The HFT Guy

A developer in London

TECH

What Does It Really Take To Track A Million Cell Phones?

19 JULY 201716 JULY 2017
THEHFTGUY
2 COMMENTS

i
18 Votes

You can find anything and everything on the internet, yet nothing that explains how to track cell phones.

Let us clarify right away, we are not talking about how to track your own cell phone in case it's lost or stolen. We are talking about tracking everyone that lives, breathes and wears a cell phone.

This is actually incredibly easy and we think that people should be aware of that.

If a representative of a phone service provider with 10 million customers came into my office and asked this question "What would it take to track every move of our 10 million customers?". My answer would be "A*n intern and 6 months*". Then we'd insist the intern will need a desk, a computer, basic programming and algebra skills. That's all it takes.

Imagine for a minute that you are the intern in question. Congratulations and welcome to our company! Your internship begins now, this document will introduce you to everything you need to know.

We'll go over the basics of cellular networks, geolocation principles, technologies readily available in every cell phone and how to leverage all of that into a truly real-time planet-scale mass surveillance system.

Spoiler Alert: If you are scared of 1984 like scenarios, you may want to stop reading this and bounce to a video with Darth Vader playing the accordion (https://www.youtube.com/watch?v=BgAlQuqzl8o).

A) Foreword

We are in a unique position with cross domain expertise. We combine experience in state-of-the-art tracking systems with past experience in the telecommunication industry.

Whether it's locating an item in a warehouse, guiding people inside a shopping mall or following stolen trucks. There are many legitimate use cases for tracking with as many constraints to satisfy: indoors, outdoors, with or without battery, variable precision, etc…

A phone itself comes with numerous technologies built-in: GPS, WiFi, accelerometer, compass, etc…

We'll focus exclusively on what is needed to achieve easy, effective, reliable, mass-tracking.

B) Requirements

We want to track cell phones. Which one? ALL OF THEM.

Some constraints:

- Cell phones are out of control
  - No physical access
  - Hardware cannot be modified
  - Software cannot be installed
- Users are out of control
  - They will not perform any wanted action
  - They will not opt-in to anything
  - They will not consent to anything
- Must be scalable to millions of cell phones
  - Self-explanatory

Better precision in time and position[1] is better but does not constitute a goal by itself. It has to be balanced against more important parameters like feasibility, scalability, reliability and costs of operation.

For the avoidance of doubt, we'll call the project an utter success if we find ourselves able to pin point any cell phone being in a specific block inside a specific city, at a specific hour.

[1] A location is always a position AND a time together. It's important to keep the two dimensions in mind.

C) Multilateration

Most systems work by "*triangulation*". It's possible to triangulate a specific position by comparing some measures to some points of reference. First things first, that's actually called multilateration.

If you use a service like a GPS, it does all the work and gives out a position with a radius of error.

If you do the hard work yourself, either you are the guy making the GPS or you are trying to mix multiple sensors in a creative way, you need to do the hard work yourself.

Ultimately, it always comes down to 4 methods.

1) Power: Signal power

With information about the transmission power, the reception power and the medium. It's possible to use physics wave propagation formulas to estimate the distance traveled.

In practice however, this method is extremely unreliable for radio waves, so you NEVER want to use that.

For instance, it's typical for a long distance radio wave to go up and down 10 fold (+-10 dB) within a single second. It changes all the time and that's when you are not moving. It gets worse when walls, windows and your head goes in and out of the track.

2) AoA: Angle of Arrival

Note: It's called *triangulation* when using angles.

With the angle of a signal, it's possible to determine that the source is within a line (or a cone). Obviously, it works better with highly directive signals.

You can surely picture a rotating radar like you've seen a thousand times in movies.

3) ToA: Time of Arrival

With the time and the speed of a signal, it’s easy to determine the distance. t = d/s.

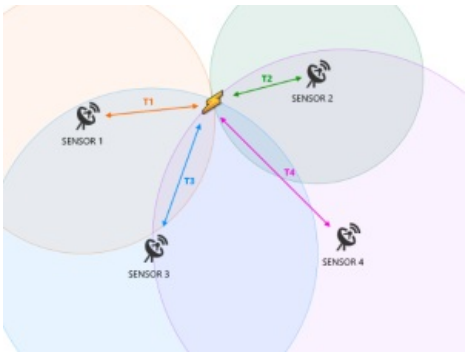
Challenge: Radio waves travel at the speed of light 299 792 458 m/s.

