Chapter 9
The American Way

The huge success that followed the introduction of GSM, first in Europe and then globally, caused some anxiety amongst the American technology companies. Although the United States had managed to build a comprehensive first-generation analog network, the major shift to digital suddenly seemed to happen elsewhere: the growth of GSM- based networks was phenomenal—GSM networks were available in 103 countries only five years after the first commercial implementation.

The history of the migration to digital networks turned out to be quite different in the United States, mainly due to the great success of the first-generation nationwide analog network roll-out. Although the capacity of the first-generation AMPS network was rapidly used up, it did cover virtually all of the continental United States. This was an impressive achievement, both in terms of the physical area covered and the relatively short time it had taken to reach this point.

The amount of hardware and money needed to reach a similar situation with the new digital networks would be painstakingly huge.

As the digital expansion had to be done in stages and in parallel with the existing analog network, the operators needed phones that supported both the analog AMPS and the new digital networks in the same device. Hence, specially tailored analog/digital hybrid handsets had to be created for the US market to support this transition period.

This parallel existence of both digital and analog circuitry added complexity and cost to the handsets and initially took away the possibility of benefiting from the global economies of scale that were fueled by the digital cellular explosion internationally. As a result, it seemed like the most active pioneer of mobile technology had suddenly become trapped in its own early success.

As one of the first analog systems, AMPS suffered from all the usual problems: it was prone to interference and dropped calls, easy to eavesdrop, and was an easy target for *subscriber cloning*, in which non-encrypted subscriber information was intercepted off the air and used to initialize another phone. All subsequent calls with the cloned phone would then be charged on the account of the owner of the original phone.

P. Launiainen, *A Brief History of Everything Wireless*,
https://doi.org/10.1007/978-3-319-78910-1_9

To fight the cloning problem, some highly complicated countermeasures had to be taken. For example, the network had to be made aware of certain characteristics of the user's phone model on the actual radio interface level so that it could be distinguished from another phone from a different manufacturer that was trying to use the same, cloned identity. This was a kludge at best, sometimes causing a false alarm and cutting a totally legal customer off the network.

All this extra hassle made it look like the country that had invented the cellular phone concept seemed to be losing the game to a new European digital standard, and to make it look even worse, GSM was not a result of a single, innovative company, but instead a concoction devised by a pan-European committee, with strong contribution from government bureaucrats.

Not only was this development happening outside of the United States, but the way it came into existence appeared to be totally against the expected norm of how innovation was supposed to work in a capitalist market.

But as the AMPS networks saturated, there was no way to stick to the first-generation networks. A gradual shift to digital had to be initiated and at first, in order to maintain maximum compatibility with the existing AMPS network, a system called *Digital AMPS (D-AMPS)* was created, coexisting on the same frequency band as AMPS. This method of interleaving digital users amongst the existing analog users provided a "soft upgrade", but it was soon obvious that this could not be a lasting solution.

What happened next was extremely unusual and risky:

In order to give a boost to domestic technologies and get local companies up to speed on digital mobile communications technology, the US market eventually turned into a test bed of four different digital cellular technologies.

The first homegrown, non-AMPS compatible standard was born in the West Coast, where the local operator, *Pacific Telesis (PacTel)*, invested heavily in a budding San Diego-based company, *Qualcomm*, which had come up with a new standard proposal called *Code Division Multiple Access (CDMA).*

CDMA used a military-style rapid *frequency-hopping* scheme, which was officially called *spread spectrum*, providing better immunity against interference, together with a theoretically more optimal utilization of the precious slice of the frequency spectrum that was available for cellular telephone use.

One particular area in which CDMA proved to be effective was dealing with long-distance *multipath propagation interference,* which refers to a situation where the transmitted signal is both received directly and reflected from various physical obstacles, like mountains, large bodies of water or high-rise buildings. This is the same effect that was the cause of the derisive "*Never the Same Color*" acronym for the American NTSC color television standard.

Another noteworthy improvement in CDMA was the more robust way to support the necessary handoff from base station to base station, effectively keeping two base stations connected to the handset until a solid handoff could be verified. This *soft handoff* feature helped in reducing the amount of dropped calls along the edges of the cells. Hence, in some crucial details of the implementation, CDMA had the potential to offer improvements over the GSM system.

