# A brief history of wi-fi privacy vulnerabilities

robertheaton.com/2019/01/15/a-brief-history-of-wi-fi-privacy-vulnerabilities

15 Jan 2019

Your good friend, Kate Kateberry, used to work as a technological consultant to dictators. She travelled the world, teaching repulsive regimes the value of a good wi-fi surveillance network. She built them systems that collected the wi-fi signals emitted by their subjects' smartphones and that used this data to track their entire population's locations in real-time. Her clients used her work to keep their dissidents downtrodden and to show them precisely targeted online ads for consumer goods.

A few months ago, after one pang of conscience too many manifested itself as a serious heart episode, Kate retired from the despot-device-tracking business. Nowadays she only uses her powers for good, and occasionally for pranks on sufficiently deserving victims.

Which brings us to your mutual friend and mortal enemy, Steve Steveington.

After many years of trading gentle, good-natured pranks, this time the Stevester has gone too far. He managed to convince both you and Kate that Hobert Reaton (serial sketchy adtech entrepreneur and founder of both WeSeeYou and I Might Be Spartacus) was starting a new company, and was looking for thoughtful investors interested in making huge, guaranteed, and immediate returns on their life savings.

As you now know, this was nothing more than a ruse, designed to trick you and Kate into transferring all of your money to the general fund of the North Korean government. You are now both almost entirely broke, and Kate has been indicted by the Federal government for economic sanctions breaking. The two of you have sworn a terrible revenge on the Steveington responsible for your downfalls.

At first Kate just wants to kill him - keep things simple. You try to explain to her that that's not how you normally handle these kinds of things. Eventually you compromise.

You will not kill Steve - for now - but you will turn his life into a real-world Truman Show. You will set up a city-wide array of wi-fi monitoring devices called Pineapples. You will combine them with Kate's world-class knowledge of privacy vulnerabilities in the wi-fi protocol in order to track Steve and his smartphone's every move. You will follow him as he shambles from his home, to work, back home, and then hopefully onward to somewhere embarrassing and ideally illegal. You will broadcast these movements live at `thesteveingtonshow.com`. If all goes well then the show will become a cult hit and Steve a world-wide laughing stock. You will then reveal to him what terrible revenge you have wrought over coffee and bagels. You will fund the venture through an illegal gambling operation based on Steve's life and by selling weird, inappropriate merchandise with his stupid face on it.

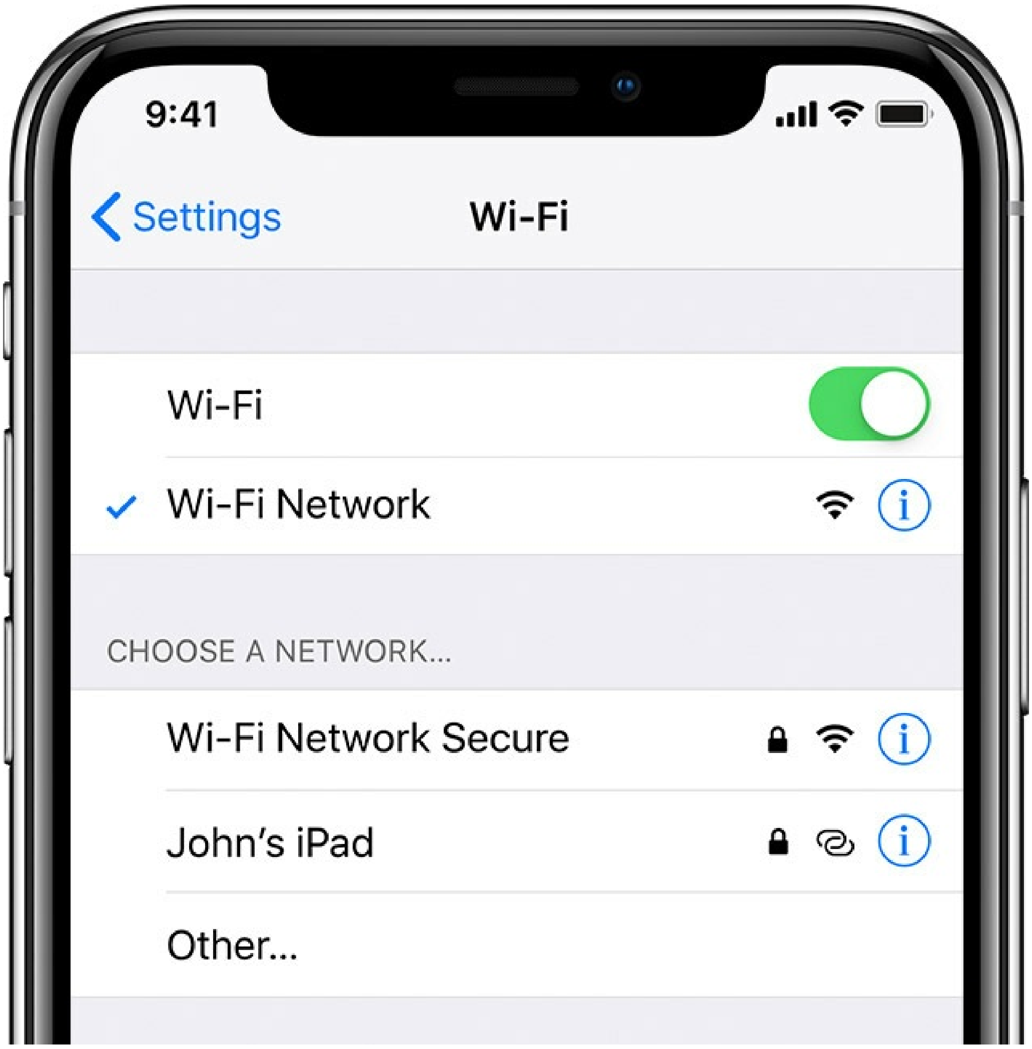
Before you start nailing down the technical details of the Steveington Show, the Katester runs you through a brief history of wi-fi-based tracking attacks and how she used them during her consulting years. After she has outlined the first few attacks, you quietly pull out your battered iPhone 4, still running an unpatched, unsupported version of iOS7, and gingerly put it into airplane mode.

