链滴

httpclient 的陷阱

作者：mrkid
原文链接：https://ld246.com/article/1366541892260
来源网站：链滴
许可协议：署名-相同方式共享 4.0 国际 (CC BY-SA 4.0)

      最近搭了一个账号系统，并且提供了接口给各个子系统进行访问。

      子系统使用api进行访问账号系统来获取用户的详细信息, 使用账号系生成的 remember-me cookie 进行区分。
      测试了几下，发现串号了，请求B用户的信息，结果返回了A。 这是怎回事呢。
      追查了一下，原因如下:
      子系统使用httpclient同账号系统交互，复用了同一个httpclient， 使 MultiThreadedHttpConnectionManager 就以为万事大吉了，实则不然。
       A. 子系统   ------ user1.remember-me cookie ------>   账号系统
            子系统 <----- user1.info & JESSIONID ----------账号系统
       子系统向账号系统发起了请求，并携带了user1的 remember-me cookie， 账号系统使用remember-me cookie 进行了自动登录，并返回 登录者的详细内容因为效率等其它原因，启用了session，把用户信息保存在对应的session中， 并在响应头里 携带了 se cookie: jessionid = xxx。
       httpclient自动将这个jessionid cookie保存了下来。
       B.   子系统   ------ user2.remember-me cookie & JESSIONID ------>   账号系统
             子系统 <----- user1.info ---------账号系统
       httpclient又向账号系统发起了请求，请求user2的信息， httpcient自动将前一个请求里返回的JSESSIONID 也添加到请求头。 账号系统一看，哦，有JSESSIONID, 直接将 JSESSIONID 对应的的session中的用户信息返回给了子系统，因此就串号了。
       问同事，为何当初要在把 httpclient 的 COOKIE_POLIC 设为 BROWSER_COMPATIBILITY， 答曰，这段代码原来是用在爬虫里的， 部分网爬虫需要先登录再爬取，因此使用BROWSER_COMPATIBILITY。
       爬虫明显同现在的子系统不同，在不同场景下需要选用合适的cokiepolicy，问题最后通过设置 httpclient 的 COOKIE_POLICY为 IGNORE_COOKIES 得以决。