How Does GPS work?

Kehan Chen

2017/7/21

Math 190S

Duke summer college

1.Introduction

With nearly a thousand satellites hanging above and rotating, humans, residents of Earth, now have capability of obtaining various kinds of information like weather and scale, and are easy to communicate with everyone around the world.

Among all functions satellites bring us, the most important and fundamental one is satellite positioning, which is known for its application----GPS (Global Positioning System). It is a global navigation satellite system that provides geolocation and time information to a GPS receiver anywhere on or near the Earth where there is an unobstructed line of sight to four or more GPS satellites. [1]

With satellite positioning technology, human can utilize satellites to position objects anywhere on Earth. To achieving that, it has to be guaranteed that at least four satellites can be observed anywhere on Earth at any moment; as a result, precise positioning can be achieved and applied to practice in the form of navigation, positioning, time service, etc.

Since satellite positioning technology plays a significant role in humans' life, we are going to talk about how it works and how it is applied to reality to impact our lives.

2.Segments of Satellite Positioning

To know about the segments of satellite positioning, let’s take GPS as an example. The Global Positioning System (GPS) is a U.S.-owned utility that provides users with positioning, navigation, and timing services. This system consists of three segments: the space segment, the control segment, and the user segment. The U.S. Air Force develops, maintains, and operates the space and control segments. [2]

2.1 Space segment

The GPS space segment consists of a constellation of satellites transmitting radio signals to users.

GPS satellites fly in medium Earth orbit (MEO) at an altitude of approximately 20,200 km (12,550 miles). Each satellite circles the Earth twice a day.

The satellites in the GPS constellation are arranged into six equally-spaced orbital planes surrounding the Earth. Each plane contains four "slots" occupied by baseline satellites. This 24-slot arrangement ensures users can view at least four satellites from virtually any point on the planet. (Figure 1)

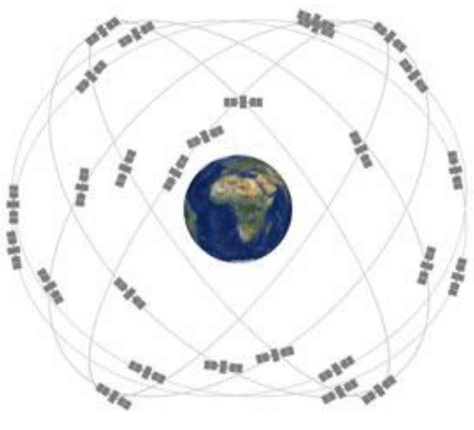

Figure 1: Expandable 24-Slot satellite constellation of GPS

The Air Force normally flies more than 24 GPS satellites to maintain coverage whenever the baseline satellites are serviced or decommissioned. The extra satellites may increase GPS performance but are not considered part of the core constellation.

2.2 Control segment

GPS control segment consists of a global network of ground facilities that track the GPS satellites, monitor their transmissions, perform analyses, and send commands and data to the constellation.

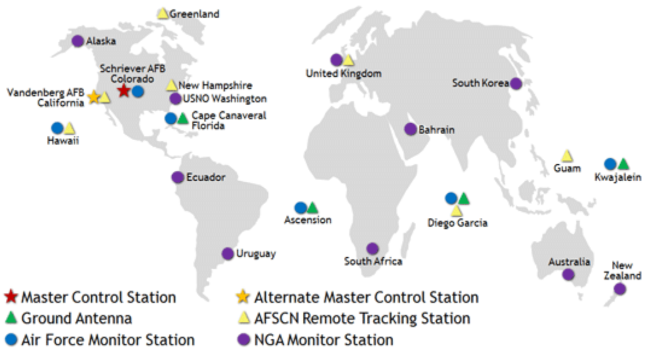

Figure 2

The current operational control segment includes a master control station, an alternate master control station, 11 command and control antennas, and 16 monitoring sites. (Figure 2)

2.3 User segment

GPS now appears everywhere in our daily lives. Mobile phones, for example, is the most common user segment for GPS. Navigation for not only cars, but also ships, and 3-D aircrafts have also been broadly achieved by receivers of GPS signals in those transportation tools.

Other functions like precise positioning in Surveying or time dissemination in

Astronomy can also be realized by different kinds of user segments. In particular, user segment for military use can help army to locate and guide themselves as well as search for and locate their enemy, which is a vital tool for military.

3. Basic Principles for GPS

After knowing about the GPS segment, it's time to know how it works.