The only issue was that *Qualcomm* lacked the money to provide a proof of concept, but this was soon remedied through the afore-mentioned initial investment from *PacTel*, a Californian operator struggling with the enormous growth of their customer base. Thanks to this collaboration, *Qualcomm* was able to move rapidly forward with its field testing, and the first US CDMA commercial network was opened in 1996.

CDMA was actively marketed across the world and also gained some international appeal outside of the United States—the very first commercial CDMA network was actually opened a couple of months before the *PacTel* offering by *Hutchinson Telecom* in Hong Kong.

In the years that followed, CDMA managed to gain support in a handful of locations across the world, most notably in South Korea, where the decision to select CDMA was heavily based on technopolitical reasoning in a country where the local *Chaebols*, huge industrial conglomerates, have a strong say in politics: going along with GSM would have meant that all the hardware for the Korean networks would come from existing vendors, or at least would force local manufacturers to pay heavy patent fees, so by selecting CDMA instead, the Koreans put their weight behind a new, developing standard, in which they expected to be actively participating.

In the United States, this unprecedented push of home-grown technologies initially slowed the transition to digital as compared to what was happening around the rest of the world. Having four different standards in the same geographical area resulted in a lot of overlapping, non-compatible infrastructure building, patchy networks, more costly handsets and a multitude of interoperability issues.

But as always, competition is fundamentally beneficial for technological development, and some interesting, novel features emerged from this potpourri: a good example of this application of American ingenuity was the *Integrated Digital Enhanced Network (IDEN)*, a competing network technology that was expanding the user experience in another direction by supporting so-called *push-to-talk* operation—you press a button, wait briefly for a beep indicating that an end-to-end voice channel has been opened, say what you want to say, and your voice is instantly played out at the destination handset.

No ringing, no answering. Direct connect.

This was exactly how traditional Walkie-Talkies work, except with IDEN it worked on a national scale, and proved to be a very useful feature for some particular user segments. This feature was deemed so successful that there was even an effort to update GSM to offer this kind of quick connectivity between handsets, but despite *Nokia* releasing a phone with this feature, it never caught on. This was partly due to inherent technical limitations that made the GSM version operate much more clumsily than the original IDEN version—the network logic in GSM simply was never designed to provide this kind of lightweight on-off activation for end-to-end connections, and there were too many fundamental issues that would need to be changed for optimal push-to-talk operation.

But perhaps more importantly, the operators didn't like it, as providing a "global Walkie-Talkie" feature could potentially have messed up their precious roaming

revenues. The operators already had the *Short Message Service (SMS)* for text-based, instant messaging, which offered astronomical *cost per bit* data revenues, especially when the users were roaming in another country. So why add a new, even easier way of instant communications that potentially could lead to a serious loss of revenue?

I remember one particular push-to-talk system sales meeting, in which the discussion was immediately focused on the SMS revenues the operator was gaining from the text messages generated by the citizens of the country that were working abroad. Introducing a push-to-talk feature in their network could potentially decimate this lucrative cash flow, so although the company otherwise had all the newest gizmos from *Nokia* in their laboratories, push-to-talk was dead on arrival.

As with so many cases where the manufacturers and operators had tightly shared interests, it took a total outsider to finally break this global SMS goldmine, more than a decade later: *WhatsApp, Inc*.

WhatsApp even offers a quasi-Walkie-Talkie audio messaging facility in a form of instant voice mail feature and finally made it easy to share images between users: something that the operator-driven *Multimedia Messaging Service (MMS)* extension never managed to do.

All in all, MMS is probably the biggest dud that came out of the GSM world: whereas SMS has mostly proven to be a very reliable channel of communications even in international cross-operator cases, MMS turned out to be fragmented and extremely unreliable, and hence never really took off.

In hindsight, MMS came to existence just at the borderline between traditional, telecom-oriented, custom-tailored services and generic mobile data-oriented services, and on the wrong side of that threshold.

WhatsApp did not only solve the expensive international text messaging and MMS use cases: with the recent introduction of zero cost, mobile data-based audio calls, *WhatsApp* is becoming a real *one-stop-shop* threat to the very core business of existing operators.

On the audio connectivity side, this kind of pure data-based solution has already existed for quite some time through services like *Skype*, but *WhatsApp* has really taken this concept to new heights with their billion active daily users.

More detailed discussion about this purely data-driven connectivity approach follows in Chapter 10: ***Internet in Your Pocket***.