To measure a distance with 30 cm accuracy requires to measure the time with +- 0.000000001 seconds (1 nanosecond). That is a hard problem.

4) TDoA: Time Difference of Arrival

Also based on time measurement.

It’s possible to use time differences instead of an absolute time.



The item to be tracked emits a pulse that is received by multiple receivers (Picture Source: Locating Lightning Strikes (http://www.microsferics.com/index.php/How_it_works))

The item to be tracked emits a pulse that is received by multiple receivers. The receivers are at known locations and synchronized in time.

By measuring the time difference between the reception of the signal at the receivers, it’s possible to determine the relative distance of the source to the receivers.

Challenge: It doesn’t only require to measure time with crazy precision but also to synchronize clocks across systems.

D) Cellular Networks Principles

We’ll go through some basics about cellular networks.

1) Base Station (BTS)

A cell phones communicates with a base station.

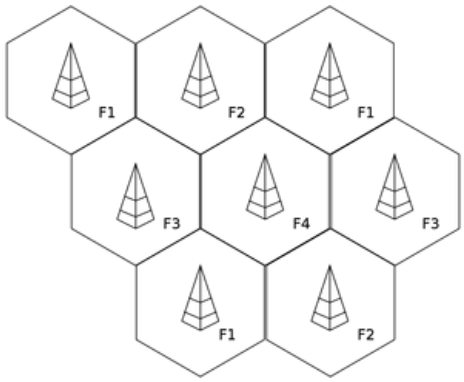
There are two channels. One for emission (to the BTS), one for reception (from the BTS). They operate at different frequencies.

The emission channel (to the BTS) is shared by all devices. At any time, there can only be one device emitting.

2) Cellular Network

A BTS covers an area around it. Adjacent BTS form a cellular network.

Two adjacent BTS need to have different frequencies to avoid interference.



Cellular Network

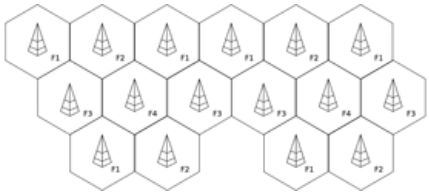
Each operator runs its own network. It may share or resell network service to other operators.

Some operators are virtual (called MVNO (https://en.wikipedia.org/wiki/Mobile_virtual_network_operator)). They have no physical infrastructure, they exist on top of another provider. For example, giffgaff (https://www.giffgaff.com/) [1] runs on top of O2.

[1] Highly recommended provider in the UK.

3) Cell Density

A base station can only cover a limited amount of users. What happens when there are too many users, like in a city center instead of a village?



Double the density. Quadruple the capacity.

Trivial, cells can be arranged more densely to increase the capacity.

E) Locating A Cell Phone

We saw the basics of cellular networks and the basics of multilateration.

1) Base Station

Your phone has to be in range of a BTS to work. By the simple virtue of having your phone "*online*", the operator knows that you are within the range of his station.

As we said before, the density of towers can be adjusted to accommodate the density of users.

A tower has a theoretical range of up to 35 km radius. In a major city, there could be one every km; in the empty country side, there could be one every 10 km.

That's enough to locate a phone down to one city.

BTS have to be located carefully to manage their coverage and not jam one another. An operation knows the locations of its BTS. They have to be registered officially to some sort of radio tower registry (the execution varies slightly by country).

P.S. We would like to give some free sites where you can see BTS but they tend to not live long. There is value in providing a good database so it's never given for free (and if it does, someone will realize their mistake soon).

2) Base Stations x 6

Back to when we were in telecom, a long time ago, we had special test phones provided by the manufacturers.

Think of an old school Nokia phone, except it comes with build-in hardware and software for debugging purpose. One of the build-in tool shows detailed connectivity information, that are otherwise not available to consumers.

With that at hands, we can see that the cell phone, right in ours hands, is able to detect and maintain connectivity with 4 towers simultaneously, at all times.

Why 4? Because there are 4 in our area. The phone could do more!

A $50 cell phone, even one from a decade ago, can be simultaneously "connected" to 6 stations. This may include stations slightly beyond range, having a signal just strong enough to be detected but too weak to be used for actual communications.