3.1 Measure the distance between satellites and receivers

GPS positioning is based on measurement of the distance between satellites and receivers. Thus, to know where something is, we have to figure out how far it is between that point and the satellite first. Only after measuring this distance precisely can we do the positioning accurately.

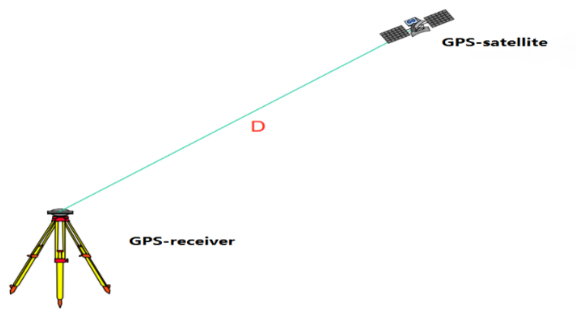

Figure 3

I define the distance between GPS-satellite and GPS-receiver as D, the speed of light as c and the traveling time of signal as t_t. Apparently, we can get:

$$\mathrm{D} = \mathrm{c} \cdot t_t$$

Then, the question is how to measure the traveling time of the signal, namely

t_t. And the solution is below:

Firstly, let both the GPS-satellite and GPS-receiver produce a same series of code at t_1.

Then, after a period of time, the code from the satellite is transmitted to the receiver on the ground at t_2.

Let $\Delta t = t_2 - t_1$

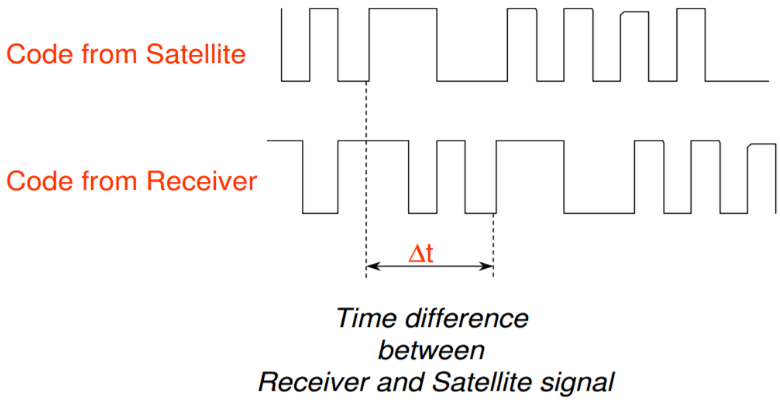

Figure 4: the time difference between two codes

Obviously,

$$D = c \cdot (t_2 - t_1) = c \cdot \Delta t$$

Now, the question is how to get Δt. It has to be known that the code above is actually a Pseudo-Random Code, a complicated mathematic code or a sequence of “on” and “off” pulse. Thus, it will produce 0 or 1 at same possibility and it has a relatively good autocorrelation. Autocorrelation, also known as serial correlation, is the correlation of a signal with a delayed copy of itself as a function of delay. Informally, it is the similarity between observations as a function of the time lag between them. The analysis of autocorrelation is a mathematical tool for finding repeating patterns, such as the presence of a periodic signal obscured by noise. [3]

In that case, code from the satellite can be aligned with code from receiver through translocation. If we can figure out how much the code from satellite moves, we can find out the what Δt is.

I implement exclusive OR to the delayed code from satellite and the synchronized code from receiver to pose a new code. Define the amount of "0" in that new code as N_0 and the amount of "1" as N_1.

After that, I define an autocorrelation function:

$$R = \frac{N_0 - N_1}{N_0 + N_1}$$

With this function, we can know whether the code from satellite is aligned with the code from receiver through translocation.

When they are aligned with each other, R should be equal to 1. For example:

$$
\begin{array}{r}
1\,1\,1\,1\,0\,0\,0\,1\,0\,0\,1\,1\,0\,1\,0 \\
\oplus \quad 1\,1\,1\,1\,0\,0\,0\,1\,0\,0\,1\,1\,0\,1\,0 \\
\hline
0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0
\end{array}
$$

$$R = \frac{N_0 - N_1}{N_0 + N_1} = \frac{15 - 0}{15 + 0} = 1$$

However, when the two codes are not aligned with each other, R will roughly equal to 0 since the amount of "0" and "1" are nearly the same. Here is an example when the two codes interlace for a digit:

$$
\begin{array}{r}
1\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 1\ 1\ 0\ 1\ 0 \\
\oplus \quad 0\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 1\ 1\ 0\ 1 \\
\hline
1\ 0\ 0\ 0\ 1\ 0\ 0\ 1\ 1\ 0\ 1\ 0\ 1\ 1\ 1
\end{array}
$$

$$R = \frac{N_0 - N_1}{N_0 + N_1} = \frac{7-8}{7+8} = -\frac{1}{15} \approx 0$$

After receiver receives signals (code) from the satellite, it can finally align the code from the satellite with the code produced by receiver.