At its heyday, IDEN offered great functionality for certain user segments, but eventually it could not keep up with the pace of what was happening on competing platforms, and it was finally decommissioned in 2013.

The huge economies of scale that were achieved in the GSM world were not overlooked in the United States either, and the first GSM-based network was activated in 1996. GSM was the only standard of the top four in use that could boast having inbuilt global roaming capability—a feature that was becoming more and more important for a highly developed, internationally connected nation like the United States.

There was only one snag. A big one.

The majority of GSM operators in the world built their networks on the 900 and 1800 MHz frequency bands, and the handsets were made to support them accordingly. But because these frequencies were already allocated to other uses in some countries, another pair of 850 and 1900 MHz had to be used instead.

The largest area with this kind of prior allocation was the North American market.

In the early days, multiple frequencies were costly to implement in the handsets, and therefore otherwise seemingly identical phones were sold with different active frequency bands for different markets. Hence, even though the GSM operators had made roaming agreements that made it possible for their users to keep their number while visiting another country, roaming with your own, physical phone was possible only if it supported the frequencies of the destination country.

With the next generation, dual-frequency phones that were made for their respective destination markets in order to address the huge user growth that had overwhelmed the original, single frequency band that was allocated for GSM, this cross-system functionality was not yet possible—900/1800 MHzEuropean GSM dual-frequency phone couldn't get into the US 850/1900 MHz dual-frequency networks, and vice versa.

For example, the legendary *Nokia 2110* was branded *Nokia 2190* for the version that supported the North American GSM frequencies, and it could not find a network to connect to if it was turned on pretty much anywhere else in the GSM world, despite being otherwise fully compatible with the local networks.

However, thanks to the GSM feature of separating the subscriber information from the handset, it was possible to remove the SIM card from your European/Asian/Australian phone and install it into a rented or purchased phone that used the North American frequencies. As was explained in Chapter 8: ***The Hockey Stick Years***, by doing this, your number and global connectivity automatically followed you into the new handset, and if you had your phone book entries stored in the SIM memory, everything looked just like it did at home.

The "can't roam with your own phone" limitation caused by the frequency differences across the world was eventually eased, thanks to ever cheaper radio circuitry that allowed adding new supported frequencies to the handsets.

Apart from just a handful of initial 900/1900 MHz hybrid solutions, the next evolutionary step came in the form of *tri-band* devices, where the third frequency was designed to be used in intercontinental roaming cases. For example, a tri-band North American phone would support 850/1800/1900 MHz bands, allowing it to use the 1800 MHz band when roaming in the "traditional" GSM regions that were prevalent elsewhere in the world.

Further down the line, full *quad-band* frequency support became available, providing the best possible network availability on a global scale. With such a device, the only limitation for using your phone anywhere in the world was whether your home operator had a roaming agreement with the destination country, and as the often exorbitant roaming charges added nicely to the operator income, all major operators had comprehensive, worldwide agreements in place.

Sometimes you still had to specifically activate the roaming feature on your account, but many operators had it switched on by default, so your phone would work the moment you stepped out of the plane in your destination country.

Naturally, as the caller could not possibly know that the person at the other end actually happened to be in a different country, the owner of the called account had to pay the extra roaming cost for the international part. Therefore, in practice, the true hard limitation for roaming use was whether the user wanted to bear the cost of this extra convenience.

Of course, some unaware or hapless roaming users managed to ramp up astronomical phone bills during their holiday stints, providing great fodder for human interest stories in the evening news, but the bottom line is that automatic, international roaming really was a groundbreaking improvement, not only for globe-trotting businessmen but also for the occasional tourist.

Roaming support provided by GSM was an especially big deal in Europe, where you could easily visit three different countries with a total of ten or even more different operators over the course of a single day.

If your home operator has comprehensive roaming agreements, your phone is able to automatically select from several operators in the destination country, and in case one of the roaming operators has a "hole" in their coverage, your phone switches seamlessly to another operator that works better in that particular area. Hence, a person roaming in another country can potentially have superior coverage compared with any local subscribers, as they are bound to a single, local operator.

To counteract the often very high cost of roaming, the aforementioned SIM portability naturally works both ways: if you want to benefit from the lower cost that results from using a local operator instead of utilizing the roaming feature of your subscription, you can purchase a local, *pre-paid* SIM card, swap it into your existing phone, and thus give your phone a new, fully local identity.

