

Capítulo 2 Introdução ao Sistema GPS

O primeiro satélite artificial colocado em órbita foi o russo Sputnik 1, lançado no ano 1957. Assim, foi demonstrada a viabilidade dos satélites artificiais. Cientistas do Laboratório de Física Aplicada (APL) da Universidade John Hopkins dos EUA monitoraram a sinal de rádio transmitido pelo Sputnik 1, notificando que o sinal do satélite tinha deslocamento Doppler devido ao movimento relativo do mesmo [1]. Seguindo esta observação, Dr. Frank T. McClure (APL) percebeu que, se a órbita do satélite é conhecida, então a localização de um usuário pode ser determinada com base na medida do deslocamento Doppler.

Utilizando este conceito, o APL propôs e desenvolveu um modelo de navegação por satélite com base no efeito Doppler. Em abril de 1960, o primeiro protótipo foi lançado exitosamente para analisar a funcionalidade. O sistema foi chamado “Navy Navigation Satellite System”, também conhecido como “Transit” [2].

O Global Positioning System (GPS) é um sistema de navegação desenvolvido pelo Departamento de Defesa (DoD) dos EUA no início da década de 1970 e foi visto como substituição do sistema “Transit”. Inicialmente, GPS só foi utilizado para fins militares, mas na atualidade o sistema pode ser acessado por usuários militares e civis, oferecendo posicionamento contínuo e sincronismo em qualquer lugar no planeta.

2.1 Características do Sistema GPS

O sistema GPS consiste de uma constelação composta inicialmente de 24 satélites operacionais [1], completada em Julho de 1993. Com a finalidade de uma cobertura contínua no planeta inteiro, os satélites GPS estavam dispostos em seis planos orbitais, cada um contendo quatro satélites (Figura 2.1). Com este tipo de distribuição, de quatro até 10 satélites GPS podem ser visíveis em qualquer lugar do planeta, se uma elevação mínima de 10 graus é considerada. Atualmente, 30 satélites GPS encontram-se em operação [3].

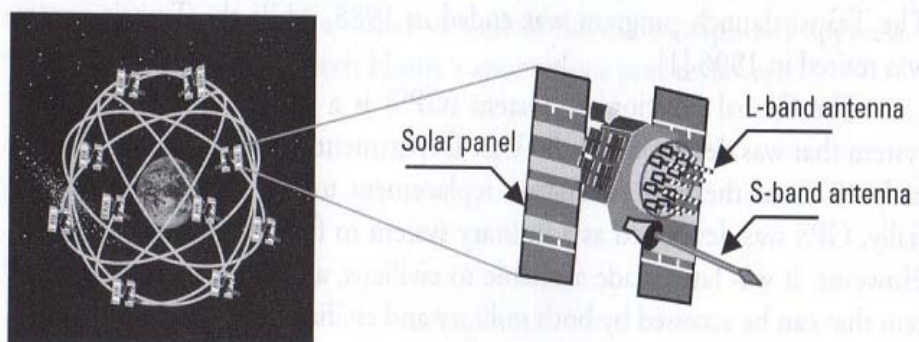


Figura 2.1- Constelação do sistema GPS (reproduzida da referência [1])

As órbitas dos satélites GPS são praticamente circulares, (uma forma elíptica com uma excentricidade máxima de, aproximadamente, 0,01) e com uma inclinação ao redor de 55 graus com o equador. O semi-eixo maior de uma órbita GPS é 26 560 km [1].

O período orbital correspondente para GPS é aproximadamente igual a 12 horas siderais (isto é, 11 horas e 58 minutos).

2.2 Segmentos do GPS

O GPS consiste de três segmentos: espacial, de controle e de usuário (Figura 2.2) [1]. O segmento espacial contempla a constelação de satélites mencionados anteriormente. Cada satélite GPS transmite um sinal que tem um número de componentes: duas ondas senoidais, também chamadas de frequências portadoras, dois códigos digitais e uma mensagem de navegação.