## An introduction to wi-fi

In order to identify and connect to each other, wi-fi enabled devices like smartphones need to belch out an enormous amount of near-continuous chatter. Wireless communication is intrinsically public and interceptable, and so some amount of information leakage is unavoidable. This makes the wi-fi protocol fraught with potential privacy pitfalls.

To start with, smartphones need to maintain an up-to-date list of nearby wireless networks that they can connect to. They seek out nearby networks in the same way as overnight visitors to your house do - by constantly and repeatedly shouting "HEY HEY ARE THERE ANY WI-FI NETWORKS AROUND HERE I CAN GET ON?". A smartphone spews out this demand every couple of seconds in the form of a *probe request*. When a wi-fi router receives one of these requests it responds with its own *probe response*, which contains its network name or *Service Set Identifier* (eg.

`StarbucksGuest`) and other metadata. The smartphone picks up the probe response, and adds the service set identifier (SSID) to the list of available networks that it displays to its user. There is a further sequence of call-and-response messages that the smartphone exchanges with a router in order to actually connect to it, but we need not be concerned with them here.



Wi-Fi

Wi-Fi ●

✓ Wi-Fi Network

CHOOSE A NETWORK...

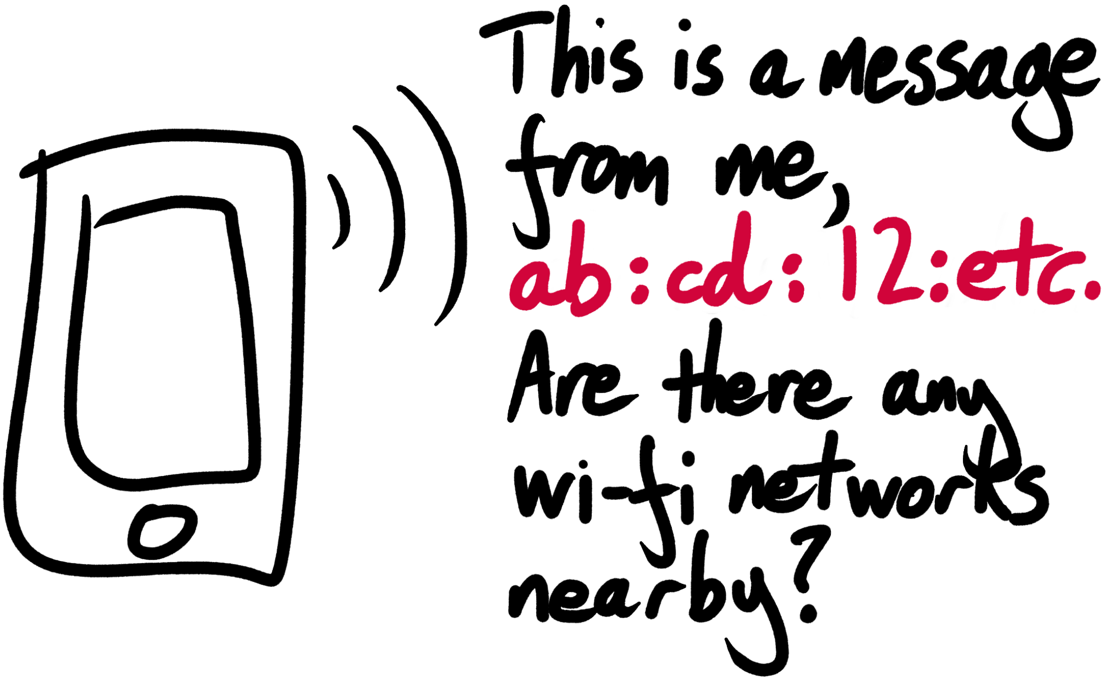
Wi-Fi Network Secure

John's iPad

Other...

Request probes are intended for routers, but can be picked up by anything or anyone. Short of disabling a device's wi-fi mode altogether, there is no good way to prevent trackers from hoovering up request probes and using them detect when a device is nearby. There isn't even any good way to prevent them from using *triangulation* (see below) to locate a device to within a few meters of accuracy.

The key battle in wi-fi privacy is therefore not whether users can go completely undetected, but whether they can remain anonymous. Can they blend into the crowd and prevent their activities from different times and spaces from being linked together into a much more detailed, holistic, and troubling profile?


This is a message from me, ab:cd:12:etc. Are there any wi-fi networks nearby?

The historical answer is no. The biography of the wi-fi protocol is lousy with privacy vulnerabilities that could have allowed tracking companies and attackers to keep track of devices and their users across large swathes of time and space. For the last 5 years or so, smartphone manufacturers have been working quite credibly on patching these goofs. Unfortunately for the Stevester, this doesn't mean that he is safe. As Kate will now demonstrate, manufacturers haven't always been fully successful in their fixes, and researchers in both industry and academia are forever developing new attacks.

Kate starts your history lesson with the basics.

## 1. Tracking devices by their MAC address

A device's MAC address is a unique identifier that it uses to describe itself whilst it communicates on a network. If connected devices ever develop consciousness and social media presences, MAC addresses will be their Twitter handles.

Before 2014, smartphones included their real MAC addresses in their request probes. As a consequence of this seemingly reasonable design choice, every smartphone in the world was constantly and publicly announcing a persistent and unique identifier for itself, and doing so in a manner that could trivially be intercepted by anyone in possession of some very basic wi-fi technology.

This made Kate Kateberry's job of identifying and tracking individuals and their devices very easy. All she had to do was lay out a few wi-fi receivers in the area she was targeting and listen as they picked up the sweet, information-rich sound of request probes and their unique, unchanging MAC addresses. She could connect new observations to those in her historical databases, and if she could somehow tie a MAC address to a real-world identity then it was game over for her target. Says Kate:

“I did one job in…actually I still shouldn’t tell you exactly where. The government had a large secret police force, one of the biggest on the continent. But it also had a even larger list of rabid anti-torture activists that it wanted to keep tabs on, and its secret police were stretched too thin to tail them all. I helped the government use scalable, cost-effective wi-fi technology to track the activists’ locations in real-time.