Figure 5

By doing this, receiver can get how many digits the code from satellite translocate to be aligned with the code from receiver. I define the quantity of that digits as N_d and N_d is a known quantity. Besides that, if we can know how long it takes to transmit a digit from satellite to receiver, the problem of measuring will be solved. So, I define this time, namely code width, as t_u and it turns out that t_u is also a known quantity since:

$$t_u = \frac{1}{f}$$

In this equation, f is the frequency of the transmitting signal, thus a known quantity. So, with the equation:

$$\Delta t = t_u \cdot N_d$$

As a result,

$$D = c \cdot \Delta t = c \cdot t_u \cdot N_d$$

Now the measurement is done and let's see how positioning works with the distance, D, measured above. [4]

3.2 Geometric principles

Now we have successfully measured the distance from a spot to a satellite and with distances between the spot and 4 different satellites, GPS can implement positioning.

When distance is measured between only one satellite and a position,

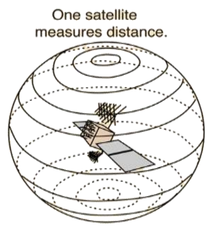

Figure 6

we can know that the position is on the sphere shown above. The radius of this sphere is the distance measured.

When two distances are measured,

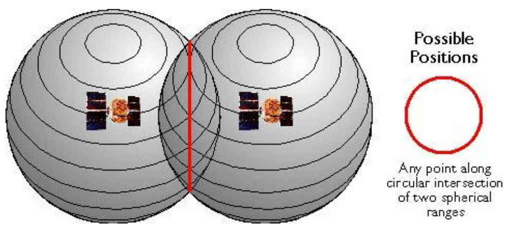

Figure 7

We can see clearly that the spot needing to be positioned is on the intersect of two spheres, which is a circle shown above.

When three distances are measured ,

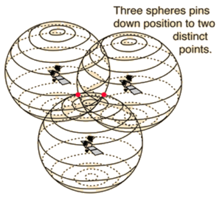

Figure 8

it is obvious that positioning point lies in either the two spots, which are the intersect of the three spheres shown above. The radius of the three spheres are distances from their respective central satellite to the positioning spot.

In most cases, only one position is actually reasonable and the other lies either inside Earth or outside in space. However, to improve the accuracy of positioning, GPS uses another forth satellite to finally decide the actual location of the spot. So, the whole process will be:

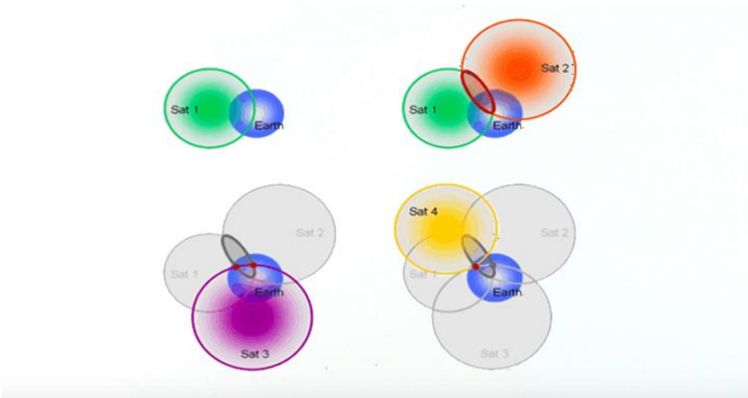

Figure 9

Because of the geometric principles, satellite positioning can be realized.

4. Math Model for Satellite Positioning

To accurately position a spot on Earth, we can establish a 3- dimensional coordinate system which takes the center of Earth as its origin.