Os códigos e a mensagem de navegação são aderidos às portadoras como uma modulação binária de duas fases. As portadoras e os códigos são principalmente usados para determinar a distância entre o usuário receptor e os satélites GPS. A mensagem de navegação tem, além de outras informações, as coordenadas (localização) dos satélites em função do tempo.

O segmento de controle do sistema GPS consiste de uma rede de estações de rastreamento distribuídas globalmente, com uma estação central localizada em Base da Força Aérea dos EUA. O segmento de usuário inclui todos os usuários civis e militares, que podem receber sinais GPS usados para determinar suas posições em qualquer lugar do planeta.

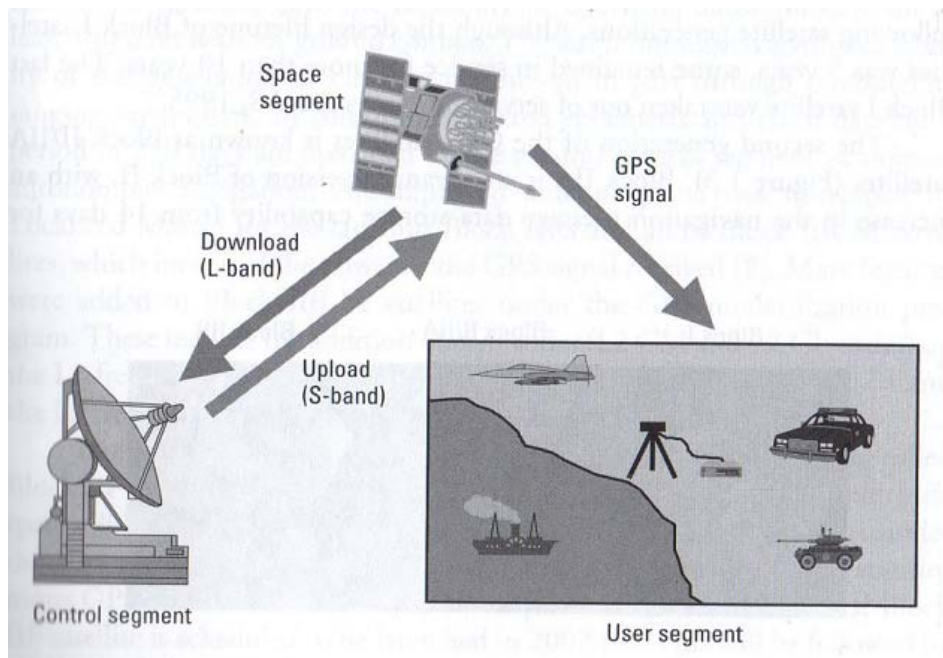


Figura 2.2- Segmentos do sistema GPS (reproduzida da referência [1])

2.3 GPS: A idéia básica

A idéia por trás do sistema GPS é simples. Se as distâncias de um ponto na Terra onde se encontra o receptor GPS a três satélites GPS são conhecidas, juntamente com as posições dos satélites, é possível determinar a localização do ponto (Receptor GPS).

Cada satélite transmite continuamente um sinal de microondas composto por duas portadoras, dois códigos e uma mensagem de navegação. Então, quando o receptor GPS é ligado, o sinal do GPS é recebido pela antena receptora, adquirido e processado. O sinal processado fornece pseudodistâncias (distâncias aproximadas) aos satélites GPS e as coordenadas dos satélites a partir de mensagens de navegação. A pseudodistância é definida como o produto da velocidade da luz no vácuo pelo retardo correspondente à máxima correlação entre o sinal recebido pelo receptor e uma replica idêntica gerada por ele. A pseudodistância, conforme indicado na discussão mais detalhada da seção 2.5, é afetada tanto pelos equipamentos utilizados quanto pelos meios nos quais os sinais GPS se propagam.

Teoricamente, só as distâncias a três satélites rastreados simultaneamente seriam necessárias para determinar a posição do receptor. Neste caso, o receptor poderia ser localizado pela interseção de três esferas, cada uma centrada em um satélite e com um raio equivalente à distância entre o receptor e satélite (Figura 2.3). Entretanto, um quarto satélite é necessário, para considerar o deslocamento de tempo (“offset”) do relógio do receptor [1]-[4].