This way, by using a local operator, your local calls are as cheap as they can be, and most importantly, your mobile data costs are just a fraction of what your roaming data would cost you, so you can now get your mandatory holiday boasting snaps into *Instagram* and *Facebook* at a much lower cost.

With the special *dual-SIM* phones, you can even keep your home number alive and accessible, while directing all data traffic through the local operator's SIM in the second slot. Usually it does require a little poking around with the configuration menus, but the potential cost savings are more than worth the hassle.

There are some positive developments that make this kind of optimization less important in certain cases: thanks to the new European Union legislation that came in force in 2017, roaming costs within the EU were slashed. This brought the European Union closer to the convenience that customers in other large geographic areas, like the United States and Brazil, had already experienced for years.

Roaming costs aside, there was also another fundamental difference regarding the billing of mobile calls: most countries separated the cellular numbers from land line numbers by new area codes, so the caller always knew that dialing a mobile number would incur higher costs.

In contrast, in some countries, most notably in the United States and Canada, mobile numbers were spread out among the existing area code space. There was no clear way of knowing whether the dialed number was a fixed line or a mobile phone, and hence all extra costs for both incoming and outgoing calls had to be borne by the owner of the mobile account.

Because of this tiny but very significant difference in the charging policy, many early cellular users in the United States kept their phones shut off, only turning them on when they themselves had to make a call, or when they got a pager message of an existing voice mail.

As a result, although the customer base grew rapidly, the air-time utilization of cellular phones in the United States initially lagged behind the rest of the world. The United States also hung on to pagers much longer than the rest of the world, where the different charging approach allowed the users to keep their mobile phones on all the time, and the existing SMS messaging of the GSM system had made paging a thing of the past.

To speed up the growth of the customer base, the American cellular operators chose to heavily promote the use of subsidized accounts. In this model, you pay either nothing or just a small sum initially for your new phone, and the operator recoups the actual cost in the form of a higher monthly fee.

To ensure that you could not take advantage of this subsidy by mixing and matching various offers from different operators, your phone was locked to work only with the SIM card that was originally assigned to you by the operator. This greatly reduced the initial cost of ownership for the customers and helped the operators grow their customer base much more rapidly than in cases where the customer would have to fork out a largish sum for an unlocked phone.

SIM lock-based subsidizing of the handset cost was also a major driver of the expansion of mobile phone use in developing countries, as paying full price for the device up front would have been too high a cost for most potential users.

The fact that the lock was in the phone and in no way limited the portability of the SIM card was hard to grasp for even some advanced users: in reality, the operator only cares about the size of the air time bill that you accumulate through the usage of your account which is assigned to your SIM card, they could not care less about the device you use to generate it with.

A good example of this lack of faith happened when my then employer *Nokia* was working on a partnership deal with another company, and I brought some of the newest *Nokia* phones as a gift to their collaboration team:

Initially, I had hard time convincing them that they really could swap in the SIM card from their existing SIM locked devices, and everything would still work just fine, without any wrath from their operator. But in the end, as usual, the lure of a new, shiny thing proved to be strong enough, and soon a bunch of happy users were yapping away with their brand-new *Nokia 8890s*, also known as *Zippos*, thanks to their smooth, cigarette lighter-like design.

Nokia 8890 wasn't yet a tri-band GSM phone but a special, early generation hybrid "world" version, mainly aimed at high-end European and Asian users who needed to visit the United States occasionally. It had a peculiar 900/1900 MHz

combination, so it was pretty well covered in the traditional GSM world, whilst the 1900 MHz frequency provided decent coverage in the urban areas of the United States.

The only major visible difference between this Nokia 8890 "world" version and the Nokia *8850* 900/1800 MHz version was the fact that the former had a retractable antenna, whereas the latter had only an internal one. This was the era when internal antenna technology was not yet on par with traditional, external antennas, and this prominent difference pretty much summed up the expected second-generation network quality in the United States during the early years of the 21st century:

The vastness of the continental United States did not help with the huge task of providing new digital coverage to replace AMPS, and the first layer of GSM networks was built with only outdoor coverage in mind.

I remember visiting Dallas, Texas in 2000 and realizing that in order to make a call with my 1900 MHz-capable phone, I had to be standing next to a window on the "correct" side of the building, which the locals called "the phone side".

