Challenging Misconceptions in Location Data Science: What Marketers Should Really Be Looking for in Real-World Signals

BY ALISTAIR GOODMAN, GENERAL MANAGER OF EMODO

The places we go reveal an enormous amount about the things we are interested in. They create a map of our universe, including our preferences and behaviors, what we like to do, our personality traits, our needs, the things we are likely to buy, and more. Smartphones have given advertisers unprecedented access to this universe. By finding users in a given place and time, advertisers can deliver ads that are not only relevant to their lives, but to that specific moment.

However, location targeting is still a fairly new technology and comes with its own set of challenges. There are a lot of misconceptions in the industry about user location data, and this leads to badly targeted campaigns and fraudulent ad inventory. Another challenge is that fabricated location data is a pervasive problem in the programmatic advertising space. Many companies are passing off second-class data to marketers as something much more valuable than it really is.

In order to run effective ad campaigns and achieve positive ROIs, advertisers first need to understand what the precision of location data inventory is and what it means. Let's take a deeper look at how user location data is sourced.

[Note: This is the first in a series of brief articles looking at the different components of how to effectively and accurately obtain, filter, profile and utilize location data to run successful mobile ad campaigns. Today we cover the different types of location data. Other articles will cover accuracy and precision, the role of Point-of-Interest data, different ways of constructing audience segments and attribution.]

HOW DO LATITUDE AND LONGITUDE COME INTO PLAY?

Each of the above sources help to pinpoint a user's latitude and longitude. Latitude and longitude coordinates point to the place of a location sensor and device on a map, and thus to where that user is located in the physical world. Their coordinates are based on a grid system where the squares become smaller with each decimal added. One more decimal digit on both coordinates increases the resolution by 10x10, meaning 100X resolution; the more decimals, the more precise the location.

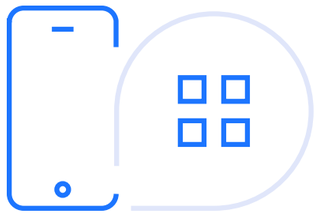
For example, the first decimal digit is worth about 11km and can distinguish the position of one large city from the

next. The third decimal digit is worth up to 110 meters and can identify a large agricultural field, campus, or city block. This is considered "privacy-safe." The fourth decimal digit is worth up to 11 m and can identify a parcel of land, and is the typical accuracy of an uncorrected GPS unit with no interference.

Lat/long is useful for locating someone but not sufficient to know "where" someone is. Lat/longs then need to be matched to the context, what is at or near that spot — such as a store, park or transit hub. Making all of these associations accurately at scale is not a trivial problem — get it wrong and a consumer looking at their phone ends up in a place that doesn't exist.

WHERE DOES USER LOCATION DATA COME FROM?

User location data comes from a variety of sources. One of the most widely used methods is via an app, which is collecting from the device itself. For example, users opt-in to allow apps like Uber, FourSquare, and Yelp to know their location. This user location data comes from the GPS chip embedded on mobile devices. While this data can be quite precise (down to less than 30 feet), it doesn't work indoors, when you are inside a mall or store (more on that later), and may require up to 30 seconds to find the satellites to get a "fix". Key to collecting location via this method is having an app that is either in use or running in the background (with associated battery drain) and the consumer must have PROVIDED consent for the use of their location.



User location can also be extracted from mobile operator networks. This method relies on cell towers to derive the location of the user by identifying which tower(s) the device is connected to. While this is less precise than GPS it has the advantage of being ubiquitously available for all subscribers on a network, and no app is required. Wi-Fi or Over-The-Air (OTA) signals can be used in a similar fashion. When a user connects to a Wi-Fi network, they are near the specific place where that hotspot is. Finally, user location data can come from beacon technology. Beacons use Bluetooth Low Energy (BLE) to transmit a signal that connects with smartphones equipped with a compatible app or operating system when at close range. This not only collects specific indoor-location data which GPS doesn't do well, but also enables businesses to push messages or offers to users when they come in range.

WHAT ARE THE PROS AND CONS OF DIFFERENT METHODS FOR SECURING USER LOCATION DATA?

Each method for securing location data has varying degrees of accuracy and precision. The reality is that no single source of user location data is sufficient for an accurate service – and multiple sources need to be combined to create the scale that advertisers need for audience segmentation, campaign delivery and measurement. Carrier location methods are useful in all types of environments (rural, suburban, urban, dense urban, indoor) and have a range of 50 m to over 1 km. GPS and Wi-Fi both have similar ranges (5 to 50 m) although GPS is a better source in rural and suburban areas, and Wi-Fi better source in indoor and urban environments. Bluetooth's range is up to 10 meters and is only useful in indoor settings.

These methods involve a trade-off between precision and accuracy. Precision reflects how closely you are able to pinpoint someone's exact location. Accuracy, on the other hand, reflects having a sufficient number of observations of a user in a place to draw a conclusion about behavior, and being able to verify that it is indeed a real consumer (and not a fraud-generating bot artificially

creating or adding to lat/long coordinates). For example, getting one fix for a user at a Dunkin Donuts at 8AM is interesting, but you will need dozens of locates most weekday mornings between 7:45-8:15 in the area to say that it's highly likely that a consumers stops for coffee and a donut most mornings on the way to work. Per this example, you actually need both a few precise fixes and a lot of accurate data to be sure that this is a coffee-and-donut consumer.