As we like to illustrate nowadays in simple terms: Your phone is a wonder of technology, it will go above and beyond to keep the communication going no matter what. When you talk, one word can go to one tower and the next one to another tower, switching as often as necessary.

On a related topic, this is why you cannot find cheap jamming devices against mobiles. Phones are intended to operate in a hostile environment with thousands of phones competing for the air. A jamming device is like a garden hose in a hurricane. It's physically impossible for any cheap pocket-size device powered by 2 AA batteries to out compete the hurricane.

To conclude this paragraph, your phone is constantly talking to multiple stations, not just one. Instead of being in a disk around a station, you can be located to the intersection of multiple disks. Handsome for tracking, not so much for your privacy.

More importantly, we need multiple points of reference to be able to perform multilateration. Here they are!

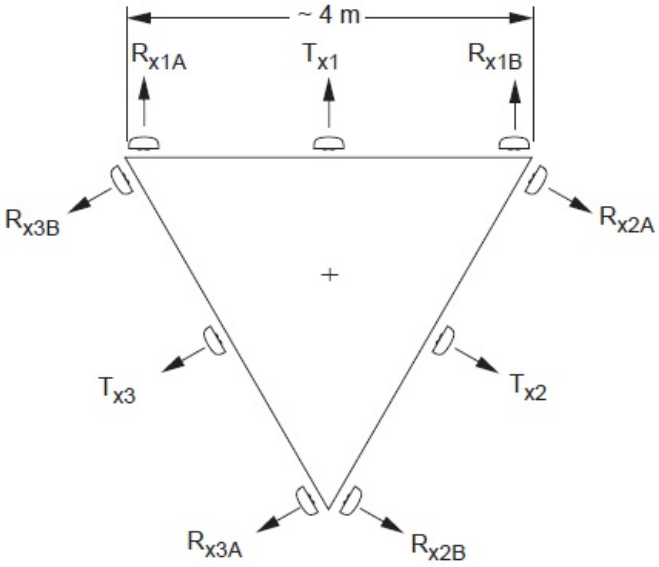
3) Angles

We said that a tower covers a radius around it. In practice, this is sub optimal so that's not how it's done.

Instead, a station is usually split in 3 independent beams of 120 degrees.

A typical base station (Source: Wikipedia (https://en.wikipedia.org/wiki/Sector_antenna))

A typical BTS. Notice the triangle shape, each face covering 120 degrees.



The arrangement of Tx and Tx. (Source: Kaithrein (http://www.kathrein.pl/down/BasicAntenna.pdf))

The technical setup, as recommended by a polish antenna manufacturer.

This allows to limit the positioning to 120 degrees. It’s actually very powerful, it just increased the accuracy a lot and allows for multilateration with only 2 BTS.

Geometry Trivia: The intersection of 2 circles gives 2 points (opposites to each other), it takes a third reference to find which point is the right one. Therefore multilateration always requires 3 references (e.g. the distances from 3 BTS). In practice, an angle is enough to do the distinction most of the time (e.g. angles and distances from 2 BTS).

This method requires information about antennas and directivity. We just checked one BTS database and it’s there so it looks like it’s not a problem to get. The precision will need to be tested in the wild (wave propagation and construction work are not perfect to the degree).

4) RSSI: Received Signal Strength Indicator

A phone emitter has a maximum power of 2 Watts (6 dB). A phone receiver has a typical sensitivity of 0.000000001 Watt (1 nW or -90 dB).

The air can attenuate a signal by a factor of 1 billion and your phone still works. Magic!

In a perfect world of undergraduate physics, the propagation loss in the air can be modeled with that equation.

$$L = 20\ \log_{10}\left(\frac{4\pi d}{\lambda}\right)$$

With L the loss in dB, lambda is the wavelength and d is the distance, lambda and d in the same unit.

In the real world, this doesn't apply at all. The air is not homogeneous and there are obstacles all over the place. The losses can vary by 2 orders of magnitude at any time (and it does). There is no meaningful value to be measured.

