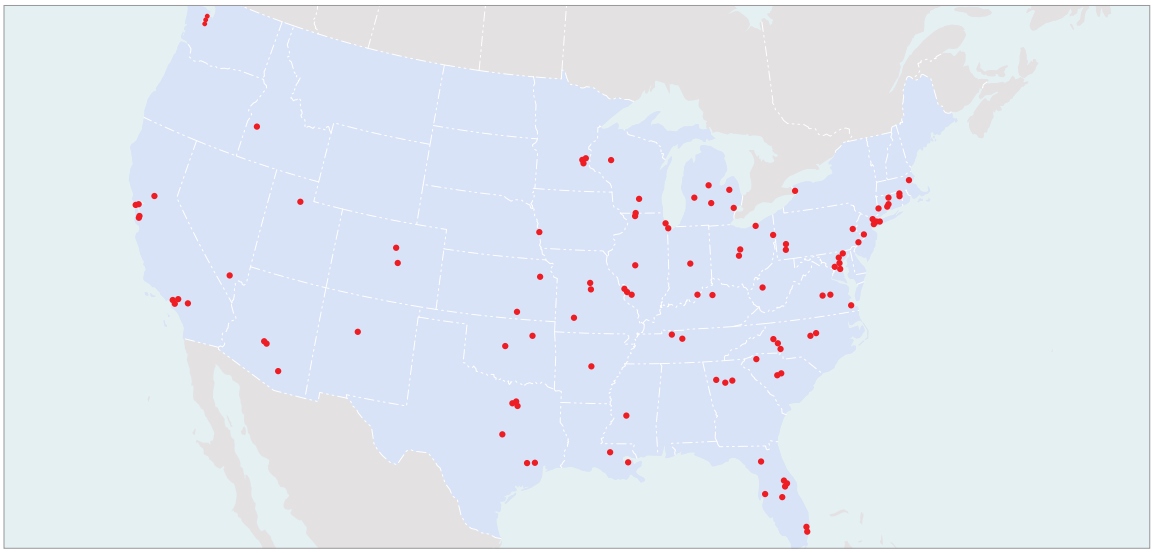
- Freshness of data: publisher may leverage location data from other apps or data pulled in the past
- Urban density: location signals and precision of data can be skewed in dense urban environments (e.g. multi-story buildings)
- Speed: Whether the device is static or moving, and if it's moving, how fast is it travelling?

- **Mistakes can happen**
  - » Only having permissions for “coarse” location tracking (as opposed to “precise”) from the OS and passing that data into the exchange.
  - » Transposed coordinates where the Lat/Lon are incorrectly inputted.
  - » Default values used by developer until a location is retrieved:



On the left is a publisher distribution indefault location used on app startup, defaulting to the center of the U.S. On the right is the same publisher distribution without the default location requests, which shows a normal distribution.

- **Publisher appending additional data**
  - » Centroid processing: publishers often derive the center-points of geographic areas such as zip codes, DMAs, or states and send as the real-time proxy for a user’s location



Unique Lat/Lon coordinate pairs corresponding to metro centroids that receive more than ten million requests per day from a very wide variety of devices.

Current Solutions

The location information and data which is currently available for mobile marketers can provide excellent value, even though there is room for greater accuracy in order to reach increasingly more granular data. All data used today to inform media decisions - whether it is traditional media ratings system, online cookies, survey based audience data etc - is not without its flaws. Through the power of mobile and the the ability to capture this powerful new location data (e.g. device location data), advertisers are getting closer to a real-time data set that can truly transform the way that data is used. By using real-time data, marketers are able to reach and understand audiences, but also to provide a way to measure the offline impact of media efforts.

Location marketing platforms, built with the challenges of mobile in mind, have taken big strides towards increased location accuracy despite the inherent issues within the ecosystem today.

Let's revisit the common location uses described above and review steps taken by the ecosystem to ensure quality.

- **Proximity Targeting, Location-based Creative:** By utilizing big data and pattern recognition technologies, most inaccurate locations, centroids and IP-based traffic can be filtered to allow for effective targeting and creative optimizations. In this way, brands are able to target end users based on their proximity to points of interest such as their own retail outlets, or even competitors' stores, and can adapt their ads to reflect different messaging at different locations.
- **Location-based Audiences:** In addition to applying big data methodologies to filter out inaccurate locations, and by analyzing historical location data as well as first and third party data sets, advertisers can further refine the information about a given device to provide useful insights and determine patterns or place usage. This data helps to build a picture of how any given location is used in real life (e.g. whether a place is residential or commercial) and provides an opportunity to target audiences based on their context and not just their physical location.
- **Attribution:** Though many different standards are currently being developed, location is increasingly being used to measure a mobile advertising campaign's effectiveness, including measuring foot traffic (either with direct integrations or by applying a statistical model) or integrating real time point of sale data (e.g. coupon redemptions).

