

开始

初始化字典，键值对为  
<客户端IP: 操作系统信息, 浏览器信息, 初始TTL值>

对于数据集中的每一条HTTP流，  
计算客户端的初始TTL值，  
解析请求报文中的User-Agent字  
段，并添加到字典中

字典创建完成后，  
依次查询字典中的所有条目

对于每一个客  
户端IP，主机  
属性信息是否  
唯一？

否

是

初始TTL值  
是否唯一？

否

从字典中删除该键值对

是

否

字典是否遍历  
完成？

是

得到清洗后的主机属性标签字典

结束