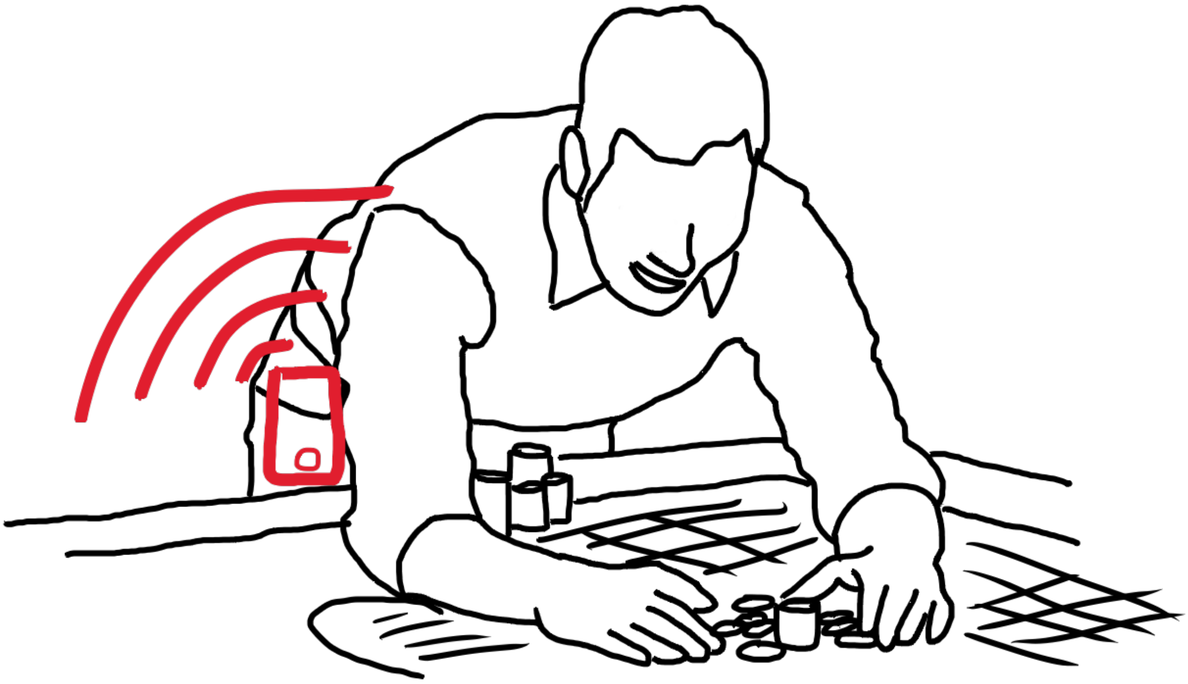
“All of the activists had smartphones. I drove my portable wi-fi listening gear around the city, pointing it at their seditious newspaper offices and meeting houses, hoovering up their devices’ request probes and MAC addresses. I showed the government how to build a city-wide network of wi-fi listeners capable of eavesdropping on all of their population’s probe requests. We used my database of dissident MAC addresses and *triangulation* to keep track of our targets’ physical locations in real-time.

“The government of [redacted] was very happy with my work. I never liked it when a client was that happy.”

## 2. Tracking devices by their MAC address, part 2

Academics were the first to consider - or at least the first to care about - the privacy implications of portable devices that squawk out a unique identifier every couple of seconds. Researchers began pestering the smartphone industry