Even many of the major highways did not have any coverage outside the city limits.

For anyone traveling with a roaming GSM phone, the perceived coverage difference between the United States and Europe or even between the United States and rural Asia was huge at the time. But to some extent, this was not similar to the experience that users in the United States had, as instead of having just a digital network support, most locally sold US second-generation devices still had the analog AMPS fallback mode built in. So they worked in locations where the roaming GSM phones could not find digital coverage.

Around the turn of the century, the mobile market had become truly global, and roaming had become a feature that many customers took for granted. Even in a market as large and self-sufficient as the United States, it was hard to ignore the success and interoperability that GSM enjoyed around the globe.

Despite some of its technological advantages, CDMA could not provide the kind of international interoperability that was automatically available to all GSM subscribers. The biggest misstep of CDMA was the fact that subscriber information was integrated in the phone rather than as a separate, detachable module as done in GSM. *Qualcomm* had also been very flexible in terms of adapting to local requests regarding frequencies and operator-specific tailored features, and initially this approach brought in lucrative sales contracts in some countries but it also kept the overall CDMA market very fragmented.

On the network implementation side, the fierce multi-provider competition and vast, homogeneous deployment of GSM networks ensured good adherence to system specifications, quickly weeding out any major bugs in the required base station components that provided the network coverage.

As a comparison, CDMA was somewhat lacking behind GSM in terms of overall robustness, requiring more fine tuning and maintenance to keep the networks in optimal working condition.

Although CDMA-based networks were widely installed globally, their overall market share topped at around 20%.

In an ever-expanding market, CDMA had a hard time fighting the economies of scale provided by GSM, especially in countries where CDMA had to compete with existing GSM networks over the same customer base.

A textbook example of this happened with *Vivo*, the only Brazilian CDMA-based operator: the three other major operators in Brazil, *Oi*, *TIM* and *Claro*, were using GSM and benefited from the cheaper handset prices that were possible due to the fact that virtually identical handsets could be sold across the GSM-based world. This kind of mass-market of identical devices results in lower unit prices, and the fact that *Vivo*'s handsets did not support international roaming was also a major negative point for those users who wanted or needed to travel abroad—customers in this category tend to have lots of disposable income and are the most coveted clients from the operator's point of view.

Therefore, despite already providing cellular coverage for a country of comparable size to the United States, *Vivo* decided to rip out all CDMA equipment from its network and replace them with GSM. This costly transition, which also put the company solidly on the path of an incoming *third-generation (3G)* upgrade, started in 2006 and proved to be a success: today, *Vivo* is the largest Brazilian operator.

Another quirk that resulted in from this change is the fact that due to *Vivo's* original use of the 850/1900 MHz frequency band, as compared with *Vivo's* competitors' 900/1800 MHz frequencies, Brazil is now the world's largest market that has all four GSM frequencies in parallel use. This kind of wide cross-use of available frequencies is prone to have unexpected interference issues, which the Brazilian operators had to collaboratively weed out by blacklisting some portions of the spectrum in heavily congested areas.

Although CDMA did have some features that made it perform better in certain circumstances, the sheer power of economies of scale and the non-existing roaming capability limited its international reach.

At the same time, with the globally exploding number of users, the growth of which was seemingly exponential, GSM became a victim of its own success: capacity limits were creeping in, and there was no way to squeeze more out of the existing networks.

On the other hand, the existing CDMA networks were facing exactly the same issues, especially in the United States.

Due to the exploding user base, both of these 2G systems were now facing the same problems as 1G, only on a much bigger scale, as the worldwide number of cellular subscribers was counted in billions instead of millions.

And it was also no longer just a question of voice calls: *mobile data* was the new, rapidly growing "must have" feature, and GSM was never optimized for data traffic.

GSM, being the first digital solution, was suffering from its pioneering status: years of operational practice and further research had revealed that despite its unparalleled global success, GSM was not utilizing the limited radio spectrum in the most effective way, and it was more susceptible to certain types of interference and handoff dropouts than CDMA.

Hence, when the discussions regarding the *third-generation (3G)* systems eventually started, the creator of CDMA, *Qualcomm*, was ready. They now had a handful of market-proven technical improvements that could be extremely valuable in the fight over 3G, and were keen to capitalize on their experience as well as possible.

The next revolutionary step from *mobile voice* to *mobile data* was about to start.