A aproximação que se tem com este método foi limitada a 100 m na componente horizontal, 156 m para a componente vertical e 340 nanossegundos para o componente do tempo até o dia 1 de maio de 2000, tendo todos estes valores a probabilidade de 95 %. Este baixo nível de aproximação foi devido aos efeitos da “SA”, uma técnica utilizada para reduzir o tempo real autônomo de posicionamento para usuários não autorizados, por motivos de segurança. Entretanto, foram feitos muitos estudos que levaram o Governo dos EUA a desativar a “SA” no dia 2 de maio do ano 2000. Com o efeito “SA” reduzido a zero, o sistema GPS melhorou sua aproximação por um fator de sete ou mais.

Para melhores aproximações da posição GPS, se utilizou os métodos de posição relativa ou diferencial, onde são usados dois receptores simultaneamente rastreados pelos mesmos satélites GPS.

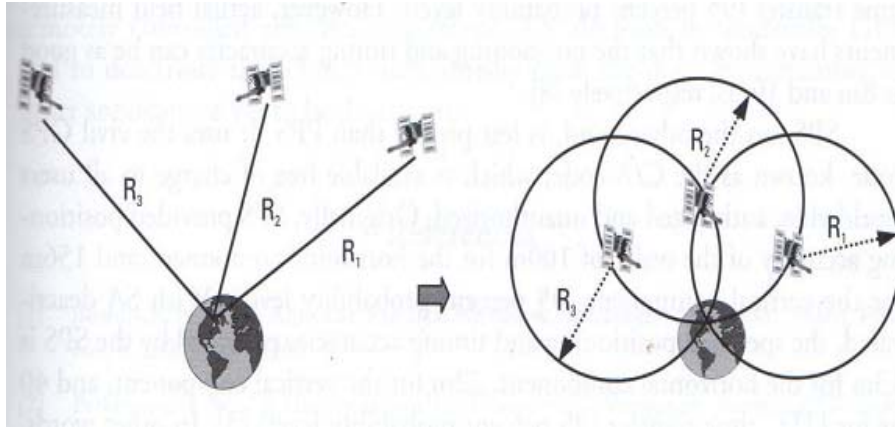


Figura 2.3- Idéia básica de posicionamento do sistema GPS (reproduzida da referência [1])

2.4 Estrutura do sinal GPS

Cada satélite GPS transmite duas ondas portadoras chamadas L1 e L2. As ondas são geradas a partir de uma frequência fundamental de 10,23 MHz, a qual é multiplicada por 154 e 120, respectivamente.

Os valores associados às ondas portadoras são:

- L1 = 1575,42 MHz de frequência e 19 cm de comprimento de onda;
- L2 = 1227,60 MHz de frequência e 24 cm de comprimento de onda [5].

A disponibilidade das frequências portadoras permite uma melhor correção do erro de GPS conhecido como retardo ionosférico.

Todos os satélites GPS transmitem as mesmas frequências portadoras L1 e L2, embora o código de modulação seja diferente para cada satélite, o qual minimiza as interferências do sinal.

Os satélites GPS transmitem dois códigos conhecidos como “coarse acquisition” (C/A-code) e de precisão (P-code). O código C/A é modulado só sobre a portadora L1, mas o código P é modulado em ambas portadoras [1].

Sistemas GPS modernos transmitem dois códigos adicionais, denominados L2 “civil-moderate” (L2 CM) e L2 “civil-long” (L2 CL). Estes códigos são comumente chamados de códigos PRN porque se assemelham a sinais aleatórios, mas são gerados usando algoritmos matemáticos. A modulação do sinal GPS é chamada modulação bifase, porque a fase da portadora é deslocada de 180 graus quando o valor do código muda de zero para um e de um para zero [1].