to improve, and in 2014 Apple announced that iOS8 would be the first mobile operating system to perform *MAC address randomization*. iOS8 devices would still include a MAC address with their request probes, but it would be a randomly generated one that changed every few minutes. This would prevent trackers from following an individual for more than a short period at a time, as devices would frequently rotate their MAC addresses and melt back into the crowd. However, this was far from the end of the story. Here’s Kate:

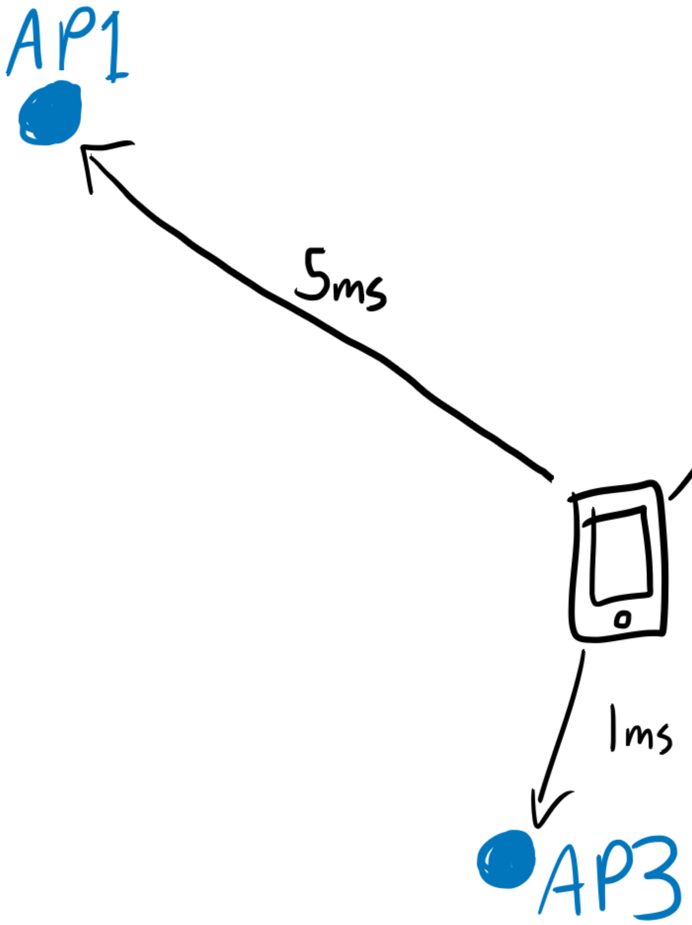
“My first job after the release of iOS8 was for an industry group of Las Vegas casinos. They wanted to use wi-fi-based tracking to build a shared list of Blackjack card-counters and their MAC addresses. Known nogoodniks often show up to casinos in deep disguise. The casinos wanted a system that used cheaters’ chattering smartphones to uncover their true identities.