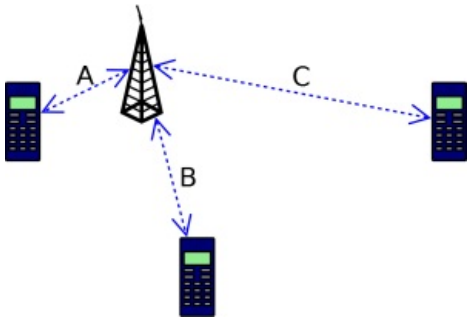
A good usage of Kalman filter (https://en.wikipedia.org/wiki/Kalman_filter) may help to filter the samples but that's both complicated and resource intensive for a mediocre result.

We've got much better to do than RSSI so let's not our waste time discussing that.

5) Timing Advance

A channel is shared between many customers, each one gets very short periods of time allocated. You can read an introduction to GSM frames (http://www.radio-electronics.com/info/cellulartelecomms/gsm_technical/gsm-radio-air-interface-slot-burst.php) for details.

The time slot might be unusable in the event of an overlap with the previous or the next slot (dedicated to another phone). One thing that could cause unwanted overlap is the propagation delay from the phone to the station.



The signal takes time to travel from a phone to the station. The delay depends how far the phone is.

Each bit is 3.69231 µs long in GSM, a radio wave can travel 1107 meters in that time. That means a phone located multiples of 1107 meters away will be multiple bits late… we don't want that!

The propagation delay is accounted for and corrected by a mechanism called the timing advance.

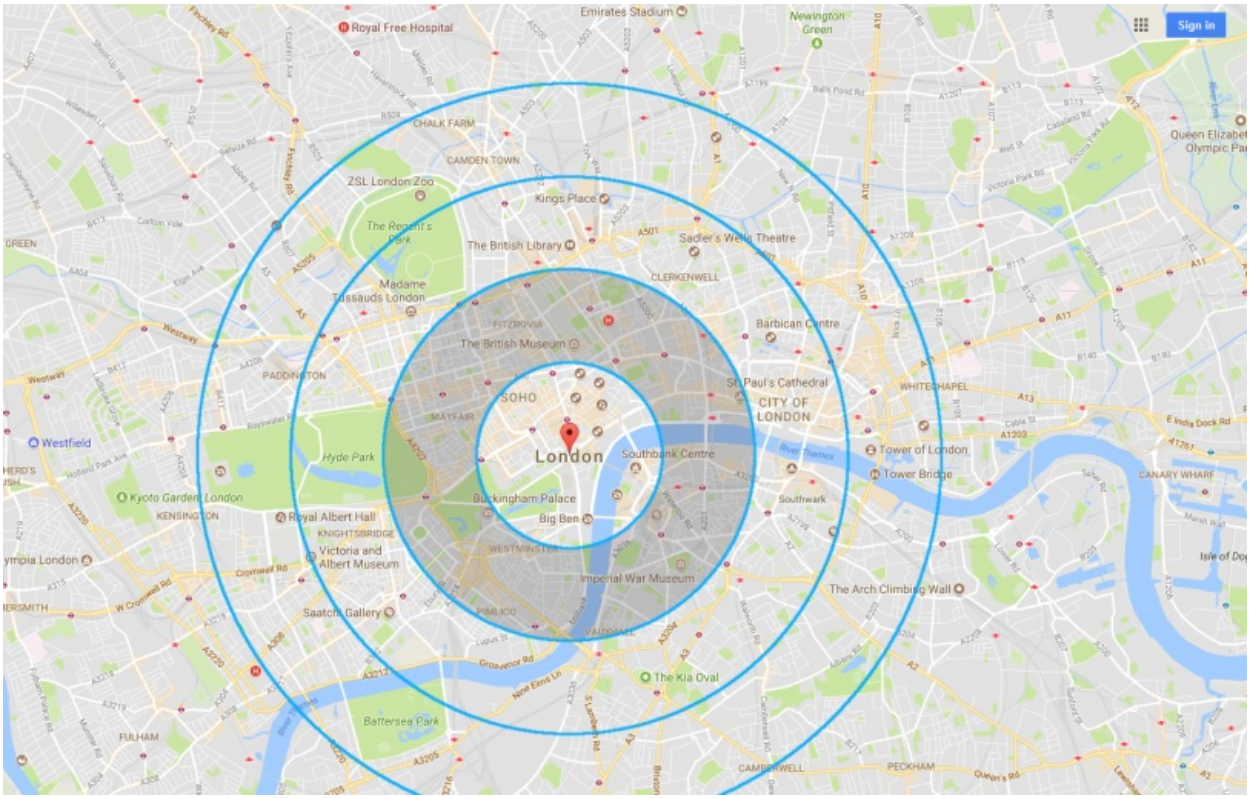
The base station measures how late messages arrive and sends a correction parameter, the timing advance, back to the phone.

