目前这个需求说明涉及的功能点很多，还需要大量相关协议的背景知识。如果是实验室做的话，在不考虑相关理论知识的学习和开发顺利的情况下，如果是两个人做至少需要三个月。

开发一个wireshark离线包分析工具的具体需求。

=============分隔符，下面是需求=========================

（基于wireshark定制脚本的实现方式，红色字体标注部分需要重点做些前期方案的调研和设计，作为关键功能实现方案需要讨论需求）

一、针对用户行为的分析：

* 功能点分散，一共8个功能点，特别是第二条描述的几个功能点，相互能复用的代码不多，都得分别处理。两个人做，全职去做，可能需要两个星期左右。
* 基于IP的在线用户统计；终端平台类型的统计；浏览器类型的统计；基于URL关键字段匹配的统计，以及相关过滤条件；
* 在线用户个数随时间变化；用户平均在线时长随时间变化；用户GET请求到达率随时间变化；TCP会话达到率随时间变化。

二、针对HTTP业务流程的分析：

* 两个人全职做，如果顺利的话，第一条在一个星期到两个星期之间，后四条在两个星期到三个星期之间。
* 区分针对单个网页的多个Get请求及其TCP连接会话，作为一个网页访问session，统计从第一个GET请求到最后一个200响应达到的时间差异。
* 针对网页访问session为单位，进行SP响应时延（从单个get请求到达到对应200响应回复）的统计、用户访问次数、上下行流量、请求到达速率随时间变化的统计，以及相关过滤条件。
* 识别浏览器的pipelining和multipart优化机制使用占比、相关参数（pipelining并发度、调度方式）的统计，结合终端平台类型和浏览器类型进行关联统计，以及相关过滤条件。
* 识别浏览器的并发TCP连接个数，pipelining并发度、multipart并发度。
* 识别网页session资源的分布模型（文本、图片、视频的种类统计、个数统计、大小统计），及其过滤条件。

以上四条功能点很多，目测十多个功能点，会需要一段时间的学习，且都需要在第一条能做的基础上进行数据的统计和分析。所以难点还在于，第一条的处理机制是否方便后四条进行数据的统计和分析，如果不方便，那么工作量会加大。

三、针对流量模型的分析：

* 针对上述过滤条件的上下行流量使用和瞬时速率统计。
* 针对TCP会话的持续时间，全部数据和有效数据（去除连接建立和连接拆除）流量使用、瞬时速率和时间占用比例统计。
* 以下两条大约也需要两个人全职做一周
* 针对网页访问session的持续时间，（上下行）全部数据和有效数据（去除协议头）流量使用、瞬时速率和时间占用比例统计。
* 针对协议类型（应用层和传输层）的（上下行）流量使用、瞬时速率统计。

四、针对拥塞控制的分析：

* 针对PCN功能标识的统计（哪些终端支持PCN拥塞控制，是否存在实际使用的PCN拥塞标识）；
* 针对TCP初始窗口大小；
* 针对TCP会话的丢包率、RTT、数据重传概率、重传流量占比的统计；
* 针对链路层的包错误率、单跳时延、帧重传概率、重传流量占比的统计。

**【我们的分析： 下面这个是非功能性需求，应该可以满足！】**

四、实现要求：

* 基于wireshark定制脚本方式，提供源码；
* 所有统计要求，提供平均数值、最小值、最大值、Mean值的统计表格（excel表格）和图形化界面；
* 部分统计要求，提供间隔10%为步长的概率分布统计表格（excel表格）和图形化界面；
* 确保提供处理海量数据的能力，或者提供数据分段整合的解决方案。