“MAC address randomization certainly made my task harder. Fortunately, even devices performing MAC address randomization for their probe requests still switch back to using their real MAC address once they have connected to a network. I therefore did whatever it took to convince casino patrons to connect to a wi-fi network controlled by a casino. I insulated the casinos from cell reception, forcing gamblers to connect to the `FreeCasinoWifi` network if they wanted to talk to the outside world. I set up fake *Evil Twin* networks that looked like common nationwide networks like `Verizon Wi-Fi`. If a patron’s phone had ever previously connected to `Verizon Wi-Fi` then it would automatically try to re-connect to my Evil Twin, inadvertently revealing its real MAC address.

“Whenever casino security identified a new card-counter, they used my tracking system to skim the grifter’s MAC address and added it to the industry blocklist. My system matched a MAC address to a person using a technique called *triangulation*.

“To perform triangulation, I distributed multiple wireless access points (APs) throughout the casino. Whenever a customer’s smartphone emitted a wi-fi message, the message was picked up by each AP at fractionally different times, depending on the smartphone’s distance from the AP. I compared these times, calculated the differences, and used them to deduce the location of the smartphone and match it to a person.

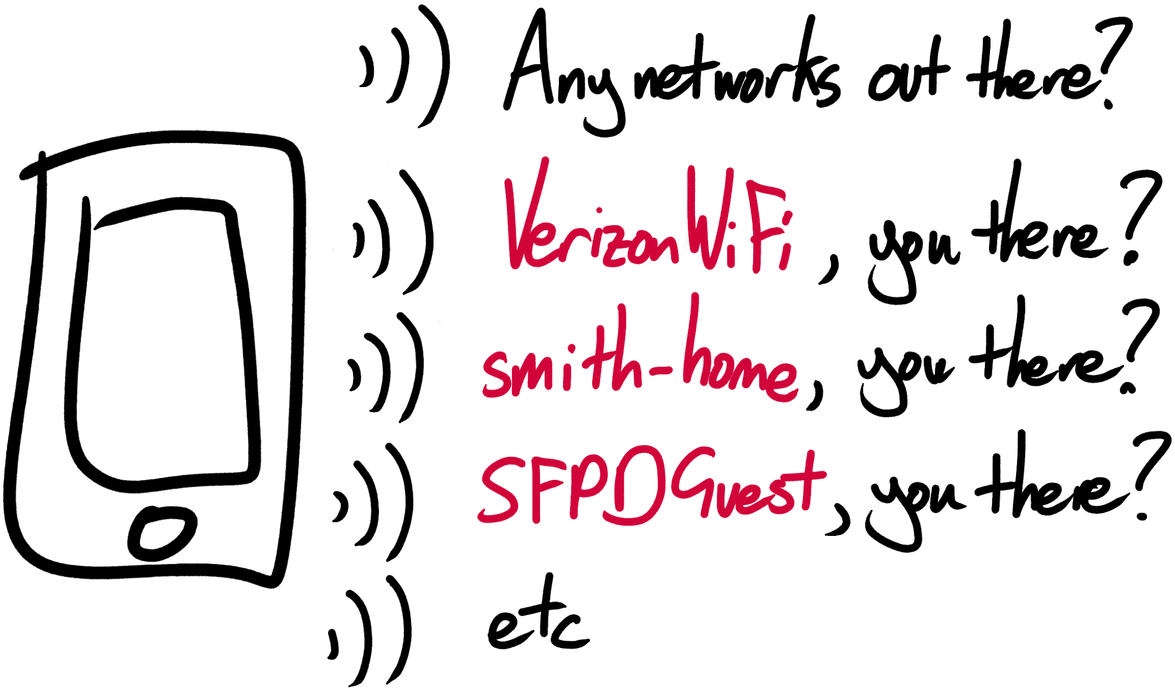


“If a swindler’s smartphone ever set foot in a casino again then an alert fired and the gentleman and the lady in the ginger wigs were swiftly escorted off the premises.