It's a number between 0 and 63 indicating how much advance it should take, in multiple of 3.69231 µs.

For the purpose of geolocation, the timing advance allows to locate a cell phone within a 1107 meters annulus around the base station.

For the purpose of being a grammar nazi, the section of a disk inside a concentric disk is called an annulus (https://en.wikipedia.org/wiki/Annulus_(mathematics)).

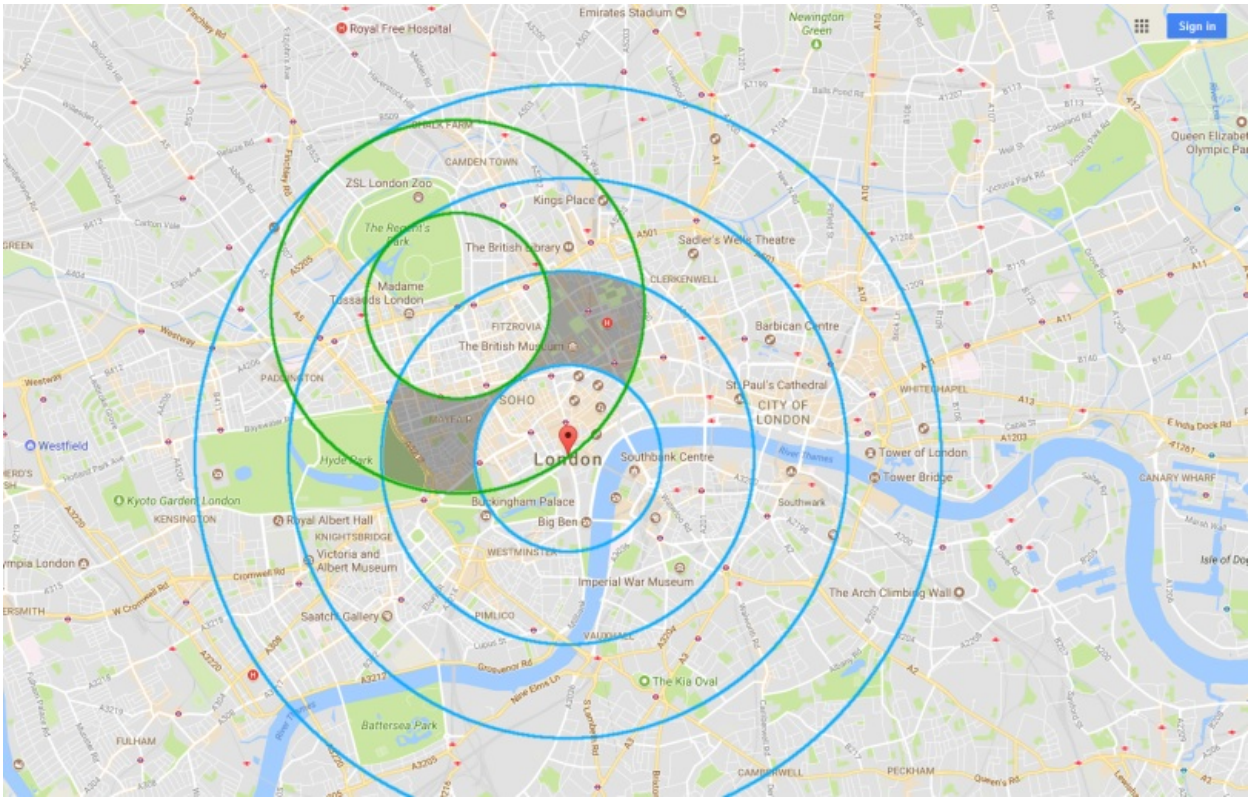
Let's see what this looks like if we put some circles on top of London.



Timing Advance Annuluses

That’s the accuracy a single tower can give with just timing advance (ignoring angles).

Let’s see what the intersection of two stations looks like.