The goal for targeting users on mobile devices is to achieve both high precision and high accuracy at scale. This does not mean you need to know where a person is to the sixth decimal point (that is the size of your hand). In fact, precision alone is the wrong topic to focus on. The more important consideration is 'user location-reach' – a significantly large enough set of observations about where an anonymous user is and when they are there from lots of different sources, to be able to deliver rich insights about behavior. It is easy to locate someone once at a precise spot – the real questions are a) are you sure the location data is not fraudulent, b) can you associate the user with their context, c) are you seeing them often enough to say with certainty that they belong in a particular segment, and d) what is the benefit from a privacy perspective to collect and store really precise data about consumers. Systematically answering all of these will yield larger, more reliable audiences for advertisers and ultimately a better performing mobile campaign.

Location Precision

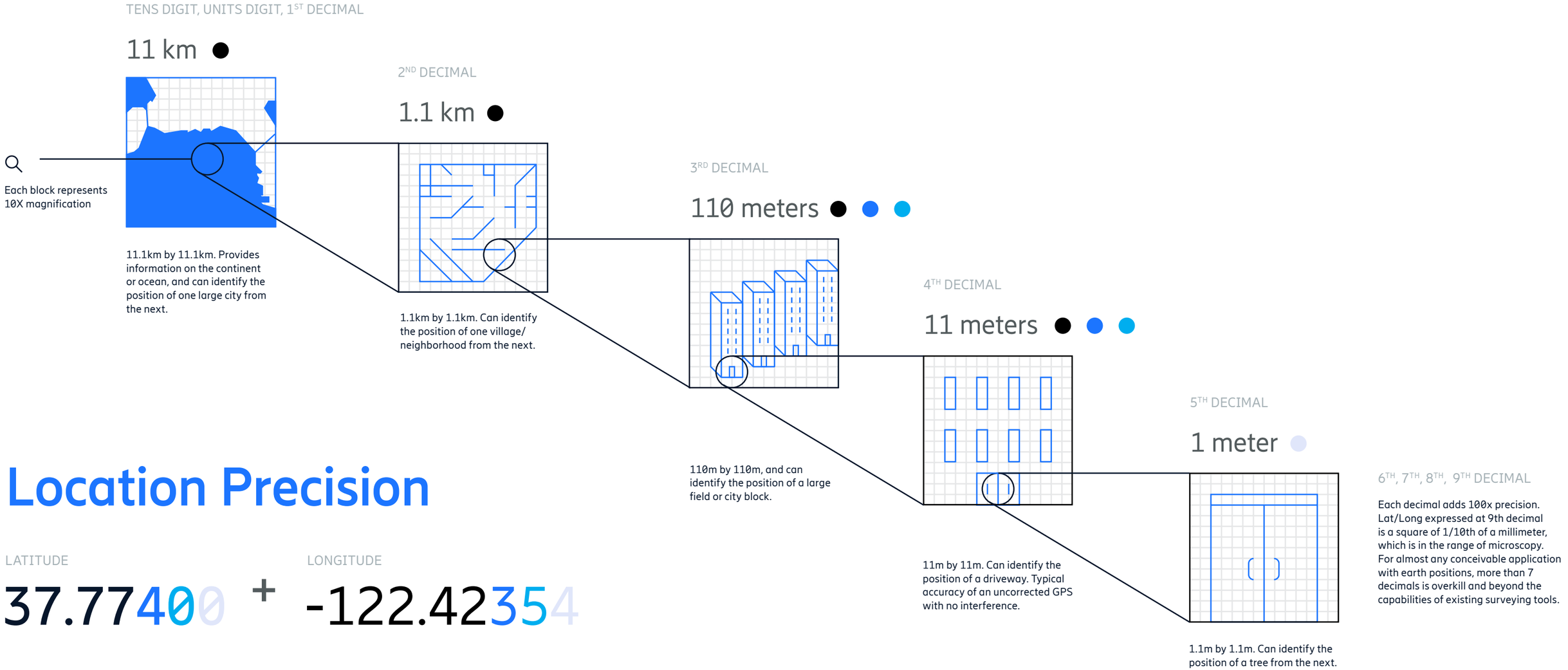
LATITUDE

37.77400

\+

LONGITUDE

-122.42354



CARRIER

Uses cell towers to determine the mobile device's location, based on the relative distance to the nearest cell phone tower

PROS

- Does not require an app, works for all phones all the time
- Best for dense, urban areas
- Privacy safe

CONS

- Precision depends on nearest cell tower
- Does not work well in rural/remote areas

WI-FI

Uses active connection to a Wi-Fi hotspot

PROS

- More precise that a cell tower signal
- Does not require user to turn on an app
- Best for indoors, dense, and urban areas

CONS

- User needs to be logged into a Wi-Fi hotspot

GPS

Uses a satellite-based system that triangulates the Lat/Long of the device

PROS

- More precise than cell tower or Wi-Fi
- May be persistent if the user has opted to share this data

CONS

- Users need to have GPS turned on to share location actively
- Does not work indoors, building structures interfere with the signal
- May drain battery
- Has privacy implications due to level of precision

BLUETOOTH/BEACON

Uses a satellite-based system that triangulates the Lat/ Long of the device

PROS

- Highly precise, but needs device to be within 10 meters
- Best for indoor use

CONS

- User needs to turn on Bluetooth
- Has privacy implications due to level of precision

EMODO INSTITUTE