B.2 使用数据包结构图

我们在第 1 章中学习到，数据包是按照协议规定的方式排列的数据。因为通用协议按照指定的规则排列包数据，所以软硬件能够解析这些数据；数据包必须遵守明确的格式规则。使用数据包结构图，我们能够识别格式并解析数据包。一个数据包结构图是数据包的图形化表现方式，使分析人员能够将任意给定协议的数据包从原始十六进制字节映射至具体字段。结构图从协议的 RFC 文档中提取，显示了协议中各字段的长度和排列顺序。

让我们查看第 6 章中的 IPv4 的数据包结构图（见图 B-3）。

<table>
<tr><td colspan="8">互联网协议 4 (IPv4)</td></tr>
<tr><td>偏移位</td><td>八位组</td><td colspan="2">0</td><td colspan="2">1</td><td colspan="2">2</td><td>3</td></tr>
<tr><td>八位组</td><td>位</td><td>0–3</td><td>4–7</td><td colspan="2">8–15</td><td>16–18</td><td>19–23</td><td>24–31</td></tr>
<tr><td>0</td><td>0</td><td>版本号</td><td>首部长度</td><td colspan="2">服务类型</td><td colspan="3">总长度</td></tr>
<tr><td>4</td><td>32</td><td colspan="3">标识符</td><td>标识</td><td colspan="3">分片偏移</td></tr>
<tr><td>8</td><td>64</td><td colspan="2">存活时间</td><td colspan="2">协议</td><td colspan="3">首部校验和</td></tr>
<tr><td>12</td><td>96</td><td colspan="7">源IP地址</td></tr>
<tr><td>16</td><td>128</td><td colspan="7">目的IP地址</td></tr>
<tr><td>20</td><td>160</td><td colspan="7">选项</td></tr>
<tr><td>24+</td><td>192+</td><td colspan="7">数据</td></tr>
</table>

图 B-3　IPv4 的数据包结构图

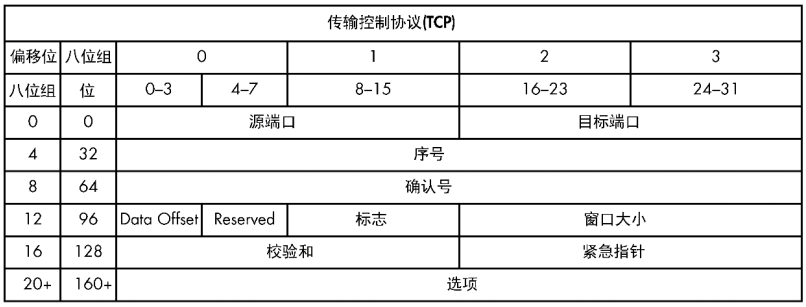
在这个结构图中横轴表示 0～31 的二进制位。换算成字节，为 0～3。纵轴也按照位和字节进行标记，每行为一个 32 位（或 4 字节）片段。我们使用数轴来计算字段地址的偏移量，根据纵轴确定字段位于哪个 4 字节片段，然后根据横轴确定给字段以字节为单位的偏移量。第一行由前四个字节组成，0～3，在横轴上进行标记。第二行由之后的四字节组成，4～7，同样在横轴上进行计数。我们从字节 4 开始计数，对应的横轴刻度为 0；下一个字节是字节 5，对应的横轴刻度为 1；以此类推。

例如，我们查看 IPv4 的结构图，0x01 字节是服务类型字段。计算方式：在纵轴上看，服务类型字段位于第一行，对应纵轴刻度为 0，所以从 0 开始计数；在横轴上看，该字段位于刻度 1；所以该字段位于，从 0 开始且偏移量为 1 字节的位置，即位于 0x01 字节。

再查看另一个例子，0x08 字节是存活时间字段。计算方式：在纵轴上看，存活时间字段位于第三行，对应纵轴刻度为 8，所以从 8 开始计数；在横轴上看，该字段位于刻度 0；所以该字段位于，从 8 开始，偏移量为 0 字节的位置，即位于 0x08 字节。

有些字段，如源 IP 字段，长度为多个字节，我们在图中看到，位于 0x12:4。其他一些字段只占用了半字节，如 0x00 字节的高位字节为版本字段，低位字节为 IP 头长度字段。0x06 字节的粒度更细，每一位都表示一个字段。当字段的值为单个二进制数值时，它通常是一个标记（flag）。如 IPv4 头中的翻转（Reversed）、不分片（Don’t Fragment）和多个分片（More Fragments）字段。标记的值为一个一位的二进制数值，1（true）或 0（false），所以当标记值为 1，表示标记生效。标记生效的具体含义根据协议和字段而定。

让我们在图 B-4 中查看另一个例子（这个结构图在第 6 章出现过）。



传输控制协议(TCP)						
偏移位	八位组	0		1	2	3
八位组	位	0–3	4–7	8–15	16–23	24–31
0	0	源端口			目标端口	
4	32	序号				
8	64	确认号				
12	96	Data Offset	Reserved	标志	窗口大小	
16	128	校验和			紧急指针	
20+	160+	选项				

图 B-4　TCP 的数据包结构图

这个图片展示了 TCP 协议头。根据这张图，我们能够在不知道 TCP 用途的情况下，回答很多与 TCP 数据包有关的问题。假设一个 TCP 数据包的协议头如下方的十六进制数据所示：

```
   0646 0050 7c23 5ab7 0000 0000 8002 2000
0b30 0000 0204 05b4 0103 0302 0101 0402
```

使用包结构图，我们能够定位和解析特定的字段。如我们能够发现以下信息。

- 源端口号位于 0x00:2 字节（0x00 至 0x01），十六进制值为 0646（十进制：1606）。
- 目的端口号位于 0x02:2 字节（0x02 至 0x03），十六进制值为 0050（十进制：80）。
- 数据偏移量字段表示协议头长度，位于 0x12 字节的高位字节，十六进制值为 8。

让我们将这些知识用来分析这个神秘的数据包。