Timing Advance with two stations.

That gives two possible areas. It takes a third measure to decide for sure (either an angle or a timing advance).

It’s intuitive enough. The more measures, the better.

Remember: Your cell phone is able to talk to 6 towers at all times, that can cooperate in tracking it.

It’s not always accurate but when it is, it can pinpoint you to the block you are walking in.

6) Geometry Quick Thoughts

Two dimensional intersections of disks[1] is high complexity both in terms of computational power and in terms of what a cheap intern might be able to understand.

Intersection of circles is a trivial problem though. There are known formulas (http://mathworld.wolfram.com/Circle-CircleIntersection.html) that can be computed in constant time.

It can be generalized to N circles by simply applying the formula to each pair of circles. Filter out the points which are not within the intended angle and distance from the station (a basic comparison in constant time[2]).

The resulting points show something that is approximate but quick and easy to compute. Remember that we have millions of people to track in real-time and only an intern for that!

Call for comment: Dear mathematician reader, please comment if you have any advice on how to find the intersection of complex shapes. [3]

[1] Strictly speaking, this should be treated in 3D. The world is a sphere. There are variations in terrains that should be accounted for, especially in mountain regions.

[2] Angles are trivial to play with in polar coordinates (or spherical coordinates).

[3] We checked how design software handle 2D and 3D intersections (SolidWorks, Catia, AutoCad). Sadly, it is advanced mathematics AND it takes a lot of computational power.

7) Summary

Locating a cell phone:

- A base station locates the phone inside its range (up to 35 km radius)
- The timing advances locates the phone in a 1107 meter annulus
- The angle splits locates the phones in a 120 degree section
- There can be many stations participating in the process
- They can be interpolated to improve the precision

8) Time

Remember that a position is always implicitly linked to a time. A phone is at a specific place at a specific time.

The phone wants to be connected in permanence. It is adjusting to the environment in real-time all the time. Typically, in a matter of seconds. It is mandatory for the phone to work (calls and messaging).

Being conservative, a phone should be able to be (re)located every minute.

Do the test.

Turn your phone off, send it a message, turn it on, how long to receive the message?

Put your phone in a tin box (to block signal), send it a message, take it out of the box, how long to receive the message?

F) Dependencies

There are some prerequisites to make that tracking system real and deploy it on a large-scale.

1) Base Station Database

The project requires a database of base stations.

Every provider know where they set up their stations, that’s part of the job of being a service provider. It’s a given if making the project as part of an ISP.

It should be easy enough to get a high quality database of base stations for anyone (not to confuse easy with inexpensive).

2) Logging BTS Information

The project requires access to BTS signal information.

First, there is an extensive authentication, roaming and payment system embedded in the network. This is necessary to provide service to the right user at the right time at the right price.

Second, almost every regulation in every country in the world require providers to save some usage information per user, for many years.

There is massive infrastructure already in place to log and audit accesses, down from the station, up to the high level customer subscription.

The values that are needed may or may not be saved already (Cell ID, TA, …), if they are not, they shouldn’t be very hard to add.

3) Matching Identities With Phones

Assuming that we track cell phones. The final step after a phone is located is to match that phone with the identity of a real person.

There is a whole authentication system made built-in the network. There are unique identifiers for customer contracts, sim cards, phones, etc…

Not sure the details of how this works and how this could be abused. Assume that an ISP can match any connected user with the subscriber.

G) The Known Unknown

We saw how to track every cell phone in service, easily done by the ISP of said customers (and by extension easily achieved by the NSA/GCHQ)

There are some unknowns that may affect the scale and the success of the operation. None that can impair it but some that can bring it up to a whole new level!

1) Near Range Tracking

A phone has to discover stations around it. It’s not possible to known which ones are right without trying.

Technically speaking, there is a possibility that the phone might have to *broadcast and try to link* to all stations in range [1].

If so, any station in an area would be able to follow any phone in proximity. National providers could track everyone everywhere since they are already cover the entire country. Rogue actors could setup dedicated networks for the sole purpose of tracking.

[1] It has to start with timing advance and authentication of the device, thus allowing for multilateration and user identity lookup right away.

2) Cross ISP Traffic

Have you ever been in an area with low reception where the phone displays "*emergency services only*".