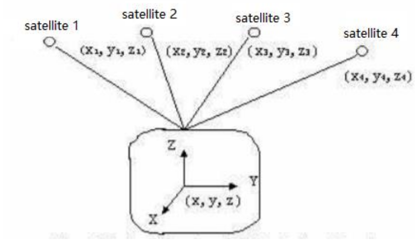

Figure 10

For the spot, its coordinate is $(\mathrm{x}, \mathrm{y}, \mathrm{z})$. And since there are clock errors between satellites and receivers (which will be discussed later in the paper), define the clock error of receiver (the spot) as v_{t_0}

For the four satellites, their coordinates are respectively$(x_i, y_i, z_i)(i = 1,2,3,4)$. Then, define the clock errors of four satellites as $v_{t_i}(i = 1,2,3,4)$, which are provided by the Satellite ephemeris. Finally, define the distances between four satellites and receiver as $d_i(i = 1,2,3,4)$ and the time to cover for signal to cover that distance as $\Delta t_i(i = 1,2,3,4)$.

So, the equation system will be

$$\begin{cases} \sqrt{(x_1 - x)^2 + (y_1 - y)^2 + (z_1 - z)^2 + c \cdot (v_{t_1} - v_{t_0})} = d_1 \\ \sqrt{(x_2 - x)^2 + (y_2 - y)^2 + (z_2 - z)^2 + c \cdot (v_{t_2} - v_{t_0})} = d_2 \\ \sqrt{(x_3 - x)^2 + (y_3 - y)^2 + (z_3 - z)^2 + c \cdot (v_{t_3} - v_{t_0})} = d_3 \\ \sqrt{(x_4 - x)^2 + (y_4 - y)^2 + (z_4 - z)^2 + c \cdot (v_{t_4} - v_{t_0})} = d_4 \end{cases}$$

In the equation system, x, y, z, v_{t_0} are four unknown quantities needing to be figured out. And since there are four unknowns as well as four equations, the system is solvable. [5]

With this math model, we can realize satellite positioning.

5. Error Sources in Precise Positioning

Due to the complicated structure of Earth and the spacetime, there are some factors that impact on the accuracy of satellite positioning. In general, there are four types of factors: Receiver's environment (multi-path), Ionospheric and tropospheric delays, Satellite constellation, and Clock errors.

5.1 Receiver's environment

Surroundings of receivers can affect positioning by blocking or interfering signal transmitting. And there are mainly two ways (Figure 10) to affect the accuracy.

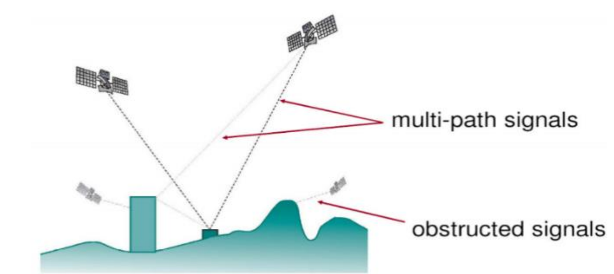

Figure 10

In that case, distances between receiver and satellites may not be measured or at least may not be measured accurately. Thus, the positioning will be error to some extent.

5.2 Ionospheric and tropospheric delays

Troposphere is the atmospheric layer placed between earth's surface and an altitude of about 60 kilometers. Ionospheric stretches from about 50 km (30 miles) from the ground to the earth's upper atmosphere of about 1000 km altitude airspace.

In ionosphere, there are quite a lot of free electrons and ions, and can change propagation of radio wave and make refraction, reflection and scattering. [6] The effect depends on the density of electrons and ions as well as temperature.

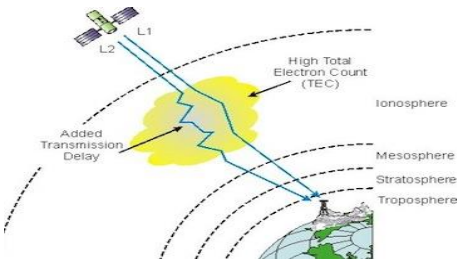

Figure 11

The effect of the troposphere on the satellite signals appears as an extra delay in the measurement of the signal traveling from the satellite to receiver. This delay depends on the temperature, pressure, humidity as well as the transmitter and receiver antennas location. [7] For example, Water vapor slows down radio waves. [8]

To sum up, as radio signals pass through the ionosphere and troposphere, they slow down based on the density of atmosphere. [9]

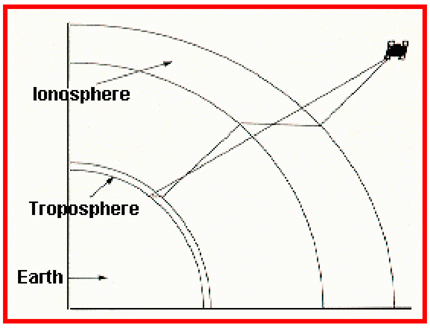

Figure 12

5.3 Satellite constellation

The distribution of satellites also will influence the precision of positioning. The professional name of this influence is “Geometric Dilution of Precision (GDOP)”, which refers to where the satellites are in relation to each other.