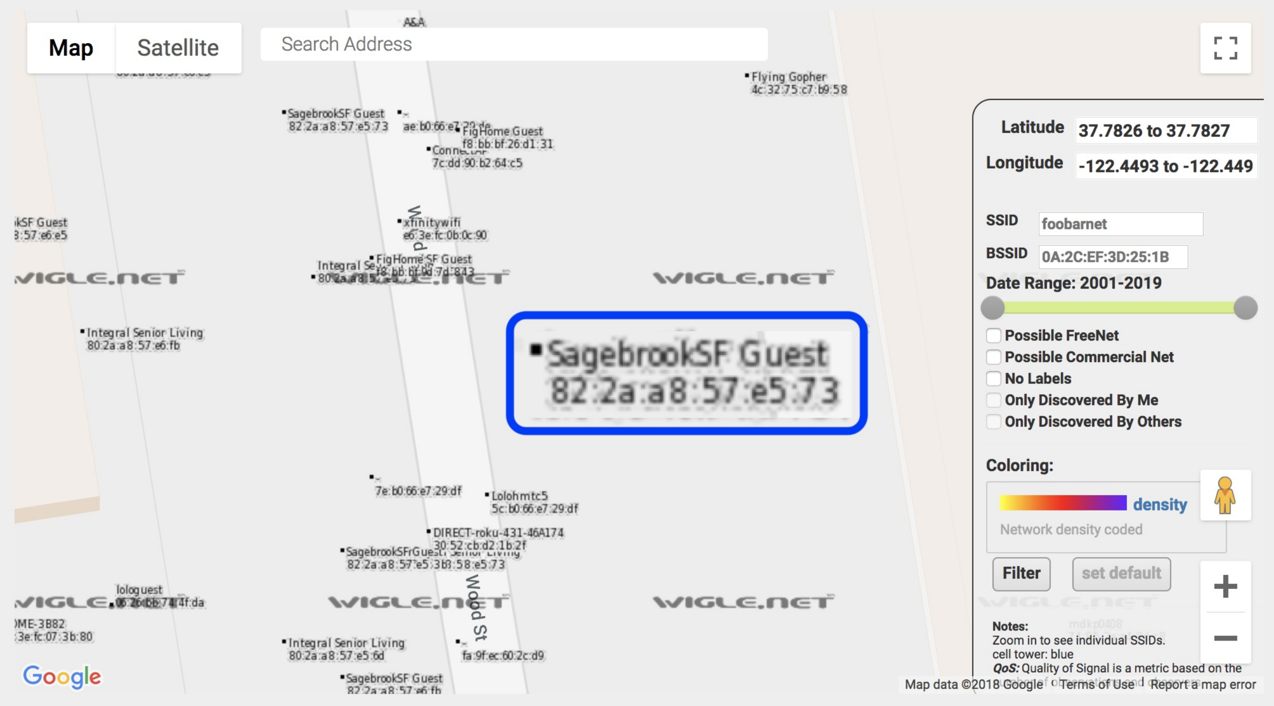
“I enjoyed this job.”

## 3. Tracking devices by the SSIDs in their probe requests

It’s not just a device’s MAC address that can give away its identity. Old smartphones used to emit much more than just the generic probe requests described above. They also fired out long series of targeted probe requests, each addressed directly to a specific wi-fi network. Instead of “hey, is anyone out there?” these requests asked “hey, `SFOAirportWifi`, are *you* out there?” Smartphones sent one of these requests for every wi-fi network they had ever previously connected to. This helped them to find and re-connect to familiar networks faster, and must have seemed like a good idea at the time.