With the appropriate measures in place to ensure the quality of the data currently available, location information becomes the crucial component for mobile advertisers. As this useful data is leveraged by location platforms and brands alike, it becomes imperative that all industry stakeholders contribute to setting standards and defining best practices, in order to ensure clean, consistent data.

More Quality Efforts Underway

With this comprehensive mapping of the location data ecosystem, we have provided mobile marketers with a clearer view of the challenges faced by both buyers and sellers when responding to calls for greater transparency and accuracy.

Though challenges do exist, it is important for marketers to understand the value of data that is currently available. By applying the correct methodologies, it is possible to refine the data to improve its accuracy. Additionally, coarse location data provides better location targeting capabilities than standard web based methods (e.g. targeting by IP address).

As we move forward to supporting industry solutions for the identified issues, the MMA Location Accuracy Committee will focus on improving current standards and recommending best practices including:

- Advocating for updates to the current OpenRTB standards to include relevant data sets that enable better evaluation of location data
    - » Specifically to include OS level “Accuracy” and “Freshness” as additional fields in the Geo object
- Working with the MRC to develop official standards for location data transparency
- Partnering with industry groups to drive adoption of these standards among publisher partners
- Continuing our drive to educate both buyers and sellers of mobile media

The MMA Location Accuracy Committee includes stakeholders from across the mobile advertising industry and the co-operation across the board will contribute to a formal set of standards going forward.

If you would like to learn more about the Committee’s work to improve location accuracy on mobile, or if your organisation would be interested to get involved with this important work, please contact us at: committees@mmaglobal.com.

MMA Overview

The MMA is the world's leading global non-profit trade mobile marketing association comprised of more than 800 member companies, from nearly fifty countries around the world. Our members hail from every faction of the mobile marketing ecosystem including brand marketers, agencies, mobile technology platforms, media companies, operators and others. **The MMA's mission is to accelerate the transformation and innovation of marketing through mobile, driving business growth with closer and stronger consumer engagement.** Anchoring the MMA's mission are four core pillars; to cultivate inspiration by driving the innovation for the Chief Marketing Officer; to build the mobile marketing capabilities for the marketing organizations through fostering know-how and confidence; to champion the effectiveness and impact of mobile through research providing tangible ROI measurement; and advocacy. Additionally, MMA industry-wide committees work collaboratively to develop and advocate global best practices and lead standards development.

Mobile Marketing is broadly defined as including advertising, apps, messaging, mCommerce and CRM on all mobile devices including smartphones and tablets. Members include: American Express, AT&T, Colgate-Palmolive, Dunkin' Brands, Facebook, Ford Motor Company, Foursquare, Google, Group M, Hewlett Packard, Hilton Worldwide, iHeartMedia, Johnson & Johnson, Kellogg Co., MasterCard, McDonalds, Mondelez International, Inc. Pandora Media, Pinterest, Procter & Gamble, Razorfish, R/GA, Starcom Worldwide, The Coca-Cola Company, The Weather Company, Unilever, Visa, VEVO, Vodafone, Walmart, xAd and many more. The MMA's global headquarters are located in New York with regional operations in Europe/Middle East/Africa (EMEA), Latin American (LATAM) and Asia Pacific (APAC). For more information about the MMA please visit www.mmaglobal.com.

Acknowledgements

The MMA wants to acknowledge the important contributions of all of the following companies, all of whom comprise the MMA Location Data Accuracy Working Group:

4INFO	OpenX
Criteo	Placecast
Dstillery	PlaceIQ
Factual	Reveal Mobile
Foursquare	The Mobile Majority
Future Of Privacy Forum	Thinknear by Telenav
Google	Ubimo
ImServices	UpSnap
Joule	Verizon Wireless
MEC	Verve Mobile
Media Rating Council (MRC)	xAd
Mobiquity Networks	YP
Near	Zebra Technologies

We want to especially thank the following individuals for their dedicated efforts on this paper:

- Monica Ho — Head of Marketing Infrastructure, xAd — Committee & Working Group Co-Chair
- Michael Lieberman — CEO, Joule — Working Group Co-Chair
- Vikas Gupta — Director of Marketing, Factual — Committee Co-Chair
- Karl Brautigam — Manager — Sales Development & Brand Strategy, Foursquare
- James Meckley — CMO, Mobiquity Networks
- Matthew Davis — VP Marketing, Reveal Mobile
- Brett Kohn — VP, Marketing, Thinknear by Telenav
- Gilad Amital — COO, Ubimo
- Ran Ben-Yair — CEO, Ubimo
- Yaffa Finkelstein — Director of Marketing, Ubimo
- James Kupernik — VP of Operations, UpSnap
- Jeffrey Goldstein — Director of Programmatic, xAd
- Pravesh Katyal — Product Manager, xAd