Phone Security

This is a very brief overview of phone security, mostly based on Chapter 22 of Anderson's *Security Engineering*

Phone networks

The earliest cell phone service, surprisingly, was developed in 1946. Usage was limited until the 1980s, when car phones started to become popular. The 1G phone technology of the 1980s was an analog system. There was no encryption, and it was easy to clone phones. Each country used a different standard, so phones from one country wouldn't work in another.

In the late 1980s, work was begun on a new standard, the Global System for Mobile Communications (GSM). This is synonymous with 2G technology. It's still in use today as a fallback in places without 4G or 5G, and those modern systems are in fact based on GSM technology. Unlike 1G technologies, GSM is digital, allowing for digital features like SMS (text) messages and internet access.

Cell phones talk to a *base station*. That base station in turn talks to the *core network*, which is managed by the cell phone carrier. Base stations are often associated with cell towers. The signal from each tower covers a particular area, roughly circular in shape. These areas are often drawn as a honeycomb of cells, which is where the *cell* in *cell phone* comes from.

Phones all contain a *SIM card*. That is a smartcard that contains a number called the *IMSI (international mobile subscriber identification)*. It is a unique number that is tied to the user's account. The SIM card also contains an authentication key that is used when connecting to a network. Besides the IMSI, there is also the *IMEI (international mobile equipment identification)*. This is a unique identifier for the phone itself, as opposed to the IMSI which is only for the SIM card. In some cases, criminals that swapped out their SIM card for a new one, while still using their own phone, were caught because of the IMEI.

Authentication and IMSI Catchers When authenticating to the network, a challenge-response protocol is used. When a phone connects to a network, it sends a message to the base station, which relays that through the core network to the user's *home location register (HLR)*, which is a database that contains the user's info. It sends back a random number to the phone. The phone's SIM card will encrypt the random number using the authentication key and send it back. The HLR also has a copy of the key, and it will verify that the encryption came out correct. This is how a user proves they have possession of the SIM card without having to send a copy of their key.

A serious weakness of the authentication process is that it is a one-way process. The user has to authenticate themselves to the network, but the network doesn't authenticate itself to the user. When a user connects to a base station, they don't know if it's the real one or now. This opened the door for a technology called an *IMSI catcher* or *StingRay*. These are essentially fake base stations. They are relatively small, often briefcase-sized devices, used by law enforcement and others to snoop on cell traffic. They trick users into connecting to them instead of the real base station by having a stronger signal.

Cryptography The creators of GSM tried to keep its cryptography secret. That usually doesn't work out well, and it didn't work out well here. Information about the ciphers leaked, and the rest of it was reverse-engineered. GSM used ciphers called A5/1 and A5/2. Weaknesses were found in A5/1, making it possible to break it in a few minutes time or less. A5/2 is a weakened form of A5/1 that was designed for export to other countries. It can be broken almost instantly and has been retired for use. In 2000, GSM vendors introduced A5/3, which seems to be somewhat secure. IMSI catchers often work by tricking the user's phone into using A5/1 or no encryption.

3G, 4G, and 5G 3G fixed many of the security weaknesses of GSM. However, IMSI catchers can still work by tricking the phone into using 2G. 3G technologies are being retired in many places in 2022. Places without 4G or 5G access in the US will still fall back to 2G. 4G is reasonably secure, though there are weaknesses. For instance, anyone that can compromise a base station can monitor all the traffic through it. 5G is still fairly new, so all of the security implications of it are not known yet. Instead of a few large base stations, 5G relies on many smaller ones. Much of the physical elements of the core network have been moved to the cloud. Likely, most of the security problems in 5G will come from weaknesses in cloud services.

Phones and Malware

The two main phones types are Google's Android and Apple's iPhone. Apple is generally considered the more secure of the two. Their devices are more locked down, both in a hardware and a software sense. iPhones are patched for up to five years, while Android phones are patched for three, and sometimes less. Android is an operating system which is run by phones from many different manufacturers, such as Google, Samsung, and Motorola. A security patch in Android needs to first be applied to Android itself, then it needs to make its way to the device manufacturers. Many of them take their time getting security patches out to users, and older devices may not get patches at all. A quick look at the Android security bulletins at `https://source.android.com/security` shows quite a few security issues being fixed each month, some of which are critical. An older, unpatched Android phone has a huge collection of ways for an attacker to do damage.

Even new phones are not safe. *Zero-day vulnerabilities*, which are things the developers don't even know about, can trade for millions of dollars on the black market if they are good enough, especially for Apple devices. Once they are used widely enough, the developers become aware of them and the vulnerabilities end up getting patched. But there are always new ones out there.

A lot of malware on Android involves taking a legitimate app and repackaging it to hold malware. This malware can root the phone to give the attacker remote access, or it can be used for phishing or to steal credentials by reading SMS text messages. Malware can also be pushed to a device through ads. Ads on a web browser can't do as much damage since they run JavaScript, which is run in a sandbox that limits its access to the system. But ads on mobile devices run native code, which allows more system access. In some cases, malware creators have gotten ads containing malware even on the biggest ad networks, like Google's AdSense.