It has to be admitted that there is error in the measurement of distance. So, when the distance between a spot and a satellite is measured, the possible location of the spot lies not exactly on the sphere but a spherical shell that has width and takes the satellite as its center. As a result, the final location of the spot will also have a scope instead of a single spot.

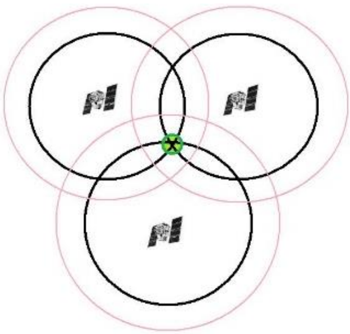

Figure 13

What the scope is like actually affect the accuracy of positioning and it is decided by the satellite constellation. A good star constellation will produce a good scope for location, which has less effect on the accuracy. (Figure 14)

While a bad one will attach much negative effect to the accuracy of positioning. (Figure 15)

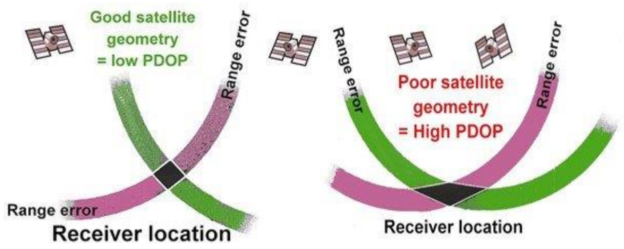

Figure 14 Figure 15

GPS estimates of elevation (Figure 15) typically have errors that are 2 to 5 times worse than the horizontal position (Figure 14) error.

To address the problem to some degree, user can set the Elevation Angle Mask on their receivers or user segments, which determines the minimum elevation angle below which the receiver will no longer use a satellite in its computations. [10] When satellites are close to horizon, there is more atmosphere, troposphere, and ionosphere for the signals to pass through; thus, more error will take place. Setting an Elevation Angle Mask can help people avoid using the satellite that will cause much error.

5.4 Clock errors

According to the theory of relativity, timing will be influenced by the velocity difference between two objects. Thus, people have to synchronize the time on the satellites with the time of receivers. To guarantee the time accuracy on the satellites, designers equip each satellite with an atomic clock, which is the most accurate timing equipment humans have so far. An atomic clock is a clock device that uses an electron transition frequency in the microwave, optical, or ultraviolet region of the electromagnetic spectrum of atoms as a frequency standard for its timekeeping element. [11] However, if also installing atomic clocks to receivers was the only way to minimize time error, then GPS would never be able to spread to people's daily life because of the unaffordable expenses. So, to address the problem, experts give the calculation.

The clock error is partly caused by the velocity of satellites relative to Earth. Let me introduce the Lorentz transformation in the theory of special relativity first

$$T = \frac{T_0}{\sqrt{1 - \frac{v^2}{c^2}}}$$

and also, the Lorentz factor.

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Suppose that the spots on Earth are all stationary, which means the velocity of the spot needing positioning is 0. Thus, the γ for the receivers, γ_r, is actually 1.

$$\gamma_r = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{1}{\sqrt{1 - \frac{0}{c^2}}} = 1$$

Now in the equation, T is the time that is experienced by satellites and T_0 is the time experienced by receivers on Earth.

For satellites, the velocity is v_s, which will be 3874 $^m/_s$ in the calculation as the average speed of satellite. Thus, comparing to the speed of light c, v_s seems to be a small quantity. So that $\frac{{v_s}^2}{c^2}$ is a small quantity.