There is no reception to make regular calls, yet it can make emergency calls, probably by using other networks (read: not the one you subscribe to). This is a legal requirement, cell and service providers have to allow that.

Technically speaking, it means that there is something built-in to allow cell phones to connect to anything through any network and your phone is trying that automatically all the time. (This is similar to the previous point).

If so, it can be abused to track your phone.

3) International Roaming

Ever been to another country? Your phone work just fine, except you're charged ten times more.

Again, this implies that the phone is connecting to anything. Better though, this implies that other providers are able to reach your current provider somehow, to confirm your access and incur your billing.

Depending on how it's done in the details, there may or may not be an opportunity to link a cell phone back to its provider and its owner, anywhere in the world.

H) The Known Known

1) Retro and Forward Compatibility

This works on all cell phones and it worked for decades.

The technology has been out and part of every cell phone at least since the first edition of GSM, circa 1991.

There is no change with 3G, 3G+, LTE. Still works like a charm!

2) This Project Can Be Done By An Intern

The technology itself is within reach of a 15 years old. Any student who attends telecom 103 is taught enough to come up with that (if only they listened instead of playing on their phones!).

20 years ago, this might have gone unnoticed or ignored. There were only a few stations and a few users. Limited accuracy, limited user impact. It's easy to imagine an early proof of concept that found it impossible at the time: "It's *gonna take an entire floppy disk to save the positions of 12000 customers! Oh my gosh. We'll never have the budget for that.*"

Nowadays, it's so trivial it's frightening. Any cell provider could take an intern and make it happen in 6 months. Gotta save some signal information? It's already done. Gotta do a bit of algebra? Nothing difficult.

3) Verizon Is Doing That Already

Feel free to read "*Verizon*" as any major phone provider.

Any service provider automatically gets incredible tracking capabilities and has to keep a history of it. It's not optional. The first half comes with the phone's infrastructure, the second half is mandated by regulations.

The core business of a provider is to provide phone service though, not to locate all customers in real-time down to the minute. There is no reason to perfect the techniques written in this document.

4) The NSA Is Doing That Already

Feel free to read the "*NSA*" as any state sponsored actor.

They want to track every people in the world. That's one of their main goals. They have lots of resources dedicated to do just that. They have the ability to infiltrate providers and/or to deploy their own rogue infrastructure.

Ironically, the most awesome mass surveillance system ever invented is out there already and quite easy to use.

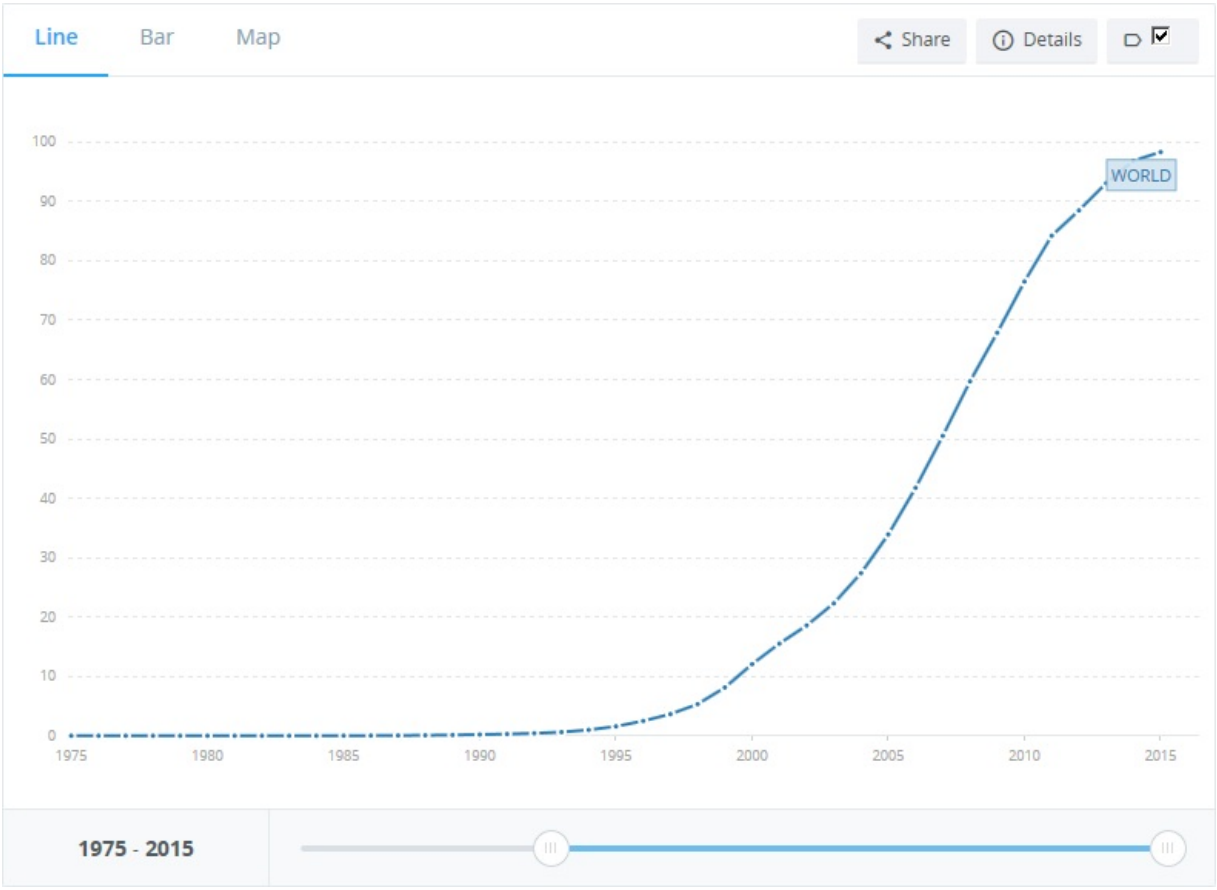
What are the odds that they figured it out? I'd say pretty high.

