Base64编码解码算是网络安全领域的很小而且很简单的一个知识点了，虽然简单但是应用场景却极其广泛，可以这样说，开发者要是不懂Base64，基本上可以告别程序员生涯了。这篇文章从原理入手，并给出java的Base64实现。方便你掌握Base64。

一、Base64由来

很早之前，电子邮件刚刚问世，那时候消息的传递都是英文，后来中国开通了互联网之后，对邮件的使用量也大量增加，这时候电子邮件就有了中文的需求。但是中文在传输的时候不能被有效地处理，这时候Base就出来了，Base64通过对这些中文进行编码，转化为服务器和网关能够识别的数据。这时候就能够使用电子邮件有效地传输了。

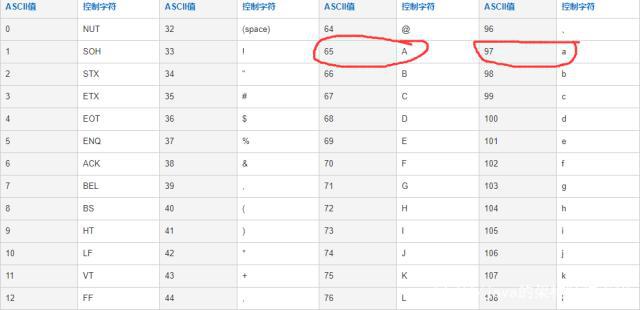
上面的这个例子不是说Base64专门为电子邮件而生的，从其诞生之初，就开始在各大领域有了广泛的应用。比如说网络上传递图片，我们可以Base64先对图片进行处理，然后就可以有效的传输了。

OK，我们大致知道其用途，然后我们就好好的分析一下他的原理，到底是如何对这些数据进行编码的。

二、Base64原理

1、Base64编码

Base64的原理超级简单，相信我们都知道ASCII 编码，从A-Z、a-z、0-9和一些其他的特殊字符，这些字符都有唯一的一个数字来表示。比如说a是97，A是65。我们来截取一部分图看一下：



同理Base64也有这样一套编码。范围是”A-Z“、”a-z“、”0-9“、”+“、”/“一共64个字符。我们给出一个表格来看一下，这个比ASCII编码要简单多了，只有64个。



由于索引是从0开始，所以最后的索引是63。在编码的时候Base64就是通过上面的进行转换编码的。下面我们就来看看Base64编码的原理。

2、基本原理

比如说有一封邮件，我们想要对其使用Base64进行编码。怎么办呢？基本步骤如下：

（1）对邮件的数据进行切分，每三个字节一组，一共24个bit。

（2）对切分后的数据重组，24个bit重组为4组，每组6个bit。

（3）对重组后的数据处理，每组最前面添加两个“0”，构成每组8个bit。此时一共32个bit。

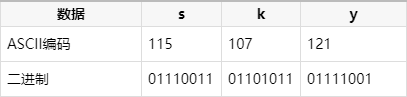
（4）根据Base64编码表，获取相应的编码值。

此时一封完整的邮件，被切分重组处理之后就变成了Base64编码了。基本原理其实很简单。不过你不理解也没关系，我们直接上个实例来解释一下。

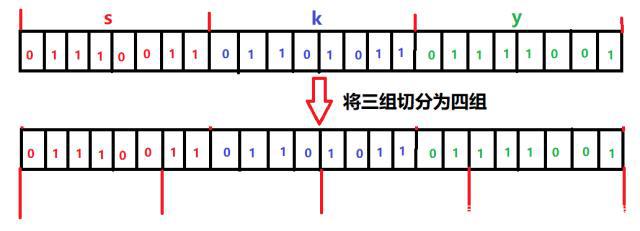
3、实例验证

比如说电子邮件里面出现了三个字母sky。我们要对这个三个字符使用Base64进行编码。

（1）对邮件的数据进行切分，每三个字节一组，一共24个bit

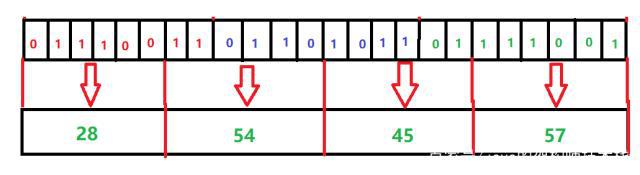


（2）对切分后的数据重组，24个bit重组为4组，每组6个bit



（3）对重组后的数据处理，每组最前面添加两个“0”，构成每组8个bit。由于在最前面添加的0，所以对数值不构成影响。

（4）根据Base64编码表，获取相应的编码值



（5）完成编码的转换



到这我们基本上就是实现了Base64编码机制从sky到c2t5的转换。

有些地方需要我们去注意一下：

（1）在第三步中，最前面添加了两个0，所以最终编码之后要比之前多出三分之一的大小。

（2）上面的例子中，我们使用的是ASCII编码，但是如果我们使用UTF-8，对应Base64编码的结果是不一样的。

（3）Base64只是进行了编码，方便数据的传输而已。这可不是加密。

原理也搞清楚了，现在我们就实现一下。

三、代码实现

你可以自己去实现一个编码解码的完整过程，但是java已经为我们封装好了，我们直接只用别人造好的轮子多好。不管是自己写还是使用别人的，原理搞清楚就OK了。



是不是很简单。其实java实现的方式有很多种，其他的还有Commons Codec和Bouncy Castle。实现的过程和JDK提供的类似，我们只需要导入相应的jar包即可。

Base64算法的应用场景有E-Mail、密钥、证书文件等等。这也只是入门，想要深入了解，可以看一些安全有关的书籍，不过很多都和数学有关，看的实在是难受，曾经看过，可惜没坚持住。惭愧，书到用时方恨少。