2.5 Pseudodistância e cálculo de posição

O valor da distância entre o satélite GPS e o receptor é obtido mediante a correlação entre o sinal recebido pelo receptor e uma réplica idêntica gerada por ele. Se os relógios tanto do satélite quanto do receptor estivessem sincronizados e a propagação ocorresse no vácuo, o produto do retardo correspondente à máxima correlação pela velocidade da luz produziria a distância entre o usuário e o satélite. Entretanto, não existe sincronismo perfeito entre os relógios dos satélites e dos receptores e a propagação ocorre na atmosfera, de modo que a distância assim determinada é errônea, comumente denominada de pseudodistância. Os erros relacionados são conhecidos como erros de distância equivalente do usuário, ou “User Equivalent Range Error” (UERE) [6].

Outros fatores que também influenciam a pseudodistância são:

- Atrasos atmosféricos (ionosfera e troposfera);
- Erro de efemérides (intencional, SA, ou não);
- Erros relativísticos (rotação da Terra);
- Erros devidos à resolução e ao ruído do receptor;
- Múltiplos percursos (sinais refletidos recebidos pelo receptor);
- Perdas de ciclo;
- Erro de bias ou área (receptores e configuração dos satélites) [4].

Com base na figura 2.4, queremos determinar o vetor $\vec{u} = (x_u, y_u, z_u)$ que representa a posição do usuário no sistema de coordenadas ECEF.

O vetor $\vec{s} = (x_s, y_s, z_s)$ representa a posição do satélite no mesmo sistema de coordenadas. O vetor $\vec{r} = \vec{s} - \vec{u}$ representa o vetor posição do satélite em relação ao usuário.

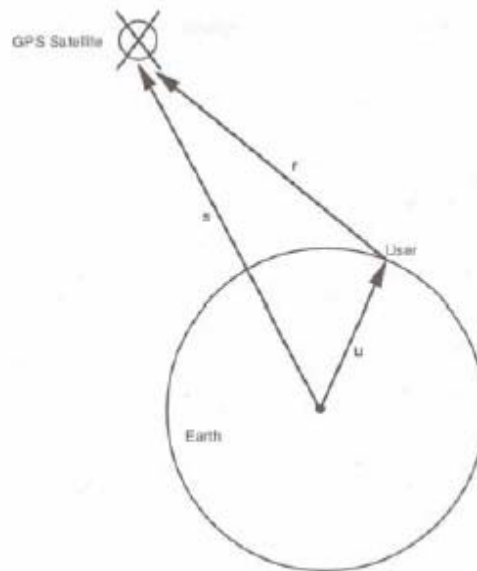


Figura 2.4 Representação dos vetores posição do usuário e do satélite (Reproduzida da referência [4])

O valor de $\|\vec{r}\|$ é calculado por medição do tempo de propagação que o código gerado no satélite utiliza entre o satélite e o receptor.

O relógio do receptor pode ter um erro de polarização (bias) do sistema. Adicionalmente, a geração de frequência do satélite e o tempo, embora baseados em um relógio atômico de césio ou rubídio, também é deslocado do tempo do sistema.

Assim, a pseudodistância ρ determinada pelo processo de correlação contém a distância geométrica entre o satélite e o usuário, a diferença entre os tempos do sistema e do relógio do usuário, e o deslocamento entre os tempos do sistema e do relógio do satélite [4].

Pseudodistância:

$$\begin{aligned}\rho &= c(T_u' + t_u) - (T_s + \delta t) \\ &= c(T_u' - T_s) + c(t_u - \delta t) = c(T_u + \delta t_D - T_s) + c(t_u - \delta t) \\ &= r + c(t_u - \delta t + \delta t_D)\end{aligned}\quad (2.2)$$

Substituindo em $\|\vec{r}\| = \|\vec{s} - \vec{u}\|$

$$r = \|\vec{s} - \vec{u}\| = \rho - c(t_u - \delta t + \delta t_D) \quad (2.3)$$