Conclusion

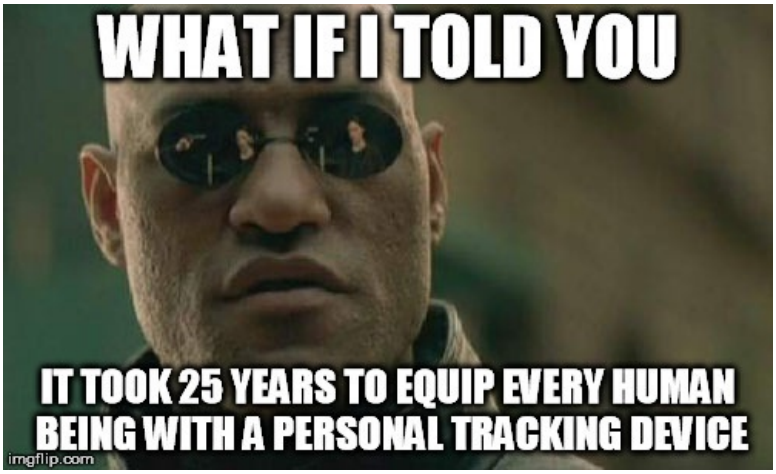
(https://thehftguy.com/2017/07/19/what-does-it-really-take-to-track-100-million-cell-phones/nokia-3310/)
(https://thehftguy.com/2017/07/19/what-does-it-really-take-to-track-100-million-cell-phones/iphone-7s/)

What's the difference between a Nokia 3310 and an iPhone 7?

There isn't any! As long as they are turned on, they can both locate you in real-time, 24/7, with a precision better than 1 square kilometer



Mobile Cellular Subscriptions per 100 people (Source: The World Bank (http://data.worldbank.org/indicator/IT.CEL.SETS.P2?end=2015&start=1975))



...and we made them pay for it!

GEOLOCATION, MASS SURVEILLANCE, TRACKING CELL PHONES, TRILATERATION

2 thoughts on “What Does It Really Take To Track A Million Cell Phones?”

newt0311 says:
19 JULY 2017 AT 16:05
Could the speed problem be taken care of by storing the raw signal data and only computing the position data on demand? Because there’s a lot more than just triangulation that could be done. This seems like the ideal use-case for something like Stan: build a bayesian probability model for the person’t movements and what the resulting signal data would look like to get a very precise probability rendering for their position. Computationally expensive but effective.

REPLY

Sven says:

19 JULY 2017 AT 17:50
Regarding the database of BTS positions, check out wigle.net

REPLY