computers are bad

https://computer.rip - me@computer.rip - fax: +1 (505) 926-5492

2022-11-27 over the horizon radar pt I

One if the most interesting things about studying history is noting the technologies that did *not* shape the present. We tend to think of new inventions as permanent fixtures, but of course the past is littered with innovations that became obsolete and fell out of production. Most of these at least get the time to become well-understood, but there are cases where its possible that even the short-term potential of new technologies was never reached because of the pace at which they were replaced.

And of course there are examples to be found in the Cold War.

Today were going to talk about Over-the-Horizon Radar, or OTH; a key innovation of the Cold War that is still found in places today but mostly lacks relevance in the modern world. OTHs short life is a bit of a disappointment: the most basic successes in OTH were hard-won, and the state of the art advanced rapidly until hitting a standstill around the 90s.

But lets start with the basics.

Radar systems can be described as either monostatic or bistatic, terms which will be important when I write more about air defense radar. Of interest to us now is monostatic radar, which is generally what you think of when someone just says radar. Monostatic radars emit RF radiation and then observe for a reflection, as opposed to bistatic radars which emit RF radiation from one site and then receive it at another site, observing for changes. Actually, well see that OTH radar sometimes had characteristics of both, but the most important thing is to understand the basic principle of monostatic radar, of emitting radiation and looking for what bounces back.

Radar can operate in a variety of parts of the RF spectrum, but for the most part is found in UHF and SHF - UHF (Ultra-High Frequency) and SHF (Super High Frequency) being the conventional terms for the spectrum from 300MHz-3GHz and 3GHz-30GHz. Why these powers of ten multiplied by three? Convention and history, as with most terminology. Short wavelengths are advantageous to radar, because RF radiation reflects better from objects that are a large portion or even better a multiple of the wavelength. A shorter wavelength thus means that you can detect smaller objects. There are other advantages of these high frequencies as well, such as allowing for smaller antennas (for much the same reason, the gain of an antenna is maximized at multiples of the wavelength, or at least at divisions by small powers of two).

UHF and SHF have a disadvantage for radar though, and that is range. As a rule of thumb, the higher the frequency (and the shorter the wavelength), the shorter the distance it will travel. There are various reasons for this, a big one is that shorter wavelengths more readily interact with materials in the path, losing energy as they do so. This has been a big topic of discussion in 5G telephony; since some 5G bands are in upper UHF and lower SHF where they will not pass through most building materials. The atmosphere actually poses the same problem, and as wavelengths get shorter the molecules in the atmosphere begin to absorb more energy. This problem gets very bad at around 60GHz and is one of the reasons that the RF spectrum must be considered finite (even more so than suggested by the fact that, well, eventually you get visible light).

Theres another reason, though, and its the more important one for our purposes. Its also the atmosphere, but in a very different way.

Most of the time that we talk about RF we are talking about line-of-sight operations. For high-band VHF and above [1], its a good rule of thumb that RF behaves like light. If you can see from one antenna to the other you will have a solid path, but if you cant things get questionable. This is of course not entirely true, VHF and UHF can penetrate most building materials well and especially for VHF reflections tend to help you out. But its the right general idea, and its very much true for radar. In most cases the useful range of a monostatic radar is limited to the radio horizon, which is a little further away than the visible horizon due to atmospheric refraction, but not that much further. This is one of the reasons we tend to put antennas on towers. Because of the low curvature of the earths surface, a higher vantage point can push the horizon quite a bit further away.

For air-defense radar applications, though, the type I tend to talk about, the situation is a little different. Most air-defense radar antennas are quite low to the ground, and are elevated on towers only to minimize ground clutter (reflections off of terrain and structures near the antenna) and terrain shadow (due to hills for example). A common airport surveillance radar might be elevated only a few feet, since airfields tend to be flat and pretty clear of obstructions to begin with. Theres a reason we dont bother to put them up on big towers: air-defense radars are pointed up. The aircraft they are trying to detect are quite high in the air, which gives a significant range advantage, sort of the opposite situation of putting the radar in the air to get better range on the ground. For the same reason, though, aircraft low to the ground are more likely to be outside of radar coverage. This is a tactical problem in wartime when pilots are trained to fly nap of the earth so that the reverse radar range, from their perspective, is very small. Its also a practical problem in air traffic control and airspace surveillance, as a Skyhawk at 2000 above ground level (a pretty typical altitude here in the mountain west where the ground is at 6k already) will pass through many blind spots in the Air Force-FAA Joint Surveillance System.

This is all a somewhat longwinded explanation of a difficult problem in the early Cold War. Before the era of ICBMs, Soviet nuclear weapons would arrive by airplane. Airplanes are, fortunately, fairly slow... especially bombers large enough for bulky nuclear munitions. The problem is that we would not be able to detect inbound aircraft until they were quite close to our coasts, allowing a much shorter warning (and interception) time than you would expect. There are a few ways to solve this problem, and we put great effort into pursuing the most obvious: placing the radar sets closer to the USSR. NORAD (North American Air Defense Command) is a joint US-Canadian venture largely because Canada is, conveniently for this purpose, in between the USSR and the US by the shortest route. A series of radar lines were constructed across Alaska, Canada, and into Greenland, culminating with the DEW (Distant Early Warning) Line in arctic norther Canada.

This approach was never quite complete, and there was always a possibility that Soviet bombers would take the long route, flying south over the Pacific or Atlantic to stay clear of the range of North American radar until they neared the coasts of the US. This is a particularly troubling possibility since even today the population of the US is quite concentrated on the coasts, and early in the Cold War it was even more the case that the East Coast *was* the United States for most purposes. Some creative solutions were imagined to this problem, including most notably the Texas Towers, radar stations built on concrete platforms far into the ocean. The Texas Towers never really worked well; the program was canceled before all five were built and then one of them collapsed, killing all 28 crew. There was an even bigger problem with this model, though: the threat landscape had changed.

During the 1960s, bombers became far less of a concern as both the US and the USSR fielded intercontinental ballistic missiles (ICBMs). ICBMs are basically rockets that launch into space, orbit around to the other side of the planet, and then plunge back towards it at terminal velocity. ICBMs are fast: a famous mural painted on a blast door by crew of a Minuteman ICBM silo, now Minuteman Missile National Historic Park, parodies the Dominos Pizza logo with the slogan Delivered worldwide in 30 minutes or less, or your next one is free. This timeline is only a little optimistic, ICBM travel time between Russia and the US really is about a half hour.

Moreover, ICBMs are hard to detect. At launch time they are very large, but like rockets (they are, after all, rockets, and several space launch systems still in use today are directly derived from ICBMs) they shed stages as they reach the apex of their trip. By the time an ICBM begins its descent to target it is only a re-entry vehicle or RV, and some RVs are only about the size of a person. To achieve both a high probability of detection and a warning time of better than a few minutes, ICBMs needed to be detected *during their ascent.* This is tricky: Soviet ICBMs had a tendency of launching from the USSR, which was a long ways away.

From the middle of the US to the middle of Russia is around 9000km, great circle distance. Thats orders of magnitude larger than the range of the best extant radar technology. And there are few ways to cheat on range: the USSR was physically vast, with the nearest allied territory still being far from ICBM fields. In order to detect the launch of ICBMs, we would need a radar that could not only see past the horizon, but see *far* past the horizon.

Lets go back, now, to what I was saying about radio bands and the atmosphere. Below VHF is HF, High Frequency, which by irony of history is now rather low frequency relative to most applications. HF has an intriguing property: some layers of the atmosphere, some of the time, will actually reflect HF radiation. In fact, complex propagation patterns can form based on multiple reflections and refraction phenomenon that allow lucky HF signals to make it clear around the planet. Ionospheric propagation of HF has been well known for just about as long as the art of radio has, and was (and still is) regularly used by ships at sea to reach each other and coast stations. HF is cantankerous, though. This is not exactly a technical term but I think it gets the idea across. Which HF frequencies will propagate in which ways depends on multiple weather and astronomical factors. More than the complexity of early radio equipment (although this was a factor), the tricky nature of HF operation is the reason that ships carried a radio officer. Establishing long-distance connections by HF required experimentation, skill, and no small amount of luck.

Luck is hard to automate, and in general there werent really any automated HF communications systems until the computer age. The long range of HF made it very appealing for radar, but the complexity of HF made it very challenging. An HF radar could, conceptually, transmit pulses via ionospheric propagation well past the horizon and then receive the reflections by the same path. The problem is how to actually interpret the reflections.

First, you must consider the view angle. HF radar energy reflects off of the high ionosphere back towards the earth, and so arrives at its target from above, at a glancing angle. This means of course that reflections are very weak, but more problematically it means that the biggest reflection is from the ground... and the targets, not far above the ground, are difficult to discriminate from the earth behind them. Radar usually solves this problem based on time-of-flight. Airplanes or recently launched ICBMs, thousands of feet or more in the air, will be a little bit closer to the ionosphere and thus to the radar site than the ground, and so the reflections will arrive a bit earlier. Heres the complication: in ionospheric propagation, multipath is almost guaranteed. RF energy leaves the radar site at a range of angles (constrained by the directional gain of the antenna), hits a large swath of the ionosphere, and reflects off of that swath at variable angles. The whole thing is sort of a smearing effect... every point on earth is reached by a number of different paths through the atmosphere at once, all with somewhat different lengths. The result is that time-of-flight discrimination is difficult or even impossible.

There are other complexities. Achieving long ranges by ionospheric propagation requires emitting RF energy at a very shallow angle with respect to the horizon, a few degrees. To be efficient (the high path loss and faint reflections mean that OTH radar requires enormous power levels), the antenna must exhibit a very high gain and be very directional. Directional antennas are typically built by placing radiating and reflecting elements some distance to either side of the primary axis, but for an antenna pointed just a few degrees above the horizon, one side of the primary axis is very quickly in the ground. HF OTH radar antennas thus must be formidably large, typically using a ground-plane design with some combination of a tall, large radiating system and a long groundplane extending in the target direction. When I say large here I mean on the scale of kilometers. Just the design and construction of the antennas was a major challenge in the development of OTH radar.

Lets switch to more of a chronological perspective, and examine the development of OTH. First, I must make the obligatory disclaimer on any cold war technology history: the Soviet Union built and operated multiple OTH radars, and likely arrived at a working design earlier than the US. Unfortunately, few resources on this history escaped Soviet secrecy, and even fewer have been translated to English. I know very little about the history of OTH radar in the USSR, although I will, of course, discuss the most famous example.

In the US, OTH radar was pioneered at the Naval Research Laboratory. Two early prototypes were built in the northeastern United States: MUSIC, and MADRE. Historical details on MUSIC are somewhat scarce, but it seems to have been of a very similar design to MADRE but not intended for permanent operation. MADRE was built in 1961, located at an existing NRL research site on Chesapeake Bay near Washington. Facing east towards the Atlantic, it transmitted pulses on variable frequencies at up to 100kW of power. MADREs large antenna is still conspicuous today, about 300 feet wide and perhaps 100 feet tallbut this would be quite small compared to later systems.

What is most interesting about MADRE is not so much the radio gear as the signal processing required to overcome the challenges Ive discussed. MADRE, like most military programs, is a tortured acronym. It stands for Magnetic-Drum Radar Equipment, and that name reveals the most interesting aspect. MADRE, like OTH radars to come, relied on computer processing to extract target returns.

In the early 60s, radar systems were almost entirely analog, particularly in the discrimination process. Common radar systems cleared clutter from the display (to show only moving targets) using methods like mercury acoustic delay lines, a basic form of electronic storage that sent a signal as a mechanical pulse through a tube of mercury. By controlling the length of the tube, the signal could be delayed for whatever period was usefulsay one rotational period of the radar antenna. For OTH radar, though, data needed to be stored on multiple dimensions and then processed in a time-compressed form.

Lets explain that a bit Further. When I mentioned that it was difficult to separate target returns from the reflection of the earth, if you have much interest in radar you may have immediately thought of Doppler methods. Indeed, ionospheric OTH radars are necessarily Doppler radars, measuring not just the reflected signal but the frequency shift it has undergone. Due to multipath effects, though, the simple use of Doppler shifts is insufficient. Atmospheric effects produce returns at a variety of shifts. To discriminate targets, its necessary to compare target positions between pulses... and thus to store a history of recent pulses with the ability to consider more than one pulse at a time. Perhaps this could be implemented using a large number of delay lines, but this was impractical, and fortunately in 1961 the magnetic drum computer was coming into use.

The magnetic drum computer is a slightly odd part of computer history, a computer fundamentally architected around its storage medium (often not only logically, but also physically). The core of the computer is a drum, often a fairly large one, spinning at a high speed. A row of magnetic heads read and write data from its magnetically coercible surface. Like delay tubes, drum computers have a fundamental time basis in their design: the revolution speed of the drum, which dictates when the same drum position will arrive back at the heads. But, they are two-dimensional, with many compact multi-track heads used to simultaneously read and write many bits at each drum position.

Signals received by MADRE were recorded in terms of Doppler shifts onto a drum spinning at 180 revolutions per second. The radar similarly transmitted 180 pulses per second (PRF), so that each revolution of the drum matched a radar pulse. With each rotation of the drum, the computer switched to writing the new samples to a new track, allowing the drum to store a history of the recent pulses20 seconds worth.

For each pulse, the computer wrote 23 analog samples. Each of these samples was range gated, meaning time limited to a specific time range and thus distance range. Specifically, in MADRE, each sample corresponded to a 455 nmi distance from the radar. The 23 samples thus covered a total of 10,465 nmi in theory, about half of the way around the earth. The area around 0Hz Doppler shift was removed from the returned signal via analog filtering, since it always contained the strong earth reflection and it was important to preserve as much dynamic range as possible for the Doppler shifted component of the return.

As the drum rotated, the computer examined the history of pulses in each range gate to find consistent returns with a similar Doppler shift. To do this, though, it was first necessary to discriminate reflections of the original transmitted pulse from various random noise received by the radar. The signal processing algorithm used for this purpose is referred to as matched filtering or matched Doppler filtering and I dont really understand it very well, but I do understand a rather intriguing aspect of the MADRE design: the computer was not actually capable of performing the matched filtering at a high enough rate, and so an independent analog device was built to perform the filtering step. As an early step in processing returns, the computer actually played them back to the analog filtering processor at a greatly accelerated speed. This allowed the computer to complete the comparative analysis of multiple pulses in the time that one pulse was recorded.

MADRE worked: in its first version, it was able to track aircraft flying over the Atlantic ocean. Later, the computer system was replaced with one that used magnetic core memory. Core memory was random access and so could be read faster than the drum, but moreover GE was able to design core memory for the computer which stored analog samples with a greater dynamic range than the original drum. These enhancements allowed MADRE to successfully track much slower targets, including ships at sea.

The MUSIC and MADRE programs produced a working OTH radar capable of surveiling the North Atlantic, and their operation lead to several useful discoveries. Perhaps the most interesting is that the radar could readily detect the ionospheric distortions caused by nuclear detonations, and MADRE regularly detected atmospheric tests at the NNSS despite pointing the wrong direction. More importantly, it was discovered that ICBM launches caused similar but much smaller distortions of the ionosphere which could also be detected by HF radar. This improved the probability of HF radar detecting an ICBM launch further.

AND THATS PART ONE. Im going to call this a multi-part piece instead of just saying Ill return to it later so that, well, Ill return to it later. Because heres the thing: on the tails of MADREs success, the US launched a program to build a second OTH radar of similar design but bigger. This one would be aimed directly at the Soviet Union.

It didnt work.

But it didnt work in a *weird* way, that leaves some interesting questions to this day.

[1] VHF is 30-300MHz, which is actually a pretty huge range in terms of characteristics and propagation. For this reason, land-mobile radio technicians especially have a tendency to subdivide VHF into low and high band, and sometimes mid-band, according to mostly informal rules.