2.6 Cálculo da posição do usuário

Com o objetivo de calcular a posição do usuário em três dimensões (x_u , y_u , z_u) e o deslocamento t_u , medidas de pseudodistâncias são realizadas para quatro satélites, resultando no seguinte sistema de equações ($j = 1, \dots, 4$):

$$\rho_j - \|s_j - u\| = ct_u \quad (2.4)$$

Tomando os valores para j entre um e quatro, o sistema de equações (2.4) pode ser reescrito na forma:

$$\begin{aligned}\rho_1 &= \sqrt{(x_1 - x_u)^2 + (y_1 - y_u)^2 + (z_1 - z_u)^2} + ct_u \\ \rho_2 &= \sqrt{(x_2 - x_u)^2 + (y_2 - y_u)^2 + (z_2 - z_u)^2} + ct_u \\ \rho_3 &= \sqrt{(x_3 - x_u)^2 + (y_3 - y_u)^2 + (z_3 - z_u)^2} + ct_u \\ \rho_4 &= \sqrt{(x_4 - x_u)^2 + (y_4 - y_u)^2 + (z_4 - z_u)^2} + ct_u\end{aligned}\quad (2.5)$$

As quatro equações formam um sistema não linear que pode ser resolvido analiticamente ou usando técnicas iterativas baseadas em linearização ou filtragem de Kalman.

De acordo com o segundo método de solução, inicialmente a posição (x_u, y_u, z_u) é aproximada por $(\hat{x}_u, \hat{y}_u, \hat{z}_u)$, com deslocamento $(\Delta x_u, \Delta y_u, \Delta z_u)$. Por consequência, o sistema de equações (2.3) é expandido numa serie de Taylor, tornando-o linear, Desta forma, se tem:

$$\hat{\rho}_j - \rho_j \approx \frac{x_j - \hat{x}_u}{\hat{r}_j} \Delta x_u + \frac{y_j - \hat{y}_u}{\hat{r}_j} \Delta y_u + \frac{z_j - \hat{z}_u}{\hat{r}_j} \Delta z_u + \Delta c t_u \quad (2.6)$$

sendo j ($j = 1, \dots, 4$) correspondente a cada satélite. Reescrevendo a equação (2.6), temos:

$$\Delta \rho_j \approx a_{xj} \Delta x_u + a_{yj} \Delta y_u + a_{zj} \Delta z_u - c \Delta t_u \quad (2.7)$$

onde:

$$\Delta \rho_j = \hat{\rho}_j - \rho_j$$

$$\Delta x_u = x_u - \hat{x}_u$$

$$\Delta y_u = y_u - \hat{y}_u$$

$$\Delta z_u = z_u - \hat{z}_u$$

$$\Delta t_u = t_u - \hat{t}_u$$

$$a_{xj} = \frac{x_j - \hat{x}_u}{\hat{r}_j}$$

$$a_{yj} = \frac{y_j - \hat{y}_u}{\hat{r}_j}$$

$$a_{zj} = \frac{z_j - \hat{z}_u}{\hat{r}_j}$$

$$\hat{r}_j = \sqrt{(x_j - \hat{x}_u)^2 + (y_j - \hat{y}_u)^2 + (z_j - \hat{z}_u)^2}$$

A equação (2.7) pode ser representada de forma matricial

$$\Delta\rho = H\Delta x \quad (2.8)$$

onde:

$$\Delta\rho = \begin{bmatrix} \Delta\rho_1 \\ \Delta\rho_2 \\ \Delta\rho_3 \\ \Delta\rho_4 \end{bmatrix} \quad H = \begin{bmatrix} a_{x1} & a_{y1} & a_{z1} & 1 \\ a_{x2} & a_{y2} & a_{z2} & 1 \\ a_{x3} & a_{y3} & a_{z3} & 1 \\ a_{x4} & a_{y4} & a_{z4} & 1 \end{bmatrix} \quad \Delta x = \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \\ \Delta x_4 \end{bmatrix}$$

Finalmente, a solução da equação (2.7), se H^{-1} existir, é obtida:

$$\Delta x = H^{-1}\Delta\rho \quad (2.9)$$