Just like normal probes, targeted probes were transmitted every few seconds, in the clear and readable by anyone with a mind to look. But unlike normal probes, these ones contained the SSIDs of the networks that a user's device had previously connected to. Attackers could intercept them and use their contents to infer the locations to which the user had previously taken their smartphone (eg. `HOOTERSFREEWIFI`, `fbi-informant-hotspot`, `russian_embassy_guest`). Helpful, open source maps of wi-fi network locations even made locating a user's home from their wi-fi network's SSID straightforward.



Latitude 37.7826 to 37.7827
Longitude -122.4493 to -122.449

SSID foobarnet
BSSID 0A:2C:EF:3D:25:1B
Date Range: 2001-2019

□ Possible FreeNet
□ Possible Commercial Net
□ No Labels
□ Only Discovered By Me
□ Only Discovered By Others

Coloring: density
Network density coded

Filter | set default

Notes:
Zoom in to see individual SSIDs.
cell tower: blue
*QoS:* Quality of Signal is a metric based on the

What's more, given enough SSIDs, a tracker could create an *SSID fingerprint* that would often uniquely identify a single device. Very few devices have connected to all 3 of `steve-steveington-home` , `UC_BERKELEY_GUEST` and `SONIC-3991_3` . This makes this short list of SSIDs a very powerful personal identifier, even in the face of defensive MAC address randomization. Suppose that a tracker picks up 3 targeted probe requests addressed to the above 3 networks. Later that day it sees 3 more probe requests, addressed to the same 3 networks, but this time on the other side of the city. It's pesos to pizza that all of these probe requests came from the same device. This would allow the tracker to link the 2 wi-fi sessions into the same target profile, without needing to know anything about the target device's MAC address.

Around 2014, the privacy implications of targeted probe requests started to become widely publicized and understood. Most new devices therefore stopped sending them. Kate has this to add:

"In 2012 my good buddy, Dr. Prithi Prithibeta, and I were in training for an arduous charity walk. However, I developed reason to suspect that she was losing her commitment to the training program I had drawn up. I had a hunch that she was staying out far too late performing 'life-saving surgeries' at the hospital instead of getting the 8 hours of sleep that my program demanded. So, after one of our training sessions, I secretly installed a wi-fi monitoring device outside her apartment.

"Prithi lives with 8 roommates, each of whom owns several phones, laptops and Xboxes. These machines all pummeled my monitoring equipment with numerous fists of noisy chatter. Nonetheless, picking out the telltale fingerprint of Prithi's phone using the long list of previously-seen SSIDs that it broadcasted was trivial. As soon as I picked up simultaneous probe requests addressed to `new_york_general` , `prithibeta-family` and `XFINITY-18819` , I knew that Prithi and her smartphone had returned home. When I analyzed the data, I saw that she was indeed staying out performing surgeries well into the early hours of the morning, in direct contravention of the strict requirements of my training program. I was disappointed but not surprised. I confronted her with my evidence. Harsh words were exchanged and not taken back, and our charity walk team disbanded. I have no regrets."

## 4. Tracking devices by their IE fingerprints

When the privacy implications of targeted request probes became widely appreciated, most new mobile devices stopped sending them altogether. This plugged one gaping hole in the wall of the living room of consumer privacy. But many other holes remained, easily large enough for a sufficiently depraved misfit or location-tracking corporation to get a good peek through if they stood on a chair.

For example, *Information Elements* (IEs) are pieces of additional information that a device can send to a router when connecting to it. IEs can include helpful information about a device's country, its power constraints, and any vendor-specific properties that might be relevant. The purpose of IEs is to improve communication between a device and a network. However, IEs come in such a wide range of different types and values that very few devices share the exact same combination. This means that the list of IEs that a device advertises can be used as another form of near-unique fingerprint, in much the same way as the list of network SSIDs that a device has previously connected to.

Empirical studies have shown that IE fingerprints are somewhat less unique than those created from SSIDs or internet browser settings (see the EFF's Panopticlick project). It appears that some devices *do* advertise the exact same set of IEs, and so have the exact same IE fingerprint. However, even devices in the same location with the same IE fingerprint can often be de-duplicated using the *ordered sequence numbers* that they increment and append to each successive probe request. Kate doesn't have the time to go into this in detail, but she does have this anecdote to add:

"I once did a contract for a major big-box electronics retailer. Customers would often come agonizingly close to buying a big-ticket item, but eventually leave with their money still in their pocket. The higher-ups at [redacted] wanted to know who these people were and how to contact them so that they could target them with a follow-up firestorm of online ads.

"I built [redacted] a state-of-the-art system that tracked customer movements throughout their stores using triangulation and a multi-layered synthesis of tracking techniques. We started by trying to follow them using their device's MAC address. We fell back to SSID fingerprinting if their device was performing MAC address randomization, and we fell back further to IE fingerprinting if their device wasn't advertising its previously connected networks.

"Once we had locked onto a customer's device, we did whatever we could to get their email address. We had generous in-store promotions that required email addresses for participation, plus big incentives to swipe loyalty cards. When a target non-buyer swiped their card in order to receive their 50 [Redacted] Club Points as a thank you for visiting, we used our wi-fi tracking system to link their loyalty account to their smartphone. We looked at the path they had taken through the store, and checked where they had been vacillating.

"We then used the email address that they had given us when they signed up for their loyalty card to perform an online advertising technique known as *audience matching*. To run an audience matching campaign you upload your target's email address to an ad platform like Facebook or Amazon, along with the targeted ads you want that person to see. If Facebook or Amazon has a user with your target's email address (and they very often do) then they hit them with your generous offers for 2% off a TV or buy-4-get-1-free on bottom-of-the-range Pentaxs.

"Our online promotions did indeed tempt many of our targets back into our stores. Almost all of them brought their smartphones with them. We recognized their devices, alerted our sales reps, and gave them targeted tips on how to close the deals.

“Honestly I think the whole setup was fragile and contrived, and not worth the time or money that [redacted] sank into it. But I could still appreciate it as a true work of privacy-invading art.”

## 5. Tracking devices today

It is unquestionably much harder than it used to be to track a device using the wi-fi protocol. The basic threat models are well-understood throughout academia and industry, and mobile operating systems are increasingly touting their privacy features as a competitive differentiator.

MAC address randomization is easy and common. Targeted probe requests are mostly a thing of the past. Mobile operating systems are even starting to burn off their IE fingerprints. The Android Oreo OS removes unnecessary Information Elements from probe requests, which no longer contain enough different options to distinguish individual devices. This makes IE-based tracking like trying to identify a person from an extremely blurry retina scan.

However, all is not lost for the Steveington Show starring Steve Steveington. Many manufacturers continue to be far too slow to implement even straightforward, well-understood privacy features, and new attacks continue to be discovered. In 2017, researchers from the US Naval Academy demonstrated a novel approach for tracking devices if you already know their MAC address. Sending a *Request-To-Send control frame* addressed to the real MAC address of a device forced it to respond with a *Clear-To-Send frame*, implicitly acknowledging that the MAC address in your original frame was correct. The researchers speculate that this may be a vulnerability of devices’ underlying wi-fi chipsets, meaning that it can only be patched by an upgrade of the hardware itself. In the same paper the authors note that no Samsung or Motorola devices appear to perform even MAC address randomization. The tide of the battle for consumer privacy may be shifting, but it’s still far from - indeed will never be - over.

## Showrunning the Dark Web Truman Show

After all of this, it turns out that Steve Steveington recently dropped his new, mostly-privacy-protected iPhone X in the toilet. He was so disappointed with himself that he didn’t even bother to fish it out, instead choosing to literalize the metaphor by flushing his thousand-dollar device down the drain. He replaced his lost gadget with a battered smart-ish-phone made in 2009 that runs a strange mobile OS that you have never heard of. At first he only intended to use it as a stopgap while he established whether his insurance would cover either dropping one’s phone down the toilet or a lie about having been robbed. However, as he now never stops telling you, he’s started to really enjoy the freedom that this new old phone gives him from today’s always-on society blah blah blah and he’s going to keep it.

Of course, Steve’s prehistoric phone does not perform MAC address randomization. It sees no reason why it shouldn’t broadcast the SSID of every wi-fi network he has ever connected to, and it takes great pride in filling out every single Information Element in great detail. And Steve, being a normal human being, never manually turns off wi-fi. Kate Kateberry is almost disappointed when she finds out. She had been hoping to use this as an excuse to implement the US Naval Academy's Control Frame attack. However, she gains renewed enthusiasm for your project and its mission when her rent and salary come in and out of her almost-empty bank account in the wrong order, and she gets hit with a substantial overdraft fee.

The first season of the Steveington Show goes off without a hitch. You show Steve what you have done. He is sad. You are happy. You get some great; if morally troubling; but very lucrative ideas for sequels.