When x is a small quantity,

$$\sqrt{1-x} \approx 1 - \frac{x}{2}$$

So,

$$\sqrt{1 - \frac{{v_s}^2}{c^2}} \approx 1 - \frac{{v_s}^2}{2c^2}$$

Thus, the Lorentz factor of the satellites γ_s will be:

$$\gamma_s = \frac{1}{\sqrt{1 - \frac{{v_s}^2}{c^2}}} \approx \frac{1}{1 - \frac{{v_s}^2}{2c^2}} = \frac{1}{1 - \frac{3874^2}{2\times(2.998\times10^8)}} = 1 - 8.349\times10^{-11}$$

As a result, when T_0 is a day, the time error, ΔT_1, is

$$\Delta T_1 = T - T_0 = T_0 \cdot (\gamma_s - \gamma_r) = 1\, day \cdot (1 - 8.349\times10^{-11} - 1) = -7214\ (ns)$$

After that, the theory of general relativity kicks in with the effect of the gravitational frequency shift being far greater than the 7 microseconds per day delay due to the velocity relative to Earth. But how much is the effect?

Define the radius of Earth as R_E and the radius of satellite orbit (roughly seen as a circle) as R_S. Also, define the mass of Earth as M_E. And time error will be ΔT_2 here. Here is the calculation:

According to the theory of general relativity, a clock which is closer to an object will be slower than a clock further away, the atomic clocks on board the GPS's are faster. Some of the parameters used above in the calculation of special relativity will be used again here; they refer to the same meaning but not the same value.

When T_0 is one day

$$\Delta T_2 \approx T_0 \cdot \left(\frac{1}{\gamma_E} - \frac{1}{\gamma_S}\right) = T_0 \cdot \left(\frac{G \cdot M_E}{R_E \cdot c^2} - \frac{G \cdot M_E}{R_S \cdot c^2}\right)$$

After using the data of G (gravitational constant) which is $6.67408 \times 10^{11}\, {m^3}/{(kg \cdot s^2)}$, R_E and R_S and the speed of light c, we now figure out:

$$\Delta T_2 \approx 45850\, ns$$

Combining the 7 microsecond a day delay due to the satellites velocity relative to Earth and the GPS's being further away from Earth, the final clock error, ΔT adds up to a 38 microseconds delay,

$$\Delta T \approx 45850 + (-7214) = 38636\, ns$$

which if left uncorrected would translate through to a 10 km/day error.

$$\Delta S = \mathrm{c} \cdot \Delta T \approx 38636\, (ns) \cdot 2.998{\times}10^8 ({}^{m}/_{s}) = 11.58\, k$$

GPS's invalid if this was not to be taken into consideration. So, this is compensated by the GPS's clocks frequencies being slightly slowed down from 10.23MHz to 10.22999999543 MHz to cancel out the effects of relativity. [12]

Conclusion

In this paper, we talk about what a satellite positioning system consists of, how satellites can position a spot of Earth, and what factors may actually affect the accuracy of satellite positioning. GPS technology is really a fascinating invention and makes a lot of contribution to human society. It is amazing that even a tiny GPS equipment in our pocket contains much knowledge of math and physics. With more and more application of these knowledge, the life of human beings will definitely be much better.

Bibliography

[1] "What is a GPS? How does it work?" What is GPS? Everyday Mysteries. N.p., n.d. Web. 21 July 2017.

[2] "The Global Positioning System." GPS.gov: GPS Overview. N.p., n.d. Web. 21 July 2017.

[3] "Autocorrelation." Wikipedia. Wikimedia Foundation, 12 July 2017. Web. 21 July 2017.

[4] GPS 定位基本原理. N.p.: n.p., n.d. PDF.

ftp://ftp.csr.utexas.edu/pub/ggfc/seogc/teacher_slides_pdf/Dong_GPS_PositioningTheory.pdf

[5] Kaplan, Elliott D., and C. Hegarty. Understanding GPS: principles and applications. Norwood: Artech House, 2006. Print.

[6] "电离层." 到百科首页. N.p., n.d. Web. 21 July 2017.

"对流层." 到百科首页. N.p., n.d. Web. 21 July 2017.

[7] "Tropospheric Delay." Tropospheric Delay - Navipedia. N.p., n.d. Web. 21 July 2017.

[8] GPS error sources. N.p.: n.p., n.d. PDF

http://www.gislab.ar.wroc.pl/server.ftp?FileID=1522&GISLabSessionID=yfbxkwuym

[9] [10] https://www.fhwa.dot.gov/publications/publicroads/95fall/images/p95a5b.gif

[11] "Atomic clock." Wikipedia. Wikimedia Foundation, 21 July 2017. Web. 21 July 2017

[12] SciBRIGHT. YouTube. YouTube, 14 Sept. 2016. Web